POST-TRANSCRIPTIONAL INHIBITION OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1) USING COMBINATORIAL RNA INTERFERENCE (RNAI) EXPRESSION VECTORS

SAMANTHA BARICHIEVY

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Doctor of Philosophy

Johannesburg, October 2009

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

Signature ----------

Date ----------

Abstract

Recent estimates indicate that globally there are over 33 million people infected with Human Immunodeficiency Virus type 1 (HIV-1). The epidemic is particularly severe in sub-Saharan Africa which accounts for 67% of all infected individuals and 72% of AIDS deaths in 2007. While current therapies, particularly in combination as a cocktail of highly active antiretroviral therapy (HAART), have had an important positive impact on the morbidity and mortality of HIV-related illness, there remain significant limitations. These include toxicities, resistance and the inability to eradicate a latent infection. In addition, most therapeutic agents have been developed to target HIV-1 subtype B, which affects individuals predominantly in Western Europe and North America. These concerns, along with ensuring patient compliance with treatment and the high cost of improved treatment regimens, have prompted the search for innovative and globally-effective therapies to eradicate HIV infection. More recently, gene therapy strategies based on the naturally-occurring RNA interference (RNAi) pathway has provided an exciting new mechanism to inhibit rogue gene expression. RNAi represents a set of highly conserved cellular pathways whereby double-stranded RNA (dsRNA) precursors are processed into shorter dsRNAs by the successive action of ribonucleases Drosha and Dicer. Processed, 21-23 nucleotide, short interfering RNAs (siRNAs) or antisense RNAs (asRNAs), associate with members of the Argonaute family of proteins to regulate gene expression at the transcriptional and post-transcriptional level.

By exploiting the biogenesis of the endogenous mammalian RNAi pathway, several exogenous RNAi pathway mimics have been developed to inhibit unique sequences, including viral targets such as HIV. In the context of HIV, a combinatorial system that allows for the simultaneous suppression of multiple targets is important in preventing viral mutational escape of this rapidly evolving pathogen. The studies presented in this thesis add significantly to the newly emerging body of research on combinatorial RNAi strategies by focusing on the two novel technological approaches using mammalian expression systems. Both RNA Pol III-generated long-hairpin RNAs (lhRNAs) and RNA Pol II-generated polycistronic primary microRNAs (pri-miRNAs) were developed as systems for generating combinatorial RNAi precursors from single transcriptional units that induce post-transcriptional silencing of several highly conserved sequences of HIV-1. These included established therapeutic sites targeted to coding and non-coding regions of the HIV-1 long terminal repeat (LTR), Polymerase, Tat and Integrase. Expressed lhRNAs with ~63 bp duplex dsRNA regions and defined 5' and 3' termini were targeted to the transcribed region of

the HIV 5' LTR and effectively suppressed two distinct sites within this region across both subtypes B and C HIV sequences. In addition, to assess whether lhRNAs could inhibit basal levels of HIV transcription, the lhRNAs were shown to suppress Tat-mediated (processive) and Tat-independent (non-processive) transcription when targeting episomal and integrated LTR-driven sequences. Portions of the lhRNAs that produced the most active siRNAs were dissected by using tiled LTR targets cloned into a luciferase reporter gene and by using northern blot analyses. Dicer-processing of expressed lhRNAs was shown to be most effective from the base of the duplex and decreased in efficiency towards the loop, suggesting that a gradient of siRNAs production is possible from a single lhRNA but with decreasing efficacy. This work laid the foundation for improved expressed lhRNAs whereby multiple unique anti-HIV siRNAs were produced from a single lhRNA. The second combinatorial RNAi strategy made use of RNA Pol II-expressed pri-miRNA mimics, where each mimic was derived from a different endogenous scaffold. Polycistronic transcripts consisting of four different pri-miRNA scaffolds and targeting four separate sites in HIV were tested in several combinatorial systems. The pri-miRNA backbone chosen was shown to dramatically affect the concentration and inhibitory efficacy of each generated effector strand, and this was largely independent of the sequence used. A strategy to combine four of the most effective pri-miRNA scaffolds into one expression cassette was developed and significant inhibition of an HIV infectious molecular clone as well as a wild type HIV isolate was demonstrated. Finally, in an attempt to uncover additional asRNAs capable of inducing inhibition via transcriptional gene silencing of the HIV LTR promoter, indiscriminate cell-wide gene activation was shown to occur as result of an unintended off-target effect. These observations demonstrated that caution should be exercised when interpreting RNA-induced gene activation results. Overall, this thesis provides a detailed description of the efficacy of two novel combinatorial RNAi approaches based on single promoter expression systems that are aimed at generating multiple RNAi effector sequences targeted to HIV. These approaches pave the way towards a better understanding of the efficacy of combinatorial RNAi and an effective and sustained gene therapy of HIV.

Acknowledgements

There are many people who helped me to complete my PhD and I would like to thank each of them.

To Marco, thank you for starting me on this journey, for convincing me that science was where I want to be and for being such a wealth of knowledge and ideas. Your support through it all made this epic journey easier and you helped me to set the bar incredibly high. Thank you for becoming a good friend and I look forward to your company for decades to come.

To Patrick, thank you for all your support and especially for your attention to detail. You taught me to aim for perfection and it is a habit I hope to keep. Thank you for your friendship as well.

To all the people who helped with various aspects of the work presented here: to Sergio Carmona for the HBV Northern blots, to Lana Schaffer for the microarray work, to Jiang Han for ChIP analysis and to Kevin Morris for the 1G5 cells, for the endless advice and for letting me share a piece of OB with you.

To my husband, Liam, thank you for the check-oils and for making me laugh. Thank you for keeping it real and for letting me rant and for being so proud of me. Thanks too for sharing so much of yourself with me, you are my favourite mielie.

To my brother, C, thanks for being my number one fan and for being so excited even when it made no sense. Your enthusiasm makes me want to be a better scientist.

To my parents, mom, dad, Pat, Ronnie, Bev and Leuns, thank you so much for all the support and pride, especially when I kept pushing the goalposts a few months further forward. Your joy for each result made the ‘optimisation steps’ somewhat easier to bear and for that alone I cannot thank you enough.

Finally, my thanks to all the various bodies that funded the lab, my project or myself personally. To the University of the Witwatersrand, the Poliomyelitis Research Foundation, the German Academic Exchange (DAAD), the Medical Research Council, the National Research Foundation and to the European Union 6th Framework, thank you.

Aspects of the work described in this thesis have been published in the following articles and PDF versions of each are available in the appendices:

- Weinberg M S, Ely A, **Barichievy S**, Crowther C, Mufamadi S, Carmona S and Arbuthnot P. (2007). “Specific inhibition of HBV replication *in vitro* and *in vivo* with expressed long hairpin RNA”. Molecular Therapy 15 (3): 534-541.

- **Barichievy S**, Saayman S, von Eije K J, Morris K, Arbuthnot P and Weinberg M. (2007). “The inhibitory efficacy of RNA Pol III-expressed long hairpin RNAs targeted to untranslated regions of the HIV-1 5' long terminal repeat”. Oligonucleotides 17: 419-431.

- Weinberg M S, **Barichievy S**, Schaffer L, Han J and Morris K V. (2007). ”An RNA targeted to the HIV-1 LTR promoter modulates indiscriminate off-target gene activation”. Nucleic Acids Research 35 (21): 7303-7312.

- Saayman S, **Barichievy S**, Capovilla A, Morris K V, Arbuthnot P and Weinberg M S. (2008). “The efficacy of generating three independent anti-HIV-1 siRNAs from a single U6 RNA pol III-expressed long hairpin RNA”. PLoS One 3(7) e1-e10.

- **Barichievy S**, Saayman S, Arbuthnot P B and Weinberg M S. (2009). “RNA interference-based gene expression strategies aimed at sustained therapeutic inhibition of HIV”. Current Topics in Medicinal Chemistry 9 (12): e1-e14.

Table of contents

Declaration 2
Abstract 3
Acknowledgements 5
Preface 6
List of figures 11

Chapter one: general introduction 13

1.1 Introduction 13
1.2 HIV biology and origins 14
1.3 HIV genome and protein organisation 15
1.4 The viral life cycle 16
1.4.1 Entry 16
1.4.2 Uncoating and reverse transcription 19
1.4.3 Integration 20
1.4.4 Proviral expression and nuclear export of transcripts 21
1.4.5 Translation, assembly and budding 24
1.4.6 Viral accessory proteins 24
1.5 HIV latency 25
1.6 HIV pathogenesis and treatment 26
1.7 Nucleic acid based therapies 28
1.7.1 Nucleic acid therapies that have been tested in clinical trials 28
1.7.2 Rationale for additional gene therapy strategies 30
1.8 RNA interference 31
1.8.1 Mammalian microRNA biogenesis 31
1.9 RNAs in transcriptional regulation 36
1.10 Exogenous RNAi mimics as gene therapy tools 37
1.10.1 Synthetic RNAi molecules 38
1.10.2 Expressed RNAi mimics 39
1.11 RNAi-mediated targeting of HIV 42
1.11.1 Incoming viral genomic RNA as a target 42
1.11.2 Inhibition of HIV mRNAs and viral escape 43
1.11.3 RNAi-mediated inhibition of host cell targets required for HIV infection 45

1.12 Pre-clinical studies for the therapeutic application of RNAi 47
1.13 Anti-HIV RNA-based clinical studies 48
1.14 Combinatorial RNAi 49
1.14.1 Multiple shRNA expression 49
1.14.2 Long-hairpin RNAs 50
1.14.3 Polycistronic pri-miRNA mimics 51
1.15 Thesis objectives 52

Chapter two: expressed long hairpin RNAs targeted to the HIV LTR as a combinatorial RNAi strategy 54

2.1 Introduction 54
2.2 Materials and methods 56
2.2.1 Cloning of expressed anti-HIV long hairpin RNAs targeted to the 5' LTR 56
2.2.2 Cloning of a subtype C 5' LTR-driven luciferase reporter target vector 59
2.2.3 Cloning of a tiling array set of target reporter vectors 60
2.2.4 Mammalian cell culture 62
2.2.5 Transfection to generate a stable cell line expressing hFluc from a subtype C HIV 5' LTR….. 63
2.2.6 Transfection and detection of luciferase activity in cultured mammalian cells 64
2.2.7 Northern blot analysis to detect processed guide sequences from anti-LTR lhRNAs…..... ……..65
2.2.8 Quantitative real-time reverse transcription PCR (qRT-PCR) of interferon-related genes in cultured mammalian cells transfected with anti-LTR lhRNAs 66
2.2.9 Viral propagation and challenge assays using p24 ELISA 67
2.2.10 Statistical analysis 68
2.3 Results 69
2.3.1 Inhibitory efficacies of anti-HIV lhRNAs targeted to subtype B and C LTR-driven reporters 69
2.3.2 Alignment of LTRb and LTRc promoter sequences 72
2.3.4 Inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA LTR 2…. 75
2.3.5 Northern blot analysis of lhRNA HBV 78
2.3.6 Assessment of possible cellular toxicity induced by anti-HIV lhRNAs 80
2.3.7 Suppression of HIV replication using LTR-specific lhRNAs 82
2.4 Discussion and conclusions 85

Chapter three: expressed polycistronic pri-miRNAs targeted to discrete HIV mRNAs as a combinatorial RNAi strategy 89

3.1 Introduction 89
3.2 Materials and methods 91
3.2.1 Cloning of individual expressed pri-miRNA mimics targeted to four separate HIV mRNAs 91
3.2.2 Cloning of HIV target reporter vectors 98
3.2.3 Cloning of multimeric polycistonic anti-HIV pri-miRNA mimics 98
3.2.4 Transfection and detection of luciferase activity in cultured mammalian cells 100
3.2.5 Northern blot analysis to detect processed guide sequences from anti-HIV pri-miRNA mimics 101
3.2.7 Computational and statistical analysis 101
3.3 Results 102
3.3.1 Analysis of the predicted secondary structures of anti-HIV pri-miRNA mimics 102
3.3.2 Inhibitory efficacies of individual anti-HIV pri-miRNA mimics targeted to a luciferase reporter to assess the ability of different pri-miRNA backbones to generate functional guide sequences 106
3.3.3 The design and construction of multimeric polycistronic anti-HIV pri-miRNA mimics 109
3.3.5 Northern blot analysis of anti-HIV pri-miRNA mimics to determine guide sequence 112
3.3.6 Suppression of HIV replication using combinatorial anti-HIV pri-miRNA mimics 114
3.4 Discussion and conclusions 116

Chapter four: promoter targeted RNAs can modulate indiscriminate off-target gene activation 120

4.1 Introduction 120
4.2 Materials and methods 121
4.2.1 Cloning of expressed asRNAs targeted to site 247 and its flanking regions within the U3 region of a subtype B HIV LTR 121
4.2.2 Cloning of HIV LTR site 247 target vectors 122
4.2.3 Mammalian cell culture 123
4.2.4 Transfection and detection of asRNA activity in cultured mammalian cells 123
4.2.5 Chromatin immunoprecipitation (ChIP) 125

4.2.6 Microarray analysis 126
4.2.7 Statistical analysis 127
4.3 Results 127
4.3.1 Selection of HIV-1 subtype B LTR-targeted asRNAs 127
4.3.2 Screening of HIV-1 LTR targeted U6-expressed asRNAs 128
4.3.3 The role of histone modifications in LTR-247+7as-mediated increase in LTR-driven luciferase expression 132
4.3.4 Single changes in the LTR-247as+7 sequence can abrogate LTR-targeted activation 133
4.3.5 Microarray analysis of LTR-247as+7-treated cells 134
4.3.6 The effect of C10orf76 on LTR-247as+7-mediated gene activation 137
4.4 Discussion and conclusions 140

Chapter five: general discussion and conclusions 143

5.1 Combinatorial RNAi and viral escape 144
5.2 HIV-mediated modulation of the RNAi pathway 146
5.3 Unintended off-target effects related to RNAi 148
5.4 Into the future: delivery of RNAi mimics to patients 151
5.4.1 Vectors for the delivery of RNAi mimics 151
5.4.2 Clinical studies in human patients 154
5.5 Concluding remarks 155

References 158

Appendices 202

List of figures

Figure 1. The HIV genome and viral proteins. 16
Figure 2. The life cycle of HIV. 17
Figure 3. The HIV virion. 20
Figure 4. The HIV promoter..... 21
Figure 5. Transcription of the HIV provirus 22
Figure 6. The mammalian RNAi pathway 34
Figure 7. Exogenous mimics in the RNAi pathway.. 40
Figure 8. Long hairpin RNA target sites within the HIV 5' LTR..... 56
Figure 9. Two-step PCR cloning strategy used to generate anti-HIV lhRNAs. 57
Figure 10. Design of expressed anti-HIV lhRNA cassettes. 70
Figure 11. Design of LTR subtype B and C luciferase-based targets. 70
Figure 12. Inhibitory efficacies of three anti-HIV lhRNAs targeted to LTRb and LTRc luciferase reporter vectors. 71
Figure 13. Alignment and comparison of LTRb (LAI) and LTRc (Du151) target sequences..... 73
Figure 14. Anti-HIV lhRNA inhibition of cognate RNAs expressed from an integrated LTRc (Du151) target..... 74
Figure 15. Tiling sequences used to generate a set of dual luciferase target reporter vectors.. 76
Figure 16. Tiling array to assess the inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA LTR 2..... 76
Figure 17. Tiling array to assess inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA HBV. 77
Figure 18. Northern blot analysis to detect putative guide sequences produced from lhRNA HBV 80
Figure 19. Quantitative real-time RT-PCR analysis to determine IFN-β levels in HEK293 cells following transfection with anti-HIV lhRNAs 81
Figure 20. Anti-HIV lhRNAs do not affect cell morphology or transfection efficiency..... 82
Figure 21. Anti-HIV lhRNAs inhibit replication of a subtype B HIV molecular clone 83
Figure 22. Challenge assay to assess anti-HIV lhRNAs against a subtype C viral isolate of HIV 84
Figure 23. Two-step cloning strategy used to generate anti-HIV pri-miRNA mimics. 96
Figure 24. Cloning steps used to generate RNA Pol II-driven anti-HIV pri-miRNA mimics..... 97
Figure 25. Cloning strategy used to generate multimeric polycistronic anti-HIV pri-miRNA constructs. 100

Figure 26. Predicted secondary structures and Gibbs free energy values (δG) for pri-miRNA mimics 103
Figure 27. Design of expressed anti-HIV pri-miRNA cassettes. 106
Figure 29. Generation of a panel of multimeric anti-HIV pri-miRNA mimics using pri-miR122 Tat, pri-miR30a LTR, pri-miR31 IN and pri-miR155 Gag 110
Figure 30. Multimeric anti-HIV pri-miRNA mimics targeted to individual HIV targets 112
Figure 31. Northern blot analysis to detect putative guide sequences produced from expressed anti-HIV pri-miRNA mimics. 114
Figure 32. Challenge assay to assess combinatorial anti-HIV pri-miRNA mimics against a subtype C viral isolate of HIV 115
Figure 33. Region within the HIV LTR targeted by U6-expressed asRNAs. 128
Figure 34. Effects of various U6-expressed antisense RNAs on luciferase expression 129
Figure 35. LTR-247as activation is independent of HIV-1 Tat 130
Figure 36. LTR-247as+7 modulates episomal gene expression 131
Figure 37. H3K4me2+ in various asRNA-treated cells 133
Figure 38. A single point mutation in LTR-247as+7 at nucleotide 7 abrogates off-target gene activation in 1G5 cells 134
Figure 39. CCR5-expressed GFP shows increased expression in the presence of LTR-247as+7 relative to the control LTR-247as 135
Figure 40. Microarray analysis of 23 deregulated genes. 136
Figure 41. Screening the miRNA binding potential of LTR-247as+7 and C10orf76. 137
Figure 42. The effect of C10orf76 on 247as+7-mediated gene activation 138
Figure 43. Select putative LTR-247as+7 target sites within C10orf76 are susceptible to LTR247as+7-mediated suppression 139
Figure A1: A schematic representation of the dilution series used to generated a TCID50/mL value in U87.CD4.CCR5 cells using p24 ELISA 207
Figure A2: HBV tiling array sequences and cloning strategy 208
Figure A3: Design of expressed anti-HIV shRNA cassettes 210
Figure A4: Diagrammatic representation of anti-HIV shRNA target sites within the 5' LTR 210
Figure A5: Inhibitory efficacies of anti-LTR shRNAs 212
Figure A6: Challenge assay to assess anti-HIV shRNAs against a subtype C viral isolate of HIV 213
Figure A7: Alignment of sequenced and expected guide regions from pCi-LTR-Tat 214

1.1 Introduction

The World Health Organisation (WHO) currently estimates that there are 33.0 million (30 million to 36 million) people infected with the human immunodeficiency virus (HIV) worldwide with that number having grown by 2.7 million (2.2 million to 3.2 million) cases in 2007 alone. South Africa has the highest number of infections globally with 5.7 million (4.9 million to 6.6 million) HIV positive people (UNAIDS 2008). Currently available drugs to treat the infection are highly effective at reducing HIV-related deaths as well as opportunistic infections that occur as part of the broader HIV-induced disease referred to as AIDS (acquired immunodeficiency syndrome). Treatment success is tightly correlated with patient compliance which can be poor as a result of toxic side effects of the drugs or their by-products, accessibility of chronic treatment and viral resistance that almost inevitably emerges. This could be dramatically improved through the use of a single intervention that would aid the patient for life but current treatment formulations require repeat administrations to be effective. An alternative strategy that could provide a solution to the problems associated with conventional HIV therapies is gene therapy.

Gene therapy is defined as the use of a nucleic acid sequence administered to human patients with the intention of regulating, repairing or replacing a targeted genetic sequence (Verma and Somia 1997). One type of gene therapy in particular utilises a naturally occurring pathway to silence expression of genes. The pathway, known as RNA interference (RNAi) is triggered by double stranded (ds) RNA and results in homology-dependent control of gene expression (Elbashir, Lendeckel et al. 2001). The dsRNAs required to initiate the pathway can be designed against any sequence of interest and can be generated synthetically or expressed within a host cell. Importantly, these RNAi triggers are highly amenable to combination thereby allowing for the simultaneous generation of multiple RNAi suppressors known as small interfering RNAs (siRNAs). Effective combinations of RNAi inhibitors present a new and exciting addition to current gene therapy strategies that warrants further investigation.

The work contained in this thesis contributes to the field of gene therapy through the use of various RNAi suppressors in combination against HIV. Of particular interest is the incorporation of an assorted set of RNAi substrates into a simple expression system that can be readily manipulated to release potent anti-HIV siRNAs. Plasmid vectors were generated that expressed multiple siRNAs from different precursor dsRNAs. These expression systems were tested using assays specifically designed to assess the abilities of each system to generate multiple functional anti-HIV siRNAs. By characterising the effects of siRNAs generated from distinct

combinatorial systems, the feasibility of developing effective concentrations of multiple siRNAs from single expressed precursor substrates was established. This strategy is of particular value in the context of rapidly evolving viruses such as HIV.

1.2 HIV biology and origins

HIV is an enveloped retrovirus belonging to the lentivirus genus and is capable of infecting primate immune cells bearing a CD4+ surface receptor, including CD4+ helper lymphocytes, CD8+ cytotoxic lymphocytes and macrophages. The term lentivirus is derived from the Latin word *lentus* which means 'slow' and refers to viruses, such as HIV, that exhibit a characteristic long incubation period. Infection with HIV causes CD4+ cells to die and their decrease in number over time leads to a weakened immune system that allows for opportunistic infections to occur. Once the CD4+ helper lymphocyte count falls to less than 350 per mm^3 of blood HIV-infected individuals enter a disease state referred to as AIDS that can span years or even decades (Levy 2009). HIV has evolved a number of strategies to elude the host immune system including integration of the viral genome into the host chromosome thereby providing a permanent blueprint for virion production [reviewed in (Taylor, Sobieszczyk et al. 2008)]. The viral reverse transcriptase enzyme, required for generating a DNA copy of the RNA genome, lacks proof-reading ability that results in a high viral mutation rate of 3.4×10^{-5} mutations per base pair (bp) per replication cycle. This translates to 10^{10} HIV variants produced in a single day within one patient (Perelson, Neumann et al. 1996; Lee, Perelson et al. 2008). In addition, recombination is common where multiple infections have occurred within a single patient and this further increases viral genomic diversity (Templeton, Kramer et al. 2009).

HIV-1 originated from three independent cross-species transmissions of Simian Immunodeficiency Virus (SIV). These occurred from infected chimpanzees (*Pan troglodytes troglodytes* and *P. t. schweinfurthii*) to humans sometime before the 20^{th} century in west central Africa (Gao, Bailes et al. 1999; Korber, Muldoon et al. 2000; Sharp, Bailes et al. 2001). Genetic analyses have revealed a single change within the viral genome that occurred independently during cross-species transfer from apes to humans (Wain, Bailes et al. 2007) and strongly implies a host-specific adaptation of the virus (Hahn, Shaw et al. 2000). An independent zoonotic event from sooty mangabeys (*Cercocebus atys*) to humans has led to HIV-2 which is geographically restricted to Cameroon (Gao, Yue et al. 1992). The chimpanzee zoonotic events gave rise to three separate groups of HIV-1 namely M (major), N (non-major) and O (outlier). A fourth HIV-1 group distinct from groups M, N and O was recently reported and shown to be closely related to gorilla SIV (Plantier, Leoz, et al. 2009). Group M viruses are responsible for the worldwide HIV pandemic although the reasons for this remain unclear (Hahn, Shaw et al.

2000; Keele, van Heuverswyn et al. 2006). Further heterogeneity exists as group M viruses can be classified into subtypes A to J (Cohen, Anderson et al. 2005; Taylor, Sobieszczyk et al. 2008). Currently 15 % of global infections are caused by subtype B viruses while 48 % are subtype C infections (UNAIDS 2008).

1.3 HIV genome and protein organisation

The HIV genome is approximately nine kilobases (kb) and comprises two copies of positive strand RNA that encode a total of nine genes. These include information for three structural, two regulatory and four accessory proteins (Figure 1). The structural genes encode Gag (group specific antigen), Pol (polymerase) and Env (envelope) proteins. The regulatory genes encode Rev (regulator of expression of viral protein) and Tat (transcriptional transactivator) proteins. The four accessory genes encode Nef (negativity factor), Vif (viral infectivity factor), Vpr (viral protein R) and Vpu proteins [viral protein U or viral protein X (Vpx) in HIV-2; (Li, Hui et al. 1992)]. The 5' and 3' long terminal repeats (LTRs) are common to all retroviruses and serve as promoter sequences for the transcription of viral mRNAs as well as full length genomic transcripts (Gifford, Katzourakis et al. 2008). The *gag* gene is transcribed as a myristoylated precursor protein that is cleaved by viral protease into core protein components, p17 (matrix), p24 (capsid), p7 and p6 (nucleocapsid). The *pol* gene overlaps the *gag* gene at a UUUUUUAG sequence and a stem loop structure is present downstream of this junction. This causes ribosomal slippage on the RNA resulting in a -1 frameshift that produces a Gag-Pol precursor protein. This polyprotein is cleaved by viral protease into integrase (IN), RNAse H (RNH), reverse transcriptase (RT) and protease (PR) (Saliou, Bourgeois et al. 2009). The Env protein is produced as a precursor glycoprotein complex (gp160) that is cleaved into surface (gp120) and membrane (gp41) subunits by host proteases. The regulatory and accessory genes are translated as single proteins that do not require proteolytic cleavage to initiate function.

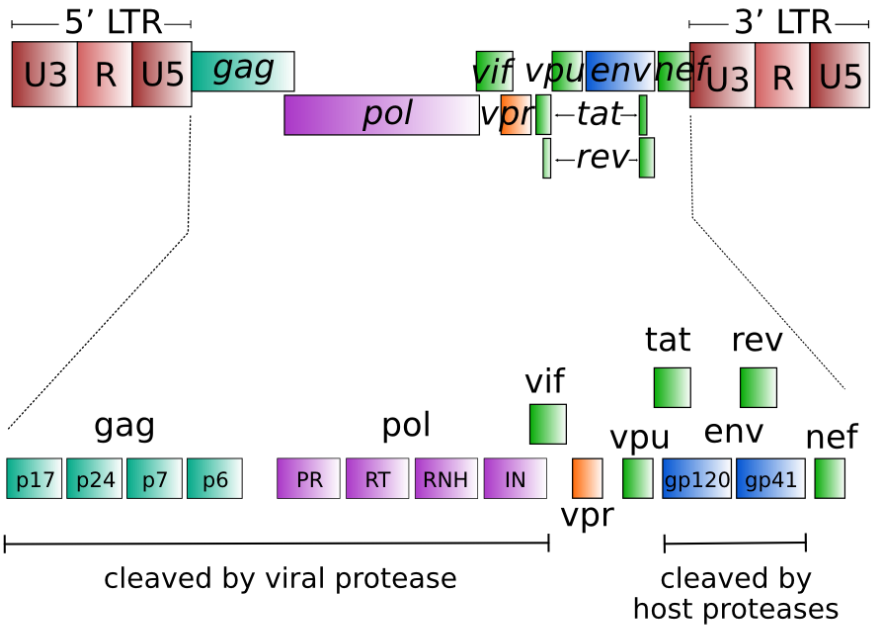


Figure 1. The HIV genome and viral proteins. Two copies of positive strand RNA make up the nine kb genome of HIV. Nine partially overlapping genes encode structural (*gag*, *pol* and *env*), regulatory (*tat* and *rev*) and accessory (*nef*, *vif*, *vpr*, and *vpu* or *vpx* in HIV-2) information. Transcription occurs at various stages throughout the viral life cycle driven by the 5' LTR promoter following integration into the host genome. The 3' LTR is used during reverse transcription of the RNA genome to generate a viral cDNA copy. The Env protein (shown in blue) is produced as a precursor glycoprotein complex (gp160) that is cleaved by host proteases into surface (gp120) and membrane (gp41) subunits. Viral polymerase (shown in purple) is a polyprotein cleaved by viral protease (PR) into reverse transcriptase (RT), RNAseH (RNH) and integrase (IN). The remaining regulatory and accessory proteins do not require cleavage to function.

1.4 The viral life cycle

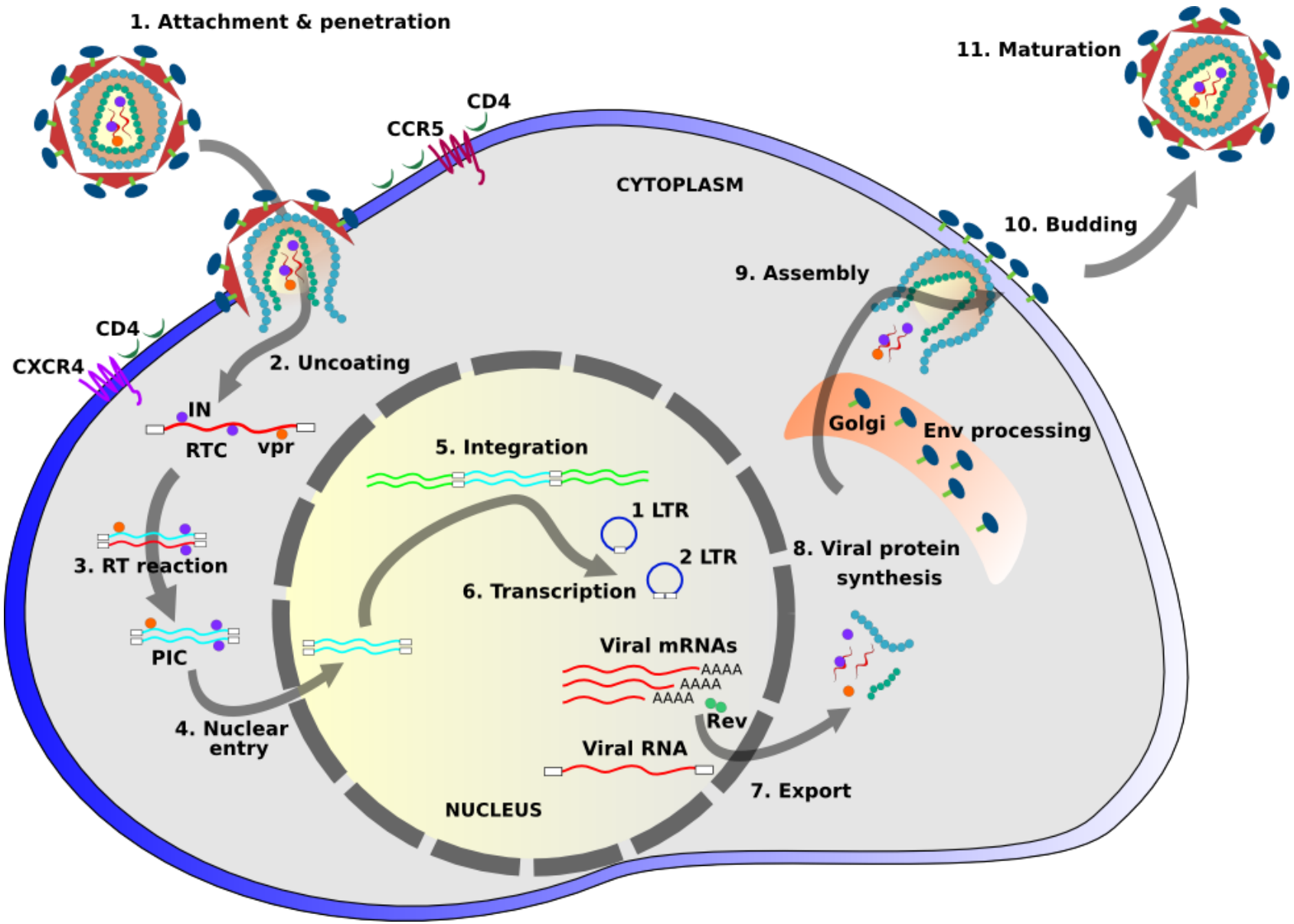
1.4.1 Entry

HIV infection occurs in CD4+ cells when HIV Env proteins bind the host CD4+ receptor and a chemokine co-receptor (Figure 2, number 1). Membrane-embedded mature gp120 and gp41 Env proteins are linked by non-covalent interactions and associate as trimeric molecules on the surface of viral particles (Wyatt, Kwong et al. 1998; Liu, Bartesaghi et al. 2008). The most variable portions of gp120 are highly glycosylated loop regions exposed at the surface of the virus. The sugars serve to mask large regions of gp120 and gp41 (Perdomo, Levi et al. 2008). Attachment of HIV is initially mediated by non-specific interactions between the glycosylated variable loops and charged glycolipids and glycoproteins of host cells. While insufficient for entry, this binding allows for subsequent receptor recognition by gp120 (Brelot and Alizon 2001).

CD4 is the major receptor for gp120 and while each monomer within the trimeric complex has a binding site for CD4, engagement at only one position is sufficient to induce conformational changes in all three glycoproteins that are required for viral entry (Salzwedel and Berger 2000). These changes allow for interactions with one of a number of chemokine co-receptors although chemokine receptor 5 (CCR5) and CXC chemokine receptor 4 (CXCR4) are commonly utilised by macrophage tropic (M-tropic) and T cell tropic (T-tropic) HIV respectively (Agrawal, Maxwell et al. 2009). The engagement of CD4-bound gp120 and its co-receptor triggers a rearrangement of gp41 trimers to form a rod-shaped bundle with three leucine zippers aligned as a coiled coil (Clapham and McKnight 2002). The exact events that lead to fusion are unclear but reorganisation of the gp41 bundle culminates in the merging of host membranes with viral envelope (Campbell and Hope 2008; Liu, Bartesaghi et al. 2008).

(Figure 2 is on the following page)

Figure 2. The life cycle of HIV. Infection occurs in cells that express CD4 receptors (shown in green) such as macrophages and T cells. Attachment of the virus (1) requires binding between HIV gp120, host CD4 and either a CCR5 (shown in dark purple) or CXCR4 (shown in light purple) co-receptor. This causes fusion of the host and pathogen membranes and penetration of the virus. Once in the cytoplasm, uncoating (2) of the viral genome from the capsid releases a reverse transcription complex (RTC). For simplicity, a single positive strand RNA genome is depicted in red bound to viral integrase (IN – purple circle) and Vpr (orange circle). The RTC triggers reverse transcription of the RNA genome (3) forming a double stranded DNA copy (shown in light blue). Vpr binds the DNA genome copy to form a pre-integration complex (PIC) and mediates nuclear entry (4). IN causes integration (5) of the viral genome (termed a provirus) into the host chromosome (shown in light green). Transcription (6) from the provirus generates primary transcripts that are either packaged as new genomes, or spliced to form numerous mRNAs. Non-transcribed 1-LTR and 2-LTR circles abound in the nucleus but do not successfully give rise to progeny. Viral Rev exports (7) nascent mRNAs into the cytoplasm where translation occurs (8). Gag proteins are transported to the host cell membrane and self-assemble (9) into capsid shells containing viral genomic RNA and accessory proteins. HIV Env protein is translated in the rough endoplasmic reticulum (ER) and cleaved into gp120 and gp41 in the Golgi complex. The latter subunit anchors the Env complex into the host membrane while covalently interacting with gp120. The concluding stage in assembly involves the coating of viral particles with Env-studded host membrane as they bud out of infected cells (10). After release, viral protease cleaves Gag/Pol precursors and initiates a structural rearrangement cascade of proteins that generates mature virus in the extracellular host environment (11).


1. Attachment & penetration
CD4
CCR5
CYTOPLASM
11. Maturation
10. Budding
9. Assembly
CD4
CXCR4
2. Uncoating
IN
RTC
vpr
5. Integration
Golgi
Env processing
1 LTR
2 LTR
8. Viral protein synthesis
3. RT reaction
6. Transcription
PIC
Viral mRNAs
AAAA
AAAA
AAAA
Rev
4. Nuclear entry
Viral RNA
7. Export
NUCLEUS

1.4.2 Uncoating and reverse transcription

The internal core of the HIV virion is composed of p17 matrix proteins that contain the viral genome and auxiliary viral proteins (Figure 3). This is surrounded by p24 capsid proteins that associate with the envelope boundary of the virion. The protective p24 shell is stable but following fusion of the viral envelope and host cell membrane, a necessary pH-induced conformational change allows the capsid to disassemble, the matrix to uncoat and the genomic RNA to be released (Figure 2, number 2). The detailed mechanism of this process is poorly understood (Ehrlich, Liu et al. 2001; Dvorin and Malim 2003). Once uncoated, the viral genome remains bound to p17 as well as nucleocapsid proteins (p7 and p6), Gag-Pol polyprotein cleavage products and Vpr in a complex referred to as the reverse transcription complex (RTC) (Freed 2001). Both p6 and p7 nucleocapsid proteins contain two zinc finger motifs (Cys-X2-Cys-X4-His-X4-Cys) and are involved in dimerisation and stabilisation of the viral genomic RNA (Zhang and Crumpacker 2002). The role of Vpr in the RTC is controversial but it may improve the fidelity of the process by modulating the mutation rate of HIV and diminishing mutations acquired during reverse transcription (Le Rouzic and Benichou 2005). For more details on the subject of reverse transcription, please refer to a review by (Harrich and Hooker 2002).

Binding of the RTC to cellular actin triggers reverse transcription of the viral genome by RT to form a 181 nucleotide (nt) single stranded (ss) negative sense DNA product (Davis, Carr et al. 2008). This serves as a primer that switches strands and hybridises to the 3' end of the viral genome to allow for completion of negative strand synthesis (Figure 2, number 3). The resultant molecule is a DNA/RNA hybrid and viral RNAse H (part of the RT complex) is responsible for degrading the RNA strand (Freed 2001). Positive strand DNA synthesis is initiated by DNA-dependent DNA polymerase and following a second strand transfer event a dsDNA copy of the viral genome is completed (Li, Stephenson et al. 1993). Importantly, the process of reverse transcription results in the copying of U3 and U5 sequences to form complete LTRs at either end of the dsDNA viral genome required for integration of HIV.

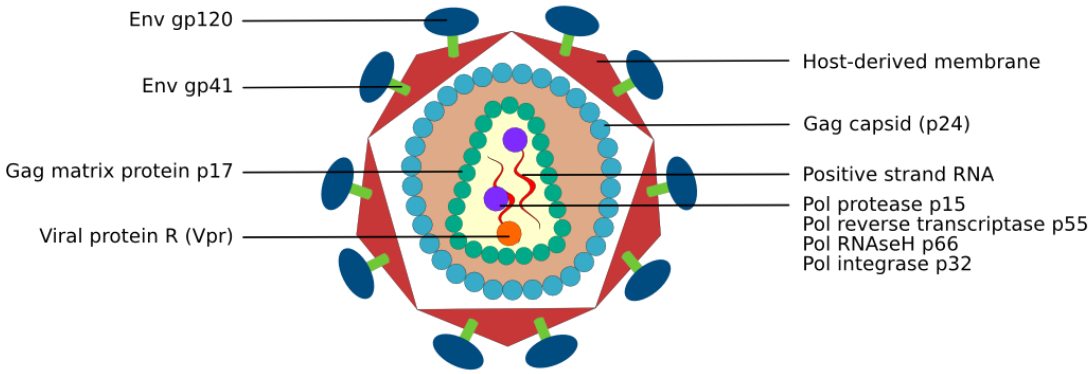


Figure 3. The HIV virion. The virion has a diameter of approximately 120 nm with surface studs of trimeric glycoproteins (gp120 and gp41) encoded by *env* that are embedded in a host-derived membrane. The viral p24 capsid (encoded by *gag*) is within the membrane and protects a conical *gag*-encoded nucleocapsid matrix (p17). Within the matrix core, two positive strand copies of the viral RNA genome and several viral proteins, including Vpr (orange circle) and polymerase (purple circle) can be found. The latter protein is cleaved into protease (p15), reverse transcriptase (p55), RNAseH (p66) and integrase (p32).

1.4.3 Integration

Synthesis of the DNA copy of the viral genome, referred to as a provirus, produces the pre-integration complex (PIC) (Li and Craigie 2009). Integration of the provirus into the chromatin of the infected host cell is a key feature of retroviruses and vital for the completion of the HIV life cycle as once integrated, the viral DNA is transcribed and replicated along with cellular DNA (Li and Craigie 2009). Unsuccessful integration events lead to re-circularised forms of the provirus termed 1-LTR or 2-LTR circles that can be detected in the nuclei of infected cells. Viral genes are not efficiently expressed from these substrates and do not give rise to successful progeny (Wiskerchen and Muesing 1995). For integration to take place the PIC is required to relocate, via microtubules, to the nucleus (Nisole and Saïb 2004). *En route* to the nucleus, the PIC protects the provirus from cellular DNAses and other host cytoplasmic sensors that may trigger an anti-viral response [Figure 2, number 4 and (Yoder, Sarasin et al. 2006; Borden, Sen et al. 2007)]. HIV Vpr forms part of the PIC and despite not encoding a canonical nuclear localisation signal may interact with host nuclear pore components to enhance transport of the PIC through the nuclear pore (Popov, Rexach et al. 1998). Notably, lentiviruses such as HIV can integrate within quiescent cells as dissociation of the nuclear membrane is not required for the virus to enter the nucleus (Heinzinger, Bukrinsky et al. 1994; Miller, Farnet et al. 1997). Additional information describing nuclear entry of retroviruses is available in the following reviews (Goff 2001; Suzuki and Craigie 2007).

Once inside the nucleus, viral IN preferentially mediates integration of the provirus at regions of host chromatin that are transcriptionally active (Engelman, Oztop et al. 2009; Felice, Cattoglio et al. 2009). During the first of two steps required for integration, two nt are removed

from the 3' end of one of the strands of the provirus to expose the terminal hydroxyl group and generate a recessed end (Figure 2, number 5). Secondly, DNA strand transfer takes place and the 3' provirus hydroxyl groups attack the phosphodiester backbone of host chromosomal DNA generating an integration intermediate in which the 3' viral ends are covalently bound to the 5' target DNA ends (Li and Craigie 2009). Host repair enzymes fill in the gaps to complete the process (Engelman 2003).

1.4.4 Proviral expression and nuclear export of transcripts

Regulation of proviral expression is mediated by the transcriptional transactivator protein, Tat, that is expressed shortly after integration and stimulates the elongation of RNA Polymerase II (RNA Pol II) to ensure full-length viral transcripts are produced [Figure 2, number 6 and (Barboric and Peterlin 2005)]. The HIV LTR promoter element is composed of U5 (unique 5' end), R (repeat) and U3 (unique 3' end) regions (Figure 4).

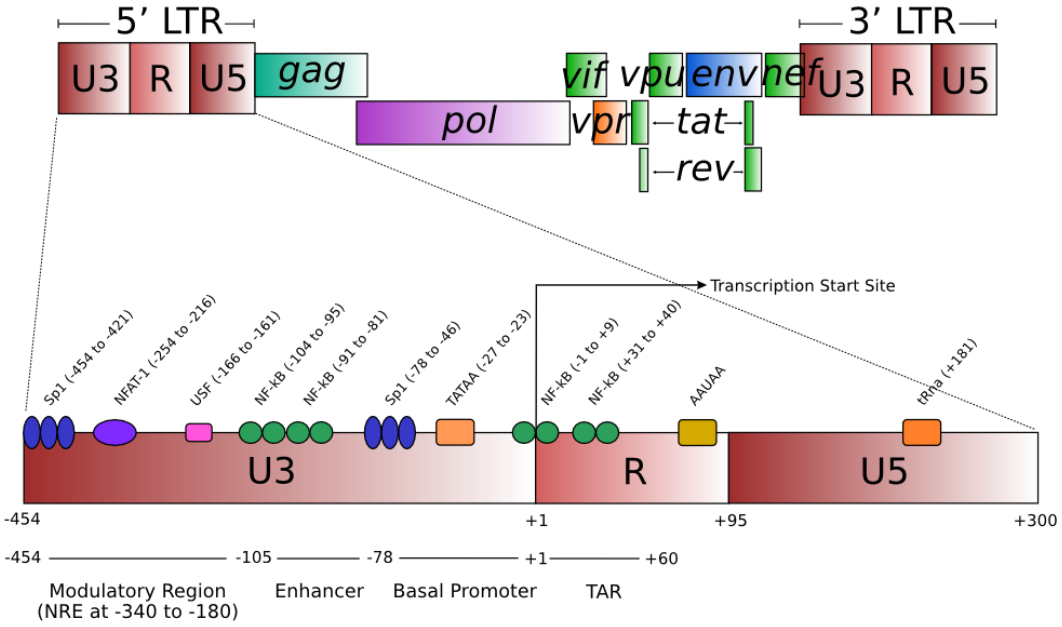


Figure 4. The HIV promoter. The HIV LTR acts as a promoter element and is composed of U3, R and U5 regions. The U3 region spans 454 bp (defined in relation to the transcription start site at +1) and can be divided into a negative regulatory (NRE) or modulatory element (-340 to -180), an enhancer (-105 to -79) and a promoter element (-78 to +1). The initial 60 bp of the R region (+1 to +95) fold into a loop structure (TAR) that binds Tat during productive transcription. The U5 region spans approximately 200 bp (+95 to +300) and contains a tRNA binding site (+181). The transcription start site of the LTR is at the R/U3 junction (+1) with a TATA box located 25 bases upstream in the U3 portion. Numerous Sp1 (blue ovals), NF-κβ (green circles) and other motifs serve as transcription factor binding sites for host proteins.

The transcription start site is at the R/U3 junction with a TATA box located 25 bases upstream in the U3 portion (Gifford, Katzourakis et al. 2008). At the start of eukaryotic transcription, a pre-initiation complex of general transcription factors, including the TATA box binding protein (TBP) and 12 to 15 TBP-associated factors, bind at the TATA box and direct RNA Pol II binding at this site. Two of the TBP-associated factors (IIE and IIH) unwind the DNA (Kim, Ebright et al. 2000) and transcription initiation proceeds whereby two nucleotide triphosphates (NTPs) are joined to form a phosphodiester bond (Dvir, Conaway et al. 1996). This allows promoter clearance of RNA Pol II although due to the instability of the early transcription complex, these transcripts are released before they have been elongated (Pal and Luse 2002). Elongation and productive transcription requires phosphorylation of serine 5 within the C-terminal domain of Rpb1, an RNA Pol II subunit [Figure 5 and (Ramanathan, Rajpara et al. 2001)].

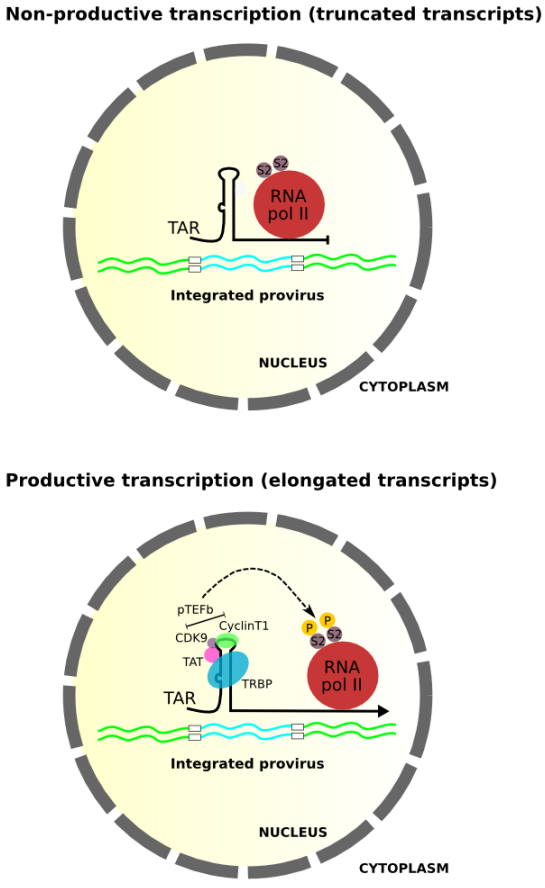


Figure 5. Transcription of the HIV provirus. Expression from the provirus (shown in light blue) can be divided into Tat-independent and Tat-dependent phases. In the absence of Tat (upper panel), RNA Pol II (shown in dark red) binds the transcription start site of the 5' LTR downstream of the TAR loop (shown in black) and initiates transcription. Un-phosphorylated RNA Pol II is non-productive and generates truncated transcripts. Some low levels of full length transcript are also produced, sufficient to generate Tat protein that transactivates transcription. During productive transcription (lower panel), Tat (pink circle) binds the TAR loop and recruits CyclinT1 (green oval), CDK9 (purple circle) and TRBP (blue oval). CDK9 phosphorylates (yellow circles) RNA Pol II allowing for elongation of transcripts.

Phosphorylation is mediated by the cyclin dependent kinase (CDK9) catalytic domain of positive transcription elongation factor b (pTEF-b) that phosphorylates RNA Pol II and ensures processive transcription [(Price 2000) and reviewed in (Sims III, Belotserkovskaya et al. 2004)]. HIV-1 Tat activates transcriptional elongation of the integrated provirus by recruiting pTEFb to the transactivating responsive (TAR) element that forms a stable RNA loop structure at the 5' end of all viral transcripts (Richter, Ping et al. 2002). The TAR loop was thought to comprise the initial 80 bp of the 5' LTR but deletion studies have established the minimal domain required for Tat binding comprise nucleotides 19 to 42 (Garcia, Harrich et al. 1989). TAR RNA folds with a six nt loop and contains two stem regions separated by a three bp pyrimidine bulge that is essential for Tat binding (Dingwall, Ernberg et al. 1990). Activation of RNA Pol II occurs following the formation of a co-operative Tat-pTEFb-TAR complex. Tat specifically interacts with the cyclin T1 subunit of pTEFb and the TAR loop adopts a conformation that allows Tat to bind at the tri-nucleotide bulge (Richter, Cao et al. 2002). This results in phosphorylation of both serine 2 and serine 5 of Rbp1 to activate RNA Pol II (Zhou, Halanski et al. 2000; Kim, Bourgeois et al. 2002). Importantly, regardless of the site of proviral integration, the 5' LTR is always packaged by two nucleosomes (Nuc 0, -413 to -253 and Nuc +1, +1 to +155) the second of which is located close to the transcriptional start site. The binding of cellular transcription factors upstream of Nuc +1 induces remodelling of this nucleosome and subsequent transcription [reviewed in (Weinberg and Morris 2006)]. Thus, the chromatin state at the site of HIV integration aids in ensuring proviral silencing.

HIV, like all retroviruses, uses a post-transcriptional mechanism to generate multiple mRNAs from a single proviral template (Lee, Culver et al. 2008). The viral transcripts expressed from the LTR include (1) un-spliced mRNAs that function as Gag-Pol precursors and genomic RNA for progeny virions, (2) incompletely spliced mRNAs that express Env, Vif, Vpr and Vpu proteins, and (3) abundantly spliced mRNAs that code for Rev, Tat and Nef proteins (Freed 2001). While the latter transcripts (~2 kb) are constitutively transported via cellular machinery to the cytoplasm, translocation of the intron-containing transcripts requires an additional mechanism that bypasses cellular splicing machinery [Figure 2, number 7 and reviewed in (Pollard 1998)]. Rev functions to shuttle un-spliced and partially spliced mRNAs to the cytoplasm by binding the Rev responsive element (RRE) conserved within these transcripts (Pond, Ridgeway et al. 2008). At low Rev concentrations, nascent un-spliced and partially spliced mRNAs are degraded but once a critical concentration of Rev has been reached export of the mRNAs to the cytoplasm via the Crm-1 pathway occurs (Fridell, Bogerd et al. 1996).

1.4.5 Translation, assembly and budding

In the host cytoplasm, expression of a complete set of viral proteins leads to concurrent assembly and budding of new virions [Figure 2, numbers 8 to 10 and (Zhang and Nguyen 2008)]. Within minutes of their production, structural Gag proteins are transported to and associate with the host cell membrane via myristoylation signals in their N-termini (Campbell and Hope 2008). Here Gag proteins self-arrange in a radial fashion and interact with matrix proteins that have enveloped viral genomic RNA (Gottlinger 2001; Lanman, Sexton et al. 2002). Incorporation of the RNA is mediated by an encapsidation signal (ψ) located in *gag* and p7 and p6 nucleocapsid proteins are involved in dimerisation of two viral genomic RNA strands (Zhang and Crumpacker 2002). HIV Env protein is translated in the rough endoplasmic reticulum (ER) of host cells as a gp160 precursor and cleaved into its gp120 and gp41 subunits in the Golgi complex before the latter becomes anchored in the host cell membrane (Liu, Bartesaghi et al. 2008). The concluding stages in assembly involve the coating of viral particles in host membrane studded with Env spikes as they bud out of infected cells and subsequent maturation of immature virions (Ganser-Pornillos, Yeager et al. 2008). Maturation is initiated by proteolysis of Gag-Pol precursors leading to a structural rearrangement cascade of proteins that finally generates mature virus in the extracellular host environment [Figure 2, number 11 and (Moore, Fu et al. 2008)].

1.4.6 Viral accessory proteins

The HIV accessory proteins interact with various host proteins in order to enhance virion production and counteract certain host-mediated anti-viral mechanisms [reviewed in (Malim and Emerman 2008)]. Briefly, Vpu aids gp160 maturation by sequestering CD4 proteins in the ER and mediating their ubiquitination and degradation. This ensures gp160 translocation to the host cell surface where it can be incorporated into budding virions (Göttlinger 2008; Salim and Ratner 2008). Vif interacts with a member of the APOBEC (apolipoprotein B mRNA-editing catalytic polypeptide) protein family to prevent deamination of cytoplasmic viral genomic RNA in budding virions [reviewed in (Goila-Gaur and Strebel 2008)]. In addition to its role in the RTC, Vpr induces cell cycle arrest in the G2 phase of infected cells. The biological relevance of this is unclear but G2 arrest may favour LTR transcription (Le Rouzic and Benichou 2005; Andersen, Le Rouzic et al. 2008). The HIV Nef protein is expressed early in the viral life cycle and ensures establishment and persistence of infection by enhancing fusion at the plasma membrane and uptake of virions via the cytosolic pathway (Schaeffer, Geleziunas et al. 2001). This action is mediated in a CD4-independent manner but Nef also mediates down-regulation of surface-

bound CD4 receptors, accelerating their degradation via lysis and preventing sequestration of the viral Env protein (Das and Jameel 2005). This surface down-regulation prevents super-infection of the host cell which would otherwise lead to cell, and subsequent viral, death (Foster and Garcia 2008).

1.5 HIV latency

A major obstacle to eradicating HIV is the establishment of a latent infection. Thymus-derived (T) cells bearing a CD4 receptor undergo a reversible resting phase during their natural life cycle and HIV infection and subsequent integration into these cells provides a blueprint for viral production that is extremely stable and can persist for the lifetime of the patient (Chun, Finzi et al. 1995; Finzi, Blankson et al. 1999; Siliciano, Kajdas et al. 2003). In addition to CD4+ T cells, macrophages and dendritic cells can serve as reservoirs for HIV as internalised virions can remain infectious for weeks within cellular endosomes and be transmitted to neighbouring T cells (Pelchen-Matthews, Kramer et al. 2003; Sharova, Swingler et al. 2005). CD4+ T cells are the main source of latent infection which can occur either pre- or post-integration of the provirus into host chromatin [reviewed in (Lassen, Han et al. 2004)]. Should HIV infection occur just prior to cellular transition from an active to resting state, integration may be blocked and this is referred to as pre-integration latency (Chiu, Soros et al. 2005). When these infected cells are activated by mitogens in their surrounding environment, the extra-chromosomal provirus can integrate and generate progeny (Pierson, Zhou et al. 2002). A more stable form of latency occurs post-integration when CD4+ T cells that carry a provirus revert to a resting state. While integration into centromeric heterochromatin partly accounts for repressive transcription (Jordan, Defechereux et al. 2001), additional host and viral factors play a role in maintaining latency [reviewed in (Marcello 2006)].

The U3 region of the HIV LTR encodes a number of consensus sequences within the modulatory region (Figure 4) for binding of host transcription factors such as NFAT (nuclear factor of activated T cells) and NF-κβ (nuclear factor kappa-beta) (Williams and Greene 2005; Williams, Chen et al. 2006). Both of these transcription factors are sequestered in the cytoplasm of resting T cells but are recruited to the nucleus following cellular activation by T-cell receptor (TCR) binding or cytokine signalling via interleukin 2 (IL-2) or tumour necrosis factor alpha (TNFα) (Pereira, Bentley et al. 2000). Once in the nucleus, NFAT and NF-κβ can initiate transcriptional activation through binding the U3 region of the HIV LTR and in this way, HIV transcription is linked to T cell activation (Scripture-Adams, Brooks et al. 2002; Hogan, Chen et al. 2003). Furthermore, the HIV LTR contains three Sp-1 sites for binding of general transcription activators and these may play a role in initiating basal levels of non-processive transcription in

resting T cells when NFAT and NF-κβ are unavailable (Lassen, Bailey et al. 2004). In addition to boosting transcription (described in section 1.4.4), the interaction of HIV Tat with the TAR loop is associated with a co-incident recruitment of histone acetyl-transferases such as p300 and CBP (cyclic-AMP responsive transcription factor binding protein). These proteins modulate various interactions at the nucleosome such as acetylation of the p50 subunit of NF-κβ (Benkirane, Chun et al. 1998; Lusic, Marcello et al. 2003). Furthermore, acetylation of Tat's activation domain by p300/CBP-associated factor (PCAF) has been shown to stabilise the TAR-Tat-cyclin T1 interaction and improve transcription from integrated provirus (D'Orso and Frankel 2009). All these interactions are evidence that HIV latency is a complex state and not simply a result of inaccessibility of the provirus to host transcription machinery in quiescent cells. Indeed, HIV has recently been shown to modulate its own expression and latency through manipulation of members of the RNAi pathway (Han and Siliciano 2007; Huang, Wang et al. 2007; Triboulet, Mari et al. 2007) and a full discussion of these modulations is described in section 5.2. The key issue with HIV latency is that unless the small pool of infected resting T cells can be eliminated, HIV cannot be eradicated within a patient and activation of the reservoir resumes productive infection and disease. Current therapies cannot solve this issue but new gene therapy strategies discussed in section 4.4 may be able to specifically affect the state of latent HIV.

1.6 HIV pathogenesis and treatment

The course of HIV infection can be divided into an acute and chronic phase, the former of which is characterised by poor to no HIV-specific antibody detection, low copies of viral RNA in the patient (termed viral load) and severe destruction of gut associated lymphoid tissue [reviewed in (Paiardini, Frank et al. 2008)]. This phase reflects immense interactions between HIV and the human immune system, with the viral load at the end of this stage determining a total burden of disease or 'set point'. The set point has considerable implications for the spread of HIV as the greater the viral load, the higher the chance of transmission (Cohen, Anderson et al. 2005). The heterogeneity of the virus may also play a role in spreading infection as within one infected individual a cosmopolitan mix of genotypes exist, some of which are favoured during transmission events – a postulate related to the overwhelming occurrence of subtype C infections worldwide [reviewed in (Taylor, Sobieszczyk et al. 2008)]. During the chronic phase, the viral load increases gradually, the CD4+ T cell count declines slowly and clinical latency occurs. Over time, but specifically in the absence of anti-HIV drugs, the viral load rebounds and CD4+ T cells die off leaving the patient severely immune-compromised.

Since the 1996 introduction of drugs targeted to HIV, morbidity and long-term prognoses of patients have substantially improved (Palella, Delaney et al. 1998; Walensky, Paltiel et al.

2006). The remarkable genetic diversity coupled with the high mutation rate of the virus led to the use of combinations of drugs for the suppression of viral replication and this is referred to as highly active antiretroviral therapy (HAART). While only a few drugs were initially available, there are currently more than 20 approved antiretroviral drugs (ARVs) that can be divided into four distinct classes based on their modes of action (Keiser, Orrell et al. 2008). Entry inhibitors block HIV binding to host cells, either by preventing the formation of a gp41 bundle (Wild, Shugrass et al. 1994) or by allosteric, non-competitive binding of the CCR5 co-receptor (Watson, Jenkinson et al. 2005; Hunt and Romanelli 2009). The biggest class of ARVs are reverse transcriptase inhibitors (RTIs) that can be nucleoside/nucleotide (NRTI) or non-nucleoside (NNRTI) based. NRTIs inhibit reverse transcription of HIV as they are incorporated into newly synthesized viral DNA and prevent elongation (Hang, Li et al. 2007). NNRTIs directly bind the RT enzyme and prevent its functioning (Hang, Li et al. 2007). Integrase inhibitors (INIs) prevent integration of the provirus into host chromatin by interfering with viral IN binding or the strand transfer events in the process (Sato, Motomura et al. 2006). The fourth class of ARVs currently available are protease inhibitors (PIs) that prevent cleavage of viral polyproteins in immature budding viruses (Flexner 1998).

A key issue with conventional drugs is resistance either from mutations developed by actively replicating HIV, transmission of resistant viral strains or stable viral reservoirs that permit low level replication in the presence of therapeutic agents and subsequent resistance mutation events [reviewed in (Richman, Margolis et al. 2009)]. Within a single patient, the impressive heterogeneity of HIV means there is a litany of possible mutants and a single mutation can induce resistance to an entire drug class (Rimsky, Shugars et al. 1998; Tuaillon, Gueudin et al. 2004; Johnson, Brun-Vézinet et al. 2008; Pillay, Pillay et al. 2008; Hunt and Romanelli 2009). As only four drug classes are currently available, reformulating HAART cocktails is limited and fresh clinical trials to reassess efficacy and safety may be required, with a significant possibility of HIV evolving resistance to the new combination. Novel drug classes, such as maturation inhibitors are being investigated but have only been evaluated for safety in healthy volunteers and require further testing before use in HIV infected patients (Martin, Blum et al. 2007; Martin, Blum et al. 2007). Furthermore, despite the lowered toxicities in the newer generation ARVs, the cumulative effects of these treatments long-term is a concern given that most of the drugs affect lipid metabolism and mitochondrial functioning [reviewed in (Mallon 2007)]. Thus while the successes of HAART regimens cannot be ignored, their shortcomings open avenues for experimental approaches such as nucleic acid based therapies that may solve some of the concerns related to current treatments.

Nucleic acid treatments differ fundamentally from conventional drugs in their mode of action as nucleic acid drugs function specifically by hybridising to unique target sequences.

Theoretically this means any specific sequence within a gene or RNA transcript can be targeted thereby hugely expanding the possibilities for HIV inhibition. In addition, as nucleic acid therapies share the same modalities, the pharmacokinetic and pharmacodynamic profiles are the same thus one safety profile should provide sufficient data for different agents to be combined. This would dramatically reduce costs involved in clinical trials and allow for therapies to reach patients in a significantly shorter period of time. Finally, the relative ease with which nucleic acids can be synthesised means regimens could be rapidly altered and customised as infection profiles changed.

1.7 Nucleic acid based therapies

Nucleic acid therapy falls into the broader class of gene therapy and involves the use of intracellularly expressed exogenous genes to inhibit unwanted viral or cellular replication. The exogenous gene (transgene) can either be introduced via a transfer vector and expressed within the target cell(s) or the transgene product (RNA or protein) can be directly introduced to initiate the silencing effect (Klug and Cummings 2000). A number of gene therapy strategies have been explored over the last fifteen years that include RNA as well as protein based agents. The former involves RNA oligonucleotides that hybridise to a complementary target while the latter uses a competitive inhibition approach akin to conventional drugs. While some examples are currently in developmental phases and remain highly experimental, certain modalities have undergone rigorous assessment and been tested in clinical trials. The following section provides a snapshot of different nucleic acid modalities used against HIV as opposed to a comprehensive review.

1.7.1 Nucleic acid therapies that have been tested in clinical trials

Antisense oligonucleotides (ASOs) are ssDNA molecules that bind via sequence complementarities to cognate mRNA forming a stable RNA/DNA hybrid that undergoes ubiquitous and constitutive RNAse H degradation (Chan, Lim et al. 2006). To induce a therapeutic silencing effect, ASOs need to avoid nuclease degradation en route to the target cell, escape lysosomal-mediated degradation and maintain integrity within the cell at clinically relevant concentrations [reviewed in (White, Anastasopoulos et al. 2009)]. Early versions of ASOs were unstable *in vivo* due to nuclease digestion. Chemical modifications such as phosphorothioate linkages improved this but overall mRNA target affinity of the modified ASOs was reduced (Geary, Henry et al. 2002). Further alterations to the ribose ring of ASOs increased their target specificities but rendered them refractory to RNAse H degradation (Prakash and

Bhat 2007). New generation ASOs include phosphorothioate linkages and a morpholino ring backbone that enhances serum stability and RNAse H cleavage *in vivo* (Summerton 2007). Various malignancies as well as a range of pathogenic viruses including HIV have been targeted by ASOs [reviewed in (Rayburn and Zhang 2008; Spurgers, Sharkey et al. 2008)]. HIV Vif (Barnor, Miyano-Kurosaki et al. 2004), RRE (Prater, Saleh et al. 2007), RNAse H (Heinrich, Mathur et al. 2009) and TAR (Chaubey, Tripathi et al. 2007) have all been targeted using expressed forms of ASOs. Interestingly, a phase I clinical trial assessing an anti-HIV *env* ASO demonstrated long-term expression of the construct *in vivo* but no antiviral effects of the ASO were detected (Levine, Humeau et al. 2006). This may have been the result of inherent inefficiencies in the RNAse H pathway, suboptimal delivery of the ASOs to their targets cells or non-specific mRNA binding (Summerton 2007). These issues relate to all ASOs and thus a number of improvements are required before they can enjoy wide-spread use.

Ribozymes are non-coding RNAs with catalytic abilities that can be designed to target specific RNA sequences [reviewed in (Sarver, Cantin et al. 1990)]. Ribozymes can cleave in *cis* or *trans* and have intrinsic requirements for specific magnesium chloride ($MgCl_2$) concentrations (Goila and Banerjea 1998) and water molecules (Rhodes, Reblova et al. 2006). Both HIV and various host factors required for infection have been targeted using ribozymes (Ojwang, Hampel et al. 1992; Goila and Banerjea 1998; Bai, Gorantla et al. 2000; Cagnon and Rossi 2000; Bai, Rossi et al. 2001) and extensive data laid the foundations for clinical trials to be conducted. Initial phase I clinical trials revealed that retroviral vectors could successfully deliver anti-HIV ribozymes to haematopoetic stem cells from patients (Michienzi, Castanotto et al. 2003; Amado, Mitsuyasu et al. 2004; Ngok, Mitsuyasu et al. 2004; Macpherson, Boyd et al. 2005). However, while the method of delivery was a viable option, the *in vivo* efficacies of the transduced cells were poor despite their long-term stable persistence. A recent phase II clinical trial also utilised autologous transplantation of transduced haematopoetic stem cells (HSCs) to assess a ribozyme targeted to the overlapping *tat*/*vpr* ORF (Mitsuyasu, Merigan et al. 2009). No adverse effects were noted and despite no change in viral load over 48 weeks, CD4+ T cell counts were higher for the patients who received the anti-HIV ribozyme. Taken together the clinical trial data suggests that ribozymes can be safely delivered to target cells but the *in vivo* efficacies of these gene therapy agents need to be improved.

RNA decoys are homologues of RNA targets such as RRE and TAR that bind natural viral proteins and competitively inhibit the native ligands necessary for HIV replication (Sullenger, Gallardo et al. 1990; Browning, Cagnon et al. 1999). U16 small nucleolar RNA (snoRNA) has been used to deliver TAR and RRE decoys with appreciable inhibition of HIV noted for both (Banerjea, Li et al. 2004; Michienzi, de Angelis et al. 2006). A clinical trial has been conducted using an RRE decoy delivered by a retroviral vector into HSCs from HIV

positive patients (Kohn, Bauer et al. 1999). No adverse effects following the procedure were detected but the number of circulating transduced cells was too low to assess any anti-HIV activity. Protein homologues of RNA decoys, referred to as trans-dominant negative derivatives, mimic viral proteins such as Rev or Tat but lack vital functional moieties. Over-expression of these trans-dominant proteins titrates out their endogenous counterparts interfering with regular viral functioning (Trono, Feinberg et al. 1989; Caputo, Grossi et al. 1996; Woffendin, Ranga et al. 1996). Clinical trials to assess the efficacies of TAR decoys and trans-dominant Rev proteins have been conducted in paediatric patients. HSCs from HIV positive children were transduced with either a TAR decoy or a transdominant Rev protein and significant long-term survival even in the presence of a high viral load was demonstrated in at least 50 % of the patients (Morgan, Walker et al. 2005; Podsakoff, Engel et al. 2005). While these data strongly suggest that decoys are effective inhibitors of HIV *in vivo* their efficacy is strongly correlated with immune reconstitution following stem cell transduction and HIV resistance can rapidly develop in the presence of suboptimal decoy concentrations.

1.7.2 Rationale for additional gene therapy strategies

Further development of the gene therapy approaches mentioned in section 1.7.1 have been hampered by serum instabilities, suboptimal delivery and toxicity as a result of the modifications required for their optimal functioning in a cellular environment [reviewed in (Rossi, June et al. 2007)]. Serum instability and toxicity are often related as the modifications required for stability of naked nucleic acids in serum may be immunogenic. However those agents that are not modified may have little effect as they are rapidly degraded in host cells. Delivery methods have typically involved retroviral vectors that carry gene therapy payloads but success depends on the transduction of a critical number of patient cells. Lentiviral vectors have an added advantage of being able to transduce non-dividing cells and could theoretically improve the overall number of genetically modified cells, but insertional mutagenesis as well as the possibility of recombination events between the vector and wild type virus remains a concern when using these vectors in the context of HIV (Morris and Rossi 2006). Ideally what is required is a nucleic acid based strategy that harnesses the simplicity and elegance of simple Watson-Crick base-pairing to achieve a highly specific and efficacious gene silencing effect with limited toxicity *in vivo*. The discovery that a natural highly conserved RNA-mediated gene regulatory pathway exists opened a new paradigm of opportunities for gene therapy known as RNA interference (RNAi). This pathway utilises dsRNA to guide homology-dependent silencing of target RNA in a manner similar to ASOs and ribozymes. Notably though, as RNAi is an endogenous pathway the canonical silencing molecules are non-toxic and highly efficacious *in vivo* with potencies far

greater than other exogenous RNA molecules (Robinson 2004; Haasnoot, Westerhout et al. 2007). Since the official discovery of RNAi in 1998 there has been an explosion in our understanding of the pathway itself allowing RNAi to be modified for multiple purposes including antiviral therapy.

1.8 RNA interference

RNAi phenomena were initially described in tobacco plants following infection (Wingard 1928) and later in petunias after the introduction of exogenous pigment genes (Napoli, Lemieux et al. 1990). In lower eukaryotes RNAi functions as an antiviral mechanism whereas higher eukaryotes utilise the pathway as a robust regulator of gene expression (Stefani and Slack 2008) with links to embryonic development, cell differentiation, proliferation, metabolism and death [reviewed in (Bartel and Chen 2004; Singh, Bhadra et al. 2008; Stefani and Slack 2008; Yekta, Tabin et al. 2008)]. Pioneering experiments in *Caenorhabditis elegans* revealed that dsRNA was the trigger of RNAi, and that the effects could be systematic and heritable (Fire, Xu et al. 1998). Viral replication, transcription of convergent genes, self-annealing cellular transcripts and mobile genetic elements can all generate dsRNA substrates that are converted to small interfering RNAs (siRNAs) to guide cleavage of cognate mRNA targets by RNAi machinery (Elbashir, Lendeckel et al. 2001). Hundreds of RNAi silencing molecules have since been identified in animals although many of their targets have not been defined.(Lagos-Quintana, Rauhut et al. 2001; Lee and Ambros 2001; Lewis, Shih et al. 2003) Two recent reviews have suggested a classification system based on RNAi precursors, biogenesis pathways and associated proteins (Ghildiyal and Zamore 2009; Kim, Han et al. 2009). Many classes of small RNAs that feed into the RNAi pathway exist but for the purposes of this thesis only the mammalian RNAi pathway encompassing microRNAs (miRNAs) and siRNAs is described.

1.8.1 Mammalian microRNA biogenesis

Expression and nuclear processing

Endogenous miRNAs are usually generated from RNA polymerase (Pol) II promoters (Lee, Kim et al. 2004) although some, usually Alu-repeat-associated, are transcribed by Pol III (Borchert, Lanier et al. 2006). Approximately 40 % of miRNA loci are located within introns of non-coding transcripts, 10 % are positioned in exons of non-coding transcripts and 40 % are situated in the introns of protein-coding transcripts (Lewis, Shih et al. 2003). Interestingly, many miRNAs occur as clusters of two or more miRNAs with highly correlated expression (Kim and Nam 2006).

Clustered miRNAs have also been shown to have functional relatedness and thus while they are not strictly polycistronic transcripts they may be referred to as ‘polycistronic miRNAs’ (Yuan, Liu et al. 2008). Pol-II generated primary miRNA transcripts (pri-miRNAs) are generally several kb in length, contain several stem-loop structures, have ssRNA flanking segments and are processed in two compartmentalised steps [Figure 6 and (Lee, Jeon et al. 2002)]. The initial step involves an RNAse-III protein family member Drosha that cleaves at the stem of the pri-miRNA to release a hairpin structure termed a pre-miRNA (Lee, Ahn et al. 2003; Han, Lee et al. 2004). Drosha is comprised of two RNAse III domains and a dsRNA binding domain (dsRBD), the latter of which is insufficient for substrate binding (Jinek and Doudna 2009). A helper protein, the DiGeorge syndrome critical region gene 8 (DGCR8) co-interacts with Drosha and the pri-miRNA to form a 650 kiloDalton (kDA) Microprocessor complex (Denli, Tops et al. 2004; Gregory, Yan et al. 2004; Han, Lee et al. 2004). DGCR8 specifically recognises the stem of a pri-miRNA (approximately 33 bp) as well as the junction between the ss and dsRNA regions and directs Drosha-mediated cleavage 11 bp from this site (Zeng and Cullen 2005; Han, Lee et al. 2006). The resultant pre-miRNA is approximately 70 bp and includes a 5' phosphate and two-nucleotide 3' overhang. Regulation of the Microprocessor complex occurs post-transcriptionally (Han, Pedersen et al. 2009). Vertebrate DGCR8 mRNA transcripts contain two conserved hairpins with features similar to pri-miRNAs. These structures are cleaved by Drosha resulting in suppression of DGCR8. Notably, the cleavage is less efficient than with genuine pri-miRNAs – a feature that is consistent with a regulatory role. DGCR8 protein in turn stabilises Drosha through a protein-protein interaction thereby completing a positive feedback loop (Han, Pedersen et al. 2009).

Interestingly, pri-miRNA processing may occur co-transcriptionally and prior to intronic splicing in the nucleus (Morlando, Ballarino et al. 2008). Drosha-mediated cleavage of intronic miRNAs does not impair transcript splicing (Kim and Kim 2007) but if the miRNA is located exonically, Drosha processing destabilises the transcript and affects downstream protein translation (Han, Pedersen et al. 2009). Cellular splicing machinery also plays a role in the biogenesis of a distinct class of miRNAs that bypass Drosha-mediated cleavage (Berezikov, Chung et al. 2007; Ruby, Jan et al. 2007)]. Termed mirtrons, these intronic miRNAs are derived from Pol II-transcribed primary-mirtron precursors that encode canonical splice sites (Figure 6). Host splicing machinery recognises the splice sites and cleaves the transcript generating a hairpin structure that resembles a Drosha product (Ying and Lin 2006; Okamura, Hagen et al. 2007).

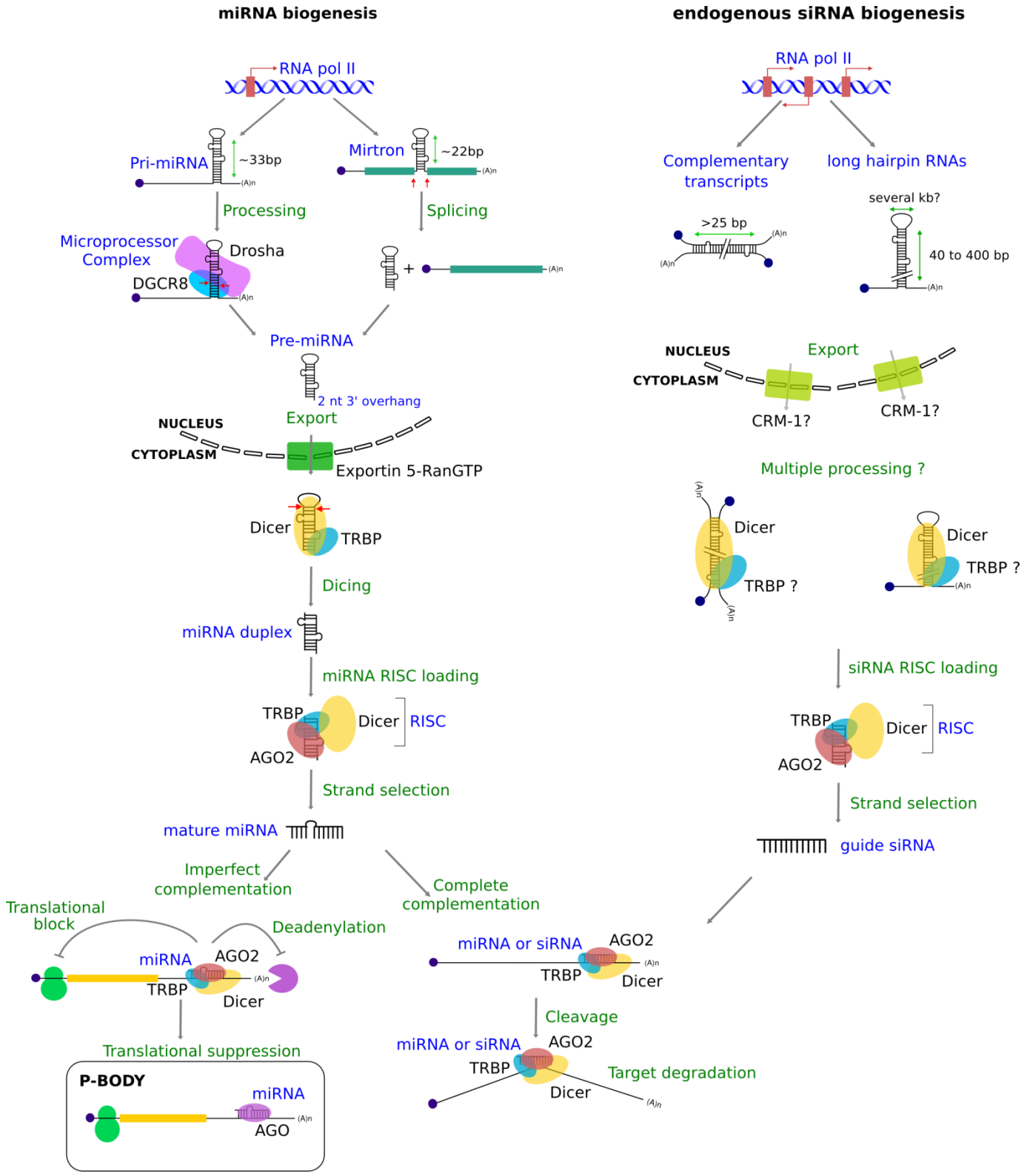

miRNA biogenesis
endogenous siRNA biogenesis
RNA pol II
Pri-miRNA
~33bp
Mirtron
~22bp
Processing
Splicing
Microprocessor Complex
Drosha
DGCR8
Pre-miRNA
2 nt 3' overhang
NUCLEUS
CYTOPLASM
Export
Exportin 5-RanGTP
Dicer
TRBP
Dicing
miRNA duplex
miRNA RISC loading
AGO2
RISC
Strand selection
mature miRNA
Imperfect complementation
Complete complementation
Translational block
Deadenylation
miRNA
Translational suppression
P-BODY
AGO
Complementary transcripts
long hairpin RNAs
>25 bp
several kb?
40 to 400 bp
CRM-1?
Multiple processing ?
TRBP ?
siRNA RISC loading
guide siRNA
miRNA or siRNA
Cleavage
Target degradation

Figure 6. The mammalian RNAi pathway. Endogenous miRNAs and siRNAs begin their genesis from RNA Pol II promoters (shown as red squares). The miRNA biogenesis pathway (left hand panel) is chiefly initiated by transcription of pri-miRNAs from gene-coding regions. Pri-miRNAs have a stem-loop structure approximately 33 bp in length flanked by single stranded regions. The RNAseIII protein Drosha (lilac square) complexed to dsRBD DGCR8 (blue oval) forms a Microprocessor complex that binds the pri-miRNA and cleaves (red arrows) to release a pre-miRNA hairpin structure with characteristic 2 nucleotide 3' overhangs. Pre-miRNAs can also be generated without Drosha. Pri-miRNAs transcribed from intronic sequences are termed mirtrons, have similar characteristics to exonic pri-miRNAs but contain splice donar and acceptor sites that are recognised by host splicing machinery to yield a pre-miRNA in the absence of Drosha. Both forms of pre-miRNA are exported from the nucleus by Exp5 (green square) that recognises and binds the 2 nucleotide 3' overhang. In the cytoplasm, the loop of the pre-miRNA is cleaved (red arrows) by Dicer (yellow oval) and its dsRBD partner TRBP (blue oval) to liberate a miRNA duplex with 2 nucleotide 3' overhangs. AGO2 (pink oval) binds the miRNA duplex to form RISC – a complex that mediates silencing of the target mRNA. The miRNA duplex is unwound leaving the guide strand (mature mRNA) to pair covalently with its cognate mRNA. If this pairing is incompletely complementary, translational suppression occurs in P-bodies. If the mature miRNA pairs 100 % with the target mRNA, AGO2 mediated cleavage of the latter occurs. Endogenous siRNAs (right hand panel) are transcribed either as long hairpin structures with stems between 40 and 400 bp or as duplex structures composed of separate complementary transcripts. It is unclear if they are processed in the nucleus prior to export or if Crm1 (light green squares) mediates their transport to the cytoplasm. There may be multiple processing steps involving Dicer/TRBP to yield a siRNA duplex that is loaded into RISC. The mature strand of the duplex (guide siRNA) pairs with 100 % complementarity to the target mRNA and guides AGO2-mediated cleavage.

Nuclear export, cytoplasmic processing and Argonaute loading

Following nuclear cleavage, pre-miRNAs need to be transported to the cytoplasm for the second processing step to occur (Lee, Jeon et al. 2002). Exportin-5 (Exp5) is responsible for transporting pre-miRNAs and does so by binding its cargo in a sequence-independent fashion at high RanGTP levels in the nucleus, translocating the Ran·Exp5·pre-miRNA complex to the cytoplasm and releasing the pre-miRNA and Ran upon GTP hydrolysis (Lund and Dahlberg 2006). The Exp5 can then return to the nucleus to facilitate another round of transport (Yi, Qin et al. 2003; Bohnsack, Czaplinski et al. 2004). Once in the cytoplasm, the loop of the pre-miRNA is cleaved by another RNAse III enzyme called Dicer to release a miRNA duplex (Bernstein, Caudy et al. 2001). Dicer comprises two copies of a conserved RNAse III domain, a dsRBD in the carboxy-terminus, an amino-terminus helicase domain, a domain of unknown function (DUF) and a PAZ (Piwi-Argonaute-Zwille) domain (Jaskiewicz and Filipowicz 2008). The distance between the PAZ and RNAse III domains within Dicer corresponds well with the average length (25 bp) of a pre-miRNA stem duplex (Macrae, Li et al. 2006). Dicer thus functions as a 'molecular ruler' (Jinek and Doudna 2009) by anchoring the 3' end of the pre-miRNA duplex in the PAZ domain and cleaving at a fixed distance from that end to generate miRNA products 21 bp in length with characteristic 5' phosphate groups and two-nucleotide 3' overhangs [(Elbashir, Harborth et al. 2001; Elbashir, Lendeckel et al. 2001) and section 1.14]. Interestingly, the helicase domain may play a regulatory role as this domain imposes a non-functional

conformation on the whole protein that must be re-arranged for catalysis to occur (Ma, Macrae et al. 2008).

Human Dicer interacts with two proteins, TRBP (TAR RNA binding protein) and PACT (Protein Kinase R activator). TRBP is a dsRBD protein (Chendrimada, Gregory et al. 2005; Daniels, Melendez-Peña et al. 2009), PACT upregulates PKR activity (Lee, Hur et al. 2006) and both proteins contribute to the formation of an RNA-induced silencing complex (RISC). Following Dicer cleavage, the miRNA duplex is released from Dicer but remains bound to TRBP and PACT (Lee, Hur et al. 2006; Daniels, Melendez-Peña et al. 2009). Specifically, the PAZ domain of TRBP recognises the two-nucleotide 3' overhangs in the miRNA duplex and anchors the substrate within RISC (Lingel, Simon et al. 2003; Kawamata, Seitz et al. 2009). A protein from the Argonaute (AGO) family binds TRBP to complete the formation of RISC (Hammond, Bernstein et al. 2000; Zamore, Tuschl et al. 2000; Martinez, Patkaniowska et al. 2002; Preall and Sontheimer 2005). One strand of the miRNA duplex remains bound within RISC as the mature miRNA or guide strand while the remaining passenger strand is degraded (Chendrimada, Gregory et al. 2005). The relative thermodynamic stability of the 5' end of the miRNA duplex determines which strand is the mature miRNA (Schwarz, Hutvagner et al. 2003). Typically, the strand with less stable 5' base pairing remains bound within RISC although this does not apply to all miRNAs (Khvorova, Reynolds et al. 2003; Han, Lee et al. 2006). The endonucleolytic activity of AGO2 is responsible for cleavage of the passenger strand and an unidentified helicase then unwinds and releases this strand (Matranga, Tomari et al. 2005; Preall and Sontheimer 2005; Rand, Petersen et al. 2005).

Target mRNA recognition and processing

Mature miRNAs regulate post-transcriptional protein synthesis by base pairing to cognate mRNAs. Depending on the degree of complementarity between the mature miRNA and its target, as well as the specific AGO protein involved in RISC loading, multiple mRNA silencing modes can occur [reviewed in (Brodersen and Voinnet 2009)]. These include accelerated exonucleolytic mRNA decay, site-specific endonucleolytic cleavage or translational repression (Azuma-Mukai, Oguri et al. 2008; Eulalio, Huntzinger et al. 2008). All four human AGO proteins bind miRNAs although endonuclease activity is a property of AGO2 alone (Liu, Carmell et al. 2004; Meister, Landthaler et al. 2004; Miyoshi, Tsukumo et al. 2005). The initial bases from positions 2 to 7 of the mature miRNA are termed the seed sequence and they provide most of the pairing specificity [reviewed in (Eulalio, Huntzinger et al. 2008; Filipowicz, Bhattacharyya et al. 2009). Complete pairing between the seed region and target mRNA is sufficient to mediate AGO2-associated RISC cleavage of the cognate phosphodiester backbone (Yekta, Shih et al. 2004; Brennecke, Stark et al. 2005). However, target site accessibility and relative

concentrations of the miRNA and its cognate sequence also play key roles in determining RISC-mediated transcriptional regulation (Doench and Sharp 2004; Ameres, Martinez et al. 2007; Kertesz, Iovino et al. 2007; Obernosterer, Tafer et al. 2008). Typically, cleavage activity of siRNA-loaded RISC is severely impaired by mismatched pairing in the seed region and translational inhibition occurs (Martinez and Tuschl 2004; Nilsen 2007) although the mechanism remains unclear (Chekulaeva and Filipowicz 2009). Notably, imperfect pairing can occur multiple times within the same transcript in animals leading to cooperative inhibition of transcription (Saetrom, Heale et al. 2007; Tay, Zhang et al. 2008).

1.9 RNAs in transcriptional regulation

The field of RNAi has revolutionised our understanding of gene regulation. However, homology-dependent RNA interactions are not exclusively linked to RNAi pathways or its protein components. Indeed, transcriptional regulation by non-coding RNAs has been linked to X inactivation, dosage compensation, imprinting, and polycomb silencing [reviewed in (Zaratiegui, Irvine et al. 2007)]. Importantly, these non-coding RNAs are distinct from siRNAs and can range in size from 100 bp to 100 kb [reviewed in (Lee 2009)]. The vast majority of non-coding RNAs are complementary to protein-coding genes and cause inhibition via an antisense mechanism (He, Vogelstein et al. 2008). Interestingly, a small group originate in promoter regions and it has been suggested that these non-coding RNAs may induce transcriptional gene silencing (TGS) through binding of promoters and subsequent heterochromatin formation (Zaratiegui, Irvine et al. 2007). This is poorly understood in mammals, although both exogenous siRNAs and endogenous miRNAs have been shown to induce TGS (Ting, Schuebel et al. 2005; Kim, Villeneuve et al. 2006; Han, Kim et al. 2007; Kim, Sætrom et al. 2008).

Despite the involvement of some RNAi pathway components such as AGO2 and siRNAs, TGS is generally not considered to be part of the canonical RNAi pathway (Hawkins, Santosom et al. 2009). Typically, small promoter-associated RNAs bind the antisense strand of complementary promoter-targeted RNAs (Han, Kim et al. 2007) and induce silencing marks such as di- or tri-methylations of lysine (K) residues on histone 3 (H3K9me2 or H3K27me3). These silencing marks can be detected downstream of the targeted site suggesting a possible mechanism for establishing and spreading of the silent state [reviewed in (Hawkins and Morris 2008)]. AGO1 and AGO2, DNA methyl transferases 3a (DNMT3a) and histone deacetylase 1 (HDAC1) are required to initiate silencing in human cells, and DNMT1 is necessary for maintaining the effect (Janowski, Huffman et al. 2006; Weinberg, Villeneuve et al. 2006; Hawkins, Santosom et al. 2009). Interestingly, the involvement of DNMT1 suggests that the silencing effects may be heritable as this protein is involved in maintaining long-lasting

epigenetic marks during DNA replication and cell division (Weber, Hellmann et al. 2007). In contrast, histone modifications have also been associated with long-term gene activation linked to small RNAs that bind promoter regions with low GC content (Kim, Villeneuve et al. 2006; Li, Okino et al. 2006; Janowski, Younger et al. 2007). This transcriptional gene activation (TGA) effect has also been found to occur in human liver cells where a liver-specific miRNA was shown to positively regulate Hepatitis C Virus (HCV) replication (Jopling, Yi et al. 2005).

Importantly, as TGS can elicit permanent epigenetic modifications, harnessing this phenomenon may offer a mechanism of sustained inhibition of gene expression that is heritable. However, to date little is known about all the factors involved in the TGS pathway or the long-term implications of chromatin remodelling. Post-transcriptional gene silencing (PTGS), the central feature of traditional RNAi, is more completely understood and provides exciting new avenues for harnessing gene expression. Manipulation of the factors involved in canonical RNAi, specifically the miRNA biogenesis pathway, allows researchers to generate a desired silencing effect of a specific target gene at the level of the mRNA. Several strategies exist whereby exogenous RNAi 'mimics' can be introduced into the RNAi pathway at various levels and induce their effects in a manner parallel to that of endogenous RNAi effectors.

1.10 Exogenous RNAi mimics as gene therapy tools

Since its discovery, the exquisite sequence-specificity of the RNAi pathway has been exploited as a tool to uncover the functions of numerous genes in a variety of organisms. This has largely been via the introduction of siRNAs that have provided a robust method to establish strong links between gene identity and gene function. In addition, RNAi has been utilised as a novel gene therapy strategy whereby siRNAs and miRNAs have been targeted to specific genes associated with disease processes [reviewed in (Haussecker 2008; López-Fraga, Wright et al. 2008)]. Typically, this involves the generation of exogenous RNAi effector sequences that mimic their endogenous counterparts. Theoretically it is possible to exploit the RNAi machinery to target any gene of interest and two main strategies can be employed to achieve this, namely synthetic siRNAs and expressed RNAi mimics. The former are introduced directly into the cytoplasm of target cells and serve as RISC substrates. While highly effective even at low concentrations, siRNAs can only induce a silencing effect for up to seven days in cells with a rapid turnover rate and several weeks in slow-diving cells because of dilution or degradation over time (Bartlett and Davis 2006). Repeat administration of siRNAs is feasible but expression of RNAi mimics that enter the pathway as substrates for Drosha or Dicer present a longer-term alternative (Figure 7). Expressed mimics can be transcribed either from RNA Pol II or Pol III promoters depending on the desired downstream effects and their design is affected by the RNAi machinery required to

process each type of mimic. Similarly to endogenous siRNAs, the most critical sequence of any RNAi mimic is the seed region that determines target match. Given that the seed sequence spans only six bp, unintended mRNA targeting [referred to as an off-target effect (OTE)] seems inevitable (Jackson, Burchard et al. 2006). Thus the choice and sequence of mimic as well as target accessibility require thorough consideration if the required silencing effect is to be achieved with a high level of efficacy and minimal unintended seed region pairing. An important objective of this thesis was to explore the development of various RNAi mimics as gene therapy modalities against HIV.

1.10.1 Synthetic RNAi molecules

Synthetic siRNAs were the first RNAi mimics tested (Elbashir, Harborth et al. 2001) and have been used to knock down target genes related to a variety of diseases [reviewed in (Whitehead, Langer et al. 2009)]. As substrates for RISC, synthetic siRNAs are generated as 21 to 23 bp duplex molecules with characteristic two nucleotide 3' overhangs (Grishok, Pasquinelli et al. 2001) and include G:U (guanine : uracil) mismatches in the stem region. These G:U wobbles cause one strand of the duplex to be less stably paired at its 5' end biasing the selection of the preferred strand into RISC (Khvorova, Reynolds et al. 2003; Schwarz, Hutvagner et al. 2003; Seitz and Zamore 2006). Longer duplex molecules (27 bp) have also been used. These enter the RNAi pathway as Dicer substrates (Kim, Longo et al. 2004; Shiota, Ikeda et al. 2006; Nishina, Unno et al. 2008) and may induce more potent inhibition at lower concentrations than synthetic 21 bp siRNAs encoding the same guide sequences (Hefner, Clark et al. 2008). Due to the ubiquitous presence of RNAses within the cellular environment, siRNAs are inherently unstable unless modified [reviewed in (Manoharan 2004; Behlke 2008)]. Traditionally phosphorothioate linkages have been introduced to prevent serum nuclease degradation but 2’ ribose modifications such as 2'-O-methyl (2'-OMe) or 2'-fluoro (2'-F) have the added advantages of enhancing duplex stability, avoiding the innate immune system and reducing OTEs (Jackson, Burchard et al. 2006; Muhonen, Tennilä et al. 2007; Robbins, Judge et al. 2007). Locked nucleic acids included at limited sites within the siRNA backbone have also been shown to improve duplex stability and nuclease protection (Mook, Baas et al. 2007) but may result in hepatotoxicity *in vivo* (Swayze, Siwkowski et al. 2007). Importantly, thorough testing of these compounds is critical as the chemical modifications applied to siRNAs can alter their efficacy or induce an immune response.

1.10.2 Expressed RNAi mimics

Pre-miRNA mimics

Early studies revealed that the guide region of endogenous pre-miRNAs could be substituted with artificial sequence and as long as the two nucleotide 3' overhangs required for Dicer recognition were present, successful inhibition of target mRNA occurred (Zeng, Wagner et al. 2002; Boden, Pusch et al. 2004; Liu, Haasnoot et al. 2007). This led to the development of pre-miRNA mimics transcribed from DNA vectors that had been introduced into target cells of interest. Although Pol II-based plasmids have been tested (Ren, Fang et al. 2007; Giering, Grimm et al. 2008) Pol III-based vectors encoding U6 or H1 transcription elements are more commonly used. These latter promoters express short RNA transcripts with defined 5' and 3' ends that mimic pre-miRNAs and are Dicer substrates (Paul, Good et al. 2002; Tuschl and Borkhardt 2002; Yu, DeRuiter et al. 2002; Cheng and Chang 2007). Referred to as short hairpin RNAs (shRNAs), pre-miRNA mimics generally comprise a sense region (anti-guide) connected to an antisense (guide) counterpart via a loop sequence (McIntyre and Fanning 2006). There is no consensus regarding an optimal shRNA structure but a duplex stem of no less than 19 nt (Li, Lin et al. 2007) with a loop of between 3 and 23 nts yields positive target suppression (Zeng and Cullen 2003; Vlassov, Korba et al. 2007). Similarly to siRNAs, disruption of three or more base pairings in the seed region of shRNAs has a marked deleterious effect on silencing although single point mutations have been shown to reduce but not eliminate target inhibition (Zeng and Cullen 2003).

Importantly, as shRNAs generally comprise a duplex stem of ~23 bp, they are processed by Dicer into single siRNA species and thus function as monotherapies. Owing to the rapid development of resistant HIV strains following the introduction of single therapeutic interventions (see section 1.6) combinatorial approaches need to be devised. Expressed pre-miRNA mimics represent promising candidates for the development of a combinatorial RNAi-based strategy, especially with regard to ease of synthesis and specificity. However, the design of pre-miRNA mimics capable of simultaneously generating multiple anti-HIV siRNAs has not been thoroughly explored.

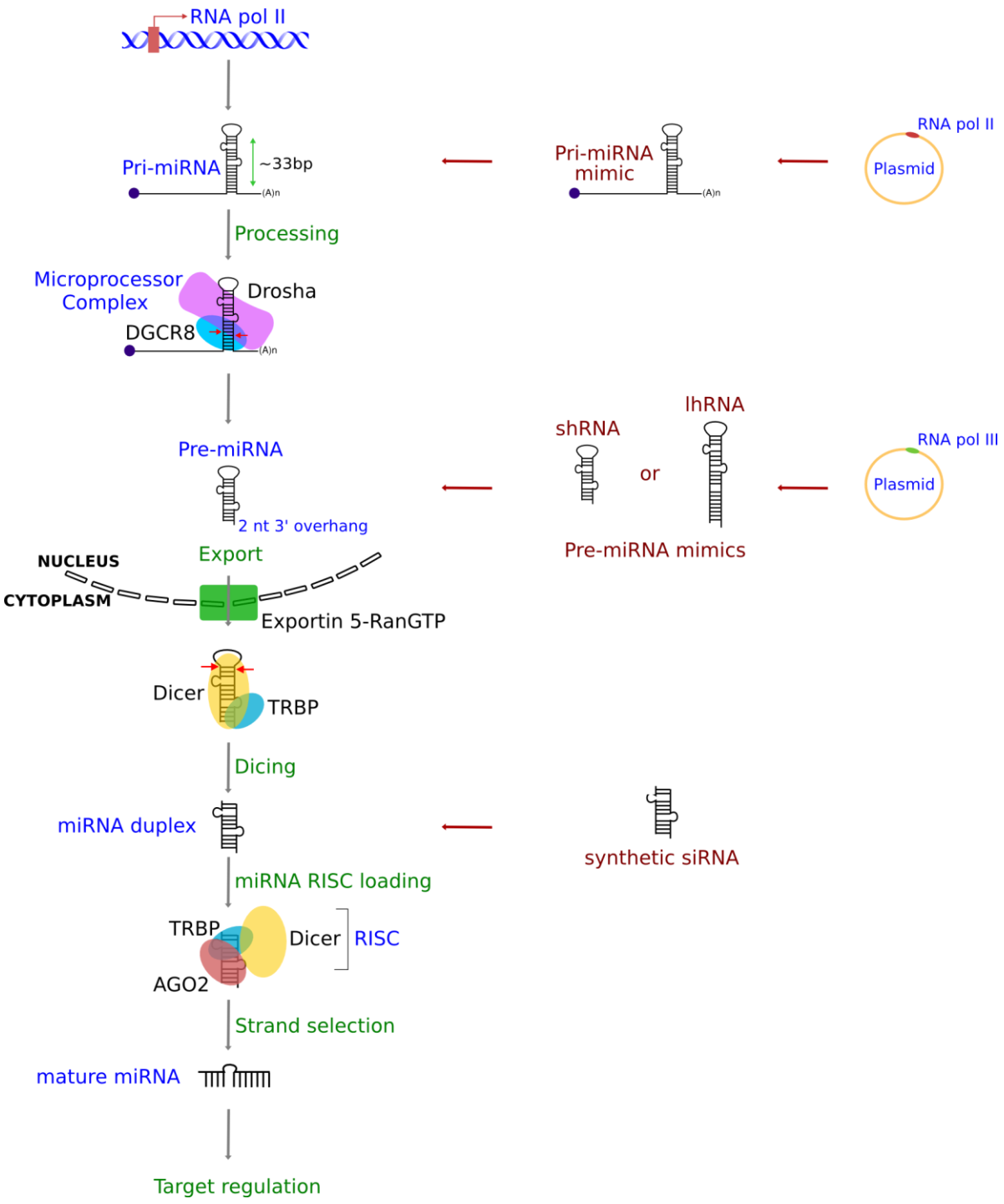


Figure 7. Exogenous mimics in the RNAi pathway. Mimics can enter the RNAi pathway as substrates for Drosha, Dicer or RISC. Pri-miRNA mimics are Drosha substrates that are expressed from RNA Pol II (shown in red) based vectors. Transcription generates a mimic with a stem loop structure and prototypical single stranded flanking sequences that is recognised by DGCR8 and cleaved by Drosha to form a pre-miRNA. RNA Pol III (shown in green) expressed constructs produce stem loop hairpin structures that resemble pre-miRNAs and enter the RNAi pathway as substrates for Dicer. These mimics are termed shRNAs or lhRNAs depending on the length of the duplex stem region. Dicer processes shRNAs into siRNAs of one species while a single lhRNA can be processed to form multiple siRNAs. RISC is the final protein complex in the pathway that interacts with mimics generated from pri-miRNAs, shRNAs, lhRNAs or with synthetic siRNA duplexes introduced directly into the cytoplasm of a cell.

Pri-miRNA mimics

Pol III promoters are constitutively active and while shRNAs expressed from these constructs are potent (Boudreau, Mas Monteys et al. 2008), they may induce lethal toxicities due to saturation of endogenous RNAi pathway components [(Grimm, Streetz et al. 2006) and described further in section 5.3]. One alternative involves the use of Pol III promoters such as tRNA that produce Dicer substrates at lower expression levels than typical cassettes driven by U6 or H1. In addition, tRNAs that are not processed by RNAse Z may be exported by exportin-t and therefore do not saturate exportin-5 transport (Boden, Pusch et al. 2003; Scherer, Frank et al. 2007). Another option involves RNA Pol II promoters that generate pri-miRNA mimics (Chung, Hart et al. 2006; Ely, Naidoo et al. 2008; Zhen, Alhousseynou et al. 2008). As substrates for Drosha these mimics comprise at least 40 bp of ssRNA flanking sequence on either side of the central hairpin structure (Figure 7). In addition, the natural loop sequences and secondary structures within the pri-miRNA ensure efficient transport and downstream processing (McManus, Petersen et al. 2002; Zeng, Wagner et al. 2002).

Several expressed pri-miRNA mimics have been developed whereby endogenous pri-miRNAs have been used as 'scaffolds' for the generation of artificial guide sequences (McManus, Petersen et al. 2002; Zeng, Wagner et al. 2002; Stegmeier, Hu et al. 2005; Chung, Hart et al. 2006). Since pri-miRNAs are naturally expressed from Pol II promoters, tissue-specific or inducible gene regulation of pri-miRNA mimics is possible, allowing for greater control of dosage and cell-specific expression. A pri-miRNA mimic under the control of a heat-shock-responsive promoter induced expression only under conditions of stress in mammalian cells (Yang and Paschen 2008). Separately, the presence of tetracycline in embryonic stem cells containing a pri-miRNA mimic under the control of a tetracycline-inducible promoter switched on expression (Wang, Theunissen et al. 2007). A similar mechanism has been used to generate transgenic mice that express tetracycline-inducible pri-miRNAs targeted to the tumour suppressor Trp53 following doxycycline addition to their drinking water (Dickins, McJunkin et al. 2007).

As Pol II promoters can be regulated by cell-specific environmental factors, they have recently become more popular for the expression of pri-miRNA mimics. However, Pol III-expressed pre-miRNA mimics (shRNAs) remain the most widely used RNAi tools for sustained viral gene suppression, and are the most robust system available for generating sufficient concentrations of siRNAs to induce potent target inhibition. Unlike shRNAs that can be readily generated based on a target sequence of interest, pri-miRNA mimics require knowledge of the endogenous sequence, as well as secondary structures within the folded RNA. Most pri-miRNA studies to date have been based on only a few endogenous backbones (pri-miR31 and pri-miR155) and thus little is known about the efficacy of using other structures as RNAi-based

mimics. Studies that explore the abilities of different endogenous pri-miRNAs to produce exogenous guide sequences will provide valuable information on the design of these mimics. In addition, as mature miRNAs can be excised from either arm of a pri-miRNA stem or even from both arms (Khvorova, Reynolds et al. 2003), results obtained using discrete pri-miRNA scaffolds that generate guide strands from different arms would be or particular importance. Finally, as single interventions against a highly mutable virus such as HIV cannot be seriously considered, data that explores the capacity for pri-miRNA mimics to be incorporated in a combinatorial strategy would be invaluable.

1.11 RNAi-mediated targeting of HIV

Viral infections caused by influenza A virus or respiratory syncitial virus have been successfully treated using siRNAs delivered directly to the lungs of infected patients (Bitko, Musiyenko et al. 2005; Bitko and Barik 2008). This mode of gene therapy is particularly suited to acute infections in the respiratory tract as siRNAs mediate a transitory effect and the upper airways can be readily targeted via aerosol delivery [reviewed in (Berkhout and ter Brake 2009)]. However, for chronic infections such as HIV, constitutive expression of shRNAs is required to ensure lifelong treatment. Theoretically two forms of HIV RNA can be targeted by RNAi. These include incoming viral genomic RNA and newly synthesised viral transcripts produced from integrated provirus. In addition, host factors required for HIV infection present alternative sites for RNAi-mediated inhibition. However, not all viral RNAs are equally sensitive to RNAi and inhibition of cellular targets may be detrimental to the host.

1.11.1 Incoming viral genomic RNA as a target

Incoming, uncoated HIV RNA is an enticing target for RNAi-mediated inhibition. Blocking viral replication at this stage would prevent proviral integration and subsequent viral gene expression and replication. There is still some debate as to whether incoming RNA is accessible to RNAi-mediated attack. Several early studies suggested that incoming RNA could be targeted using siRNAs (Capodici, Kariko et al. 2002; Coburn and Cullen 2002; Jacque, Triques et al. 2002). However, the viral mRNAs in these studies were not introduced directly into the cell but rather transcribed from a plasmid co-transfected with the siRNAs and this is significantly different to incoming viral RNA that is bound to various proteins. Incoming viral RNA, represented by two viral RNA genomes, is continually bound by proteins such as p17 matrix, integrase and reverse transcriptase (Dvorin and Malim 2003). This association may limit access of the RNAi machinery to the target HIV genomic RNA. Also, following membrane fusion between the host cell and

virion, the viral core particle is only partially dissolved while the RNA genome is reverse transcribed and transported towards the nucleus. Furthermore, reverse transcription usually occurs within 6 hours of viral entry into lymphocytes (Zack, Arrigo et al. 1990), providing a very short window of opportunity for RNAi-mediated inhibition of incoming viral RNA. Indeed two thorough investigations using synthetic siRNAs and shRNAs against a wide range of HIV targets indicated that the newly-released viral genome remains encapsidated or partially-encapsidated during reverse transcription, thus blocking access to guide strand-loaded RISC (Westerhout, ter Brake et al. 2006; Gao, Lobritz et al. 2008). These studies indicate that targeting of viral mRNAs is a better approach to achieving RNAi-mediated HIV inhibition.

1.11.2 Inhibition of HIV mRNAs and viral escape

Targeting of viral mRNA species transcribed from integrated provirus has been more successful and RNAi strategies have been used against every coding and non-coding genomic sequence to inhibit viral replication. The *rev* ORF was one of the first viral genes to be targeted using siRNAs and shRNAs (Coburn and Cullen 2002; Lee, Dohjima et al. 2002). These initial investigations showed that viral transcripts generated early during the life cycle were accessible to RNAi-mediated inhibition, but not all sites were susceptible to silencing. In addition, some non-specific decrease in *tat* gene expression was also reported but this was due to overlap between the *tat* and *rev* ORFs. The effect was alleviated by designing siRNAs to target non-overlapping portions of the coding sequences (Coburn and Cullen 2002). Further information on HIV target site accessibility was gathered using siRNAs/shRNAs against the TAR stem-loop of the 5' LTR. The TAR target was found to be refractory to RNAi knockdown and this data highlighted that secondary folding of RNAs could render them inaccessible to siRNAs (Jacque, Triques et al. 2002; Yoshinari, Miyagishi et al. 2004; Nishitsuji, Kohara et al. 2006). Notwithstanding the TAR loop, the HIV LTR remains an attractive target particularly as non-coding RNAs and siRNAs have been shown to shut down gene expression by targeting promoter sequences (section 1.9). This presents an especially attractive alternative to traditional post-transcriptional gene silencing, as chromatin changes associated with transcriptional gene silencing could cause the provirus to become stably heterochromatised. Intriguingly, the potential for TGS-mediated inhibition of the LTR has not been explored and offers an interesting prospect for a gene therapeutic against HIV.

Single nucleotide substitutions within target mRNA, especially in the seed region, are sufficient for viral protection from RNAi effectors (Gitlin, Karelsky et al. 2002). In one example, a shRNA targeted to *tat* mRNA became ineffective after 25 days and viral genome analysis showed that a single base substitution within *tat* was responsible for the resistance (Boden,

Pusch et al. 2003). Similarly, escape mutants appeared following several weeks of exposure to a shRNA targeting *nef* mRNA. Further analysis revealed that base substitutions or deletions within *nef* modified the shRNA target (Das, Brummelkamp et al. 2004). These two examples involved targeting of non-essential viral genes but designing siRNAs against essential genes also resulted in emergence of resistance (Senserrich, Pauls et al. 2008; von Eije, ter Brake et al. 2008). Inhibition of infection with three different HIV primary isolates in the presence of a single siRNA targeted to overlapping *env*/*rev* mRNA was not absolute. Escape mutants bearing two point mutations in the target sequence readily evolved (Senserrich, Pauls et al. 2008). The use of shRNAs targeted to essential viral sequences did not arrest all viral escape, although there was restriction of escape possibilities. Mutations were limited to non-silent changes that presumably allowed RNAi resistance while maintaining viral fitness (von Eije, ter Brake et al. 2008).

Viral escape from RNAi does not always involve target sequence mutations. Sequencing of two viral mutants replicating in the presence of *nef*-specific shRNAs revealed substitution mutations upstream of the siRNA target. Resistance was caused by alterations in the viral mRNA secondary structures which probably prevented access of activated RISC to its cognate sequence (Westerhout, Ooms et al. 2005). Even more unexpected was the emergence of escape through the modulation of transcription. Mutations in non-targeted LTR promoter sequences have been shown to compensate for RNAi-mediated inhibition by up-regulating viral gene transcription (Leonard, Shah et al. 2008). Notwithstanding these additional factors associated with resistance, focus has shifted to targeting conserved vital sequences within the HIV genome, since HIV is less able to mutate these sites without loss of fitness (Han, Wind-Rotolo et al. 2004; Morris, Chung et al. 2004; Chang, Liu et al. 2005; Lee 2005; Naito, Nohtomi et al. 2007; von Eije, ter Brake et al. 2008). Further evidence of the extent of RNAi tolerance to sequence polymorphisms emerged from a study using expressed shRNAs against three targets: a highly variable *rev* sequence, a *gag* site conserved across subtype B strains only, and a *vif* sequence conserved across all subtypes (Lee 2005). *Rev*-specific shRNAs inhibited the least number of isolates, regardless of subtype. The *gag*-specific shRNAs protected against subtype B virus alone and the anti-*vif* shRNAs were effective against all isolates (Lee 2005). These results are particularly intriguing as they suggest that many separate strains of HIV may be silenced by targeting conserved sequences. This is particularly relevant in the context of wild type HIV infection where the viral strains are not isogenic but rather comprise a pool of similar but non-identical sequences. Studies that explore the capacities of RNAi mimics to inhibit conserved HIV sites when assessed against wild type HIV isolates will be extremely valuable.

HIV dosage has also been shown to play a key role in determining whether RNAi-mediated inhibition is effective. Specifically, shRNAs designed to target HIV IN showed some

inhibition of HIV replication when shRNA-expressing CD4+ T cells were challenged with virus at a low dose (100 pg of p24) of infection (Nishitsuji, Kohara et al. 2006). This effect was abrogated during long-term (22 days) culture or following a high-dose (10 000 pg of p24) infection (Nishitsuji, Kohara et al. 2006). Similarly, increased dosages of HIV were used to select for various anti-HIV shRNAs that displayed robust and sustained inhibitory effects in the presence of highly concentrated virus (von Eije, ter Brake et al. 2009). Together, this data suggests that although highly conserved viral sequences have allowed for the discovery of many effective anti-HIV siRNAs/shRNAs (Naito, Nohtomi et al. 2007; ter Brake, Hooft et al. 2008; McIntyre, Groneman et al. 2009), the long-term inhibitory effects against actively evolving HIV infection need to be evaluated (von Eije, ter Brake et al. 2008).

1.11.3 RNAi-mediated inhibition of host cell targets required for HIV infection

Although using RNAi to target viral sequences is important, it is clear that targeting viral sequences alone may not be sufficient to ensure long-term viral suppression. Since HIV requires a large number of cellular host factors for successful replication, these are potential therapeutic targets for RNAi-mediated inhibition. A significant advantage of targeting host genes is that they are not under selective pressure, and therefore to gain resistance, HIV would need to evolve an entirely new set of cellular associations to bypass the requirement for an important host factor. The major CD4 receptor required for HIV docking was one of the first host cell sequences targeted using siRNAs (Novina, Murray et al. 2002). While these results were promising, there has been a shift away from targeting CD4 as a result of the ubiquitous presence of this receptor on various host cells and its function in normal immune pathways. The chemokine co-receptors required for HIV entry offered more attractive alternatives and there have been numerous studies describing RNAi-mediated suppression of CCR5 (Martinez, Gutierrez et al. 2002; Qin, An et al. 2003; Song, Lee et al. 2003) and CXCR4 (Martinez, Gutierrez et al. 2002; Anderson, Banerjea et al. 2003). CCR5 was particularly enticing as a deletion mutant allele was found to be present at a frequency of 0.092 in Caucasian populations, and individuals homozygous for the deletion are highly resistant to M-tropic HIV infection (Samson, Libert et al. 1996). Unfortunately this strategy remains inadequate for complete protection against HIV-1 as T-tropic and dual tropic strains can still gain access to cells. One study using HIV-2 showed CXCR4 to be sufficient for entry even in the absence of CD4 (Endres, Clapham et al. 1996). Furthermore, although less common, there exist primary HIV isolates that utilize CCR3 and CCR2b as co-receptors to gain access to CD4+ cells (Choe, Farzan et al. 1996; Doranz, Rucker et al. 1996). RNAi-mediated inhibition of expression of other proteins involved in HIV transcription, such as

NFκβ, cyclin T1, CDK9 and TRBP have also been shown to down-regulate viral replication (Surabhi and Gaynor 2002; Chiu, Cao et al. 2004; Christensen, Daher et al. 2007).

Differences in the viral life cycle in T cells and in macrophages present a further challenge to the development of RNAi-based HIV therapy. Specifically, in T cells, nascent viral mRNAs transcribed from the provirus are the most attractive targets for RNAi as they are shuttled to the cytoplasm by Rev. However, in macrophages new virions do not assemble and bud at the plasma membrane. Instead viruses may bud into endocytic organelles and accumulate in these vacuoles before being released as an exosome. The advantage for HIV is that exosomes escape immune surveillance but importantly these organelles are also inaccessible to antiviral therapies and RNAi-based gene silencing [reviewed in (Verani, Gras et al. 2005)]. Thus the identification of host factors that are required for HIV infection, and which are common to all cells infectable by the virus, has become more important for the development of gene expression-based antiviral silencing approaches. Recent advances in high-throughput screens using robotic instrumentation and genome-wide siRNA libraries have resulted in novel reverse genetics applications to monitor host-pathogen interactions. Last year alone, three influential papers were published describing the use of large-scale RNAi screens to uncover host dependency factors (HDFs) required for HIV replication [(Brass, Dykxhoorn et al. 2008; Konig, Zhou et al. 2008; Zhou, Xu et al. 2008) and described in section 5.1]. The genes identified in all three studies have established novel RNAi targets that may also serve as potent drug targets. However, further investigation will be required to verify a genuine role of HDFs in a physiological context.

Another important consideration is the delivery method required to introduce anti-HDF siRNAs to their target sites. Lentiviral vectors are one of the most promising candidates for delivering and stably expressing anti-HIV RNAi payloads [reviewed in (Morris and Rossi 2006)]. Retroviral and mobilisation-competent vectors have been used to deliver siRNAs or shRNAs targeted to NF-κβ sites within the HIV 5' LTR (Turner, De La Cruz et al. 2009; Yamagishi, Ishida et al. 2009). Long-term silencing (1 to 12 months) was observed and chromatin immunoprecipitation assays (ChIP) demonstrated rapid, sustained induction of heterochromatin structures enriched for histone silencing marks. In addition, no evidence of viral resistance was observed suggesting that RNAi-mediated transcriptional silencing of the HIV promoter could provide a long-term therapeutic strategy. However, no studies have been completed using promoter-directed RNAi against wild type HIV or high doses of the virus and related pre-clinical studies in animal models of HIV infection have yet to be published.

1.12 Pre-clinical studies for the therapeutic application of RNAi

Current gene therapy protocols have focused mainly on achieving long term suppression of HIV by using retroviral or lentiviral vectors that generate stably expressed RNAi transgenes. Viruses are capable of delivering payloads to a variety of host cells depending on the tropism of the viral envelope proteins. Lentiviruses are particularly useful in this regard as they have the capacity to transduce non-dividing cells and integrate within the host genome thereby ensuring cellular-mediated transcription of the provirus [reviewed in (Morris and Rossi 2006)]. In a seminal paper, primary haematopoetic progenitor cells transduced by lentiviral vectors encoding anti-*rev* siRNAs underwent successful maturation into mature macrophages or T cells (Banerjea, Li et al. 2003). Importantly, stable transduction of the progenitor cells resulted in marked resistance to wild type HIV infection that was characteristic of the mature cells as well. These *in vitro* data were further assessed using a humanised mouse model of HIV infection. In this model lentiviral vector-transduced human haematopoetic stem cells (HSCs) are used to reconstitute the T-lymphoid compartment of SCID (severe combined immunodeficiency) mice that do not express immune cells (Akkina, Rosenblatt et al. 1994). Engraftment of siRev-transduced progenitor cells into these mice (SCID-hu) resulted in reconstitution of thymocytes that matured and responded to mitogenic stimulation as expected for normal T cells. Encouragingly, significant viral resistance was also observed following *ex vivo* HIV challenge of differentiated, transduced macrophages and T cells collected from the mice (Banerjea, Li et al. 2003).

While useful, the SCID-hu mouse model has low engraftment levels with limited systemic spread of HIV following infection. This has prompted subsequent research to be performed using modified SCID mice. One model is the non-obese diabetic/SCID (NOD/SCID) mouse which harbours a complete null mutation in the common cytokine receptor γ chain that abrogates interleukin 2 binding (NOD/SCID/IL2rγ^{null}) (Ishikawa, Yasukawa et al. 2005). In these mice T and B cell development is disrupted as the γ chain receptor is required for innate immunity (Watanabe, Ohta et al. 2007). An elegant study used single chain antibodies (scFv) against the CD7 surface antigen present on human T cells as a highly specific delivery vehicle for anti-HIV siRNAs in NOD/SCID/IL2rγ^{null} mice (Kumar, Ban et al. 2008). Anti-*vif* and anti-*tat* siRNAs conjugated to scFvs via an arginine tag and introduced into both NOD/SCID/IL2rγ^{null} and SCID-hu mice successfully prevented HIV infection and prevented CD4+ T cell loss without any observed toxicity. Using a mouse model derived from a recombination activating gene (RAG)–deficient strain that also lacks a γ chain receptor (BALB-Rag2^{null}γ^{null}), HSCs tranduced with a lentivector encoding an anti-*nef* shRNA showed significant inhibition of HIV replication *ex vivo* (ter Brake, Legrand et al. 2009).

All of the mouse models described above have demonstrated the utility of lentiviral vector deliver of anti-HIV RNAi mimics into HSCs, providing a foundation for similar approaches in human patients. Using a non-human primate model, successful stable T-cell expression of shRNAs directed to CCR5 in rhesus macaques has been reported (An, Donahue et al. 2007). Following transplantation of rhesus HSCs, consistent down-regulation of CCR5 expression was observed for up to 14 months in mature T cells. These cells were also less susceptible to SIV infection *ex vivo* (An, Donahue et al. 2007). Concerns remain that macaque and murine models of HIV infection utilise marginally matured progenitor cells as opposed to true stem cells and long-term adverse effects may occur (ter Brake, Legrand et al. 2009). Despite this, a number of human clinical trials assessing anti-HIV gene therapies have been conducted or are currently underway [reviewed in (Rossi, June et al. 2007)].

1.13 Anti-HIV RNA-based clinical studies

A number of anti-HIV gene therapies including ribozymes, RRE decoys, transdominant Rev proteins and RNAi have all been tested in human phase one or two clinical trials with varying efficacies depending on the transfer vector used and whether T cells or HSCs were transduced (Woffendin, Ranga et al. 1996; Kohn, Bauer et al. 1999; Macpherson, Boyd et al. 2005; Mitsuyasu, Merigan et al. 2009). The first example of RNAi in a stem cell setting is currently being conducted by John Rossi and collaborators at the City of Hope Medical Centre in Duarte, California. This trial involves a pilot study of transduced HSCs in AIDS lymphoma patients to assess the safety and therapeutic feasibility of using a lentiviral vector encoding three anti-HIV modalities (Li, Li et al. 2006). The study was initiated when a lentiviral construct expressing a snoRNA nucleolar localising Pol III (U6)-driven TAR RNA decoy, a U6-expressed shRNA complementary to the *tat*/*rev* overlapping region and a Pol III (VA1)-promoted anti-CCR5 ribozyme was shown to successfully transduce primary HSCs (Li and Rossi 2005) and induce long-term inhibition of HIV in these cells (Li, Kim et al. 2005). Vector-transduced HSCs were also able to differentiate normally into different lymphocyte lineages without evidence of untoward effects. Protected HSCs gave rise to phenotypically normal T cells that were HIV resistant when engrafted into SCID-hu mice (Anderson, Li et al. 2007). For the human trial, bone marrow lymphocytes were removed from four patients and cultured *ex vivo* to enrich for progenitor cells. Following enrichment, the HSCs were transduced with the lentiviral vector before autologous reintroduction into patients who had undergone bone marrow ablation (Anderson, Li et al. 2007). Outcomes from these studies are keenly awaited and will be of importance for refinement of this exciting new technology. Importantly, the emergence of resistant HIV strains will be closely monitored as this presents a daunting challenge that confounds all anti-HIV therapies. HIV

resistance was observed very early on following the introduction of conventional ARV monotherapies prompting the development of combinations of drugs in the form of HAART. Viral resistance to single RNA-based gene therapy approaches have also been observed and transdominant proteins, RNA decoys and ribozymes have all been combined in strategies akin to HAART (Gervaix, li et al. 1997; Barnor, Miyano-Kurosaki et al. 2005; Li, Kim et al. 2005; Asparuhova, Barde et al. 2008). However, very few attempts have been made using RNAi mimics alone. Multi-target gene silencing presents an advantage over traditional HAART in that there are many conserved RNAi-susceptible sequences within the HIV-1 genome. This emerging field of gene therapy, referred to as combinatorial RNAi, has provided a plethora of research opportunities and this thesis specifically focuses on the simultaneous generation of multiple RNAi mimics targeted to HIV.

1.14 Combinatorial RNAi

The aim of any combinatorial anti-HIV strategy is to inhibit multiple target sites thereby maximising the overall suppressive efficacy to decrease the chance of escape mutants. The target sites could be conserved viral sequences and HDFs. At this stage it is premature to predict the long-term efficacy of multiple targeting methods, however it has been proposed that assuming equal and effective inhibition, the simultaneous targeting of at least four RNAi-susceptible sequences may be sufficient to prevent or, at the very least, severely delay the emergence of RNAi-resistant variants (Wiethoff and Middaugh 2003; Leonard and Schaffer 2005; Berkhout and Haasnoot 2006; Berkhout and ter Brake 2009). There are several strategies for inhibiting multiple target sites on HIV-1, the most promising of which involve the use of expressed RNAi mimics.

1.14.1 Multiple shRNA expression

Short hairpin RNA expression cassettes using RNA Pol III promoters (U6, H1, tRNA and VA1) can be placed adjacent to each other within a single delivery vector for the combined expression of multiple shRNAs. In an early study, shRNAs targeted to HIV integrase (shIN) showed potent short-term inhibition of HIV although shIN-resistant strains soon emerged (Nishitsuji, Kohara et al. 2006). Subsequently, a combination of expressed shRNAs designed to target wild type and escape mutant sequences showed an overall decrease in antiviral activity. This was possibly a result of diminished siRNA access to their cognate sequences and suggested a potentially detrimental effect when administering multiple shRNAs targeted to the same site (Nishitsuji, Kohara et al. 2006). Strategies which have targeted up to three separate conserved sites have

yielded more promising results (Anderson, Banerjea et al. 2003; ter Brake, Konstantinova et al. 2006; Song, Giang et al. 2007). However, most show the emergence of viral escape in culture, although at a delayed rate when compared to single shRNAs alone. In a seminal study, a single lentiviral vector was used to produce stably transduced T cells containing four tandem anti-HIV shRNA expression cassettes (ter Brake, Hooft et al. 2008). This multi-shRNA approach prevented emergence of escape mutants over an 80 day period in culture. However it was noted that some low-level viral replication occurred in the presence of the inhibiting multi-RNAi mimics, which was attributed to the lack of silencing of incoming RNA and incomplete viral RNA suppression. As this study assessed efficacy against a single molecular clone of HIV, it remains to be determined whether sustained repression of a more diverse viral population (such as primary virus from an infected individual) is possible. Additionally, different promoters were used to produce each shRNA, since multiple copies of the U6 promoter were subject to deletion during lentiviral recombination (ter Brake, Hooft et al. 2008). It has been noted that the additive effect of using multiple promoters in tandem results in an overall decrease in guide strand production (ter Brake, Konstantinova et al. 2006; ter Brake, Hooft et al. 2008). Also, a combinatorial system with heterogenous promoters produces different amounts of each RNAi hairpin precursor and subsequent guide sequences. Theoretical computational analyses of escape kinetics highlighted the importance of ensuring efficient delivery to HIV-1 susceptible cells and potent silencing by each individual RNAi mimic within a multi-targeting system (Leonard and Schaffer 2005). Finally, the introduction of multiple Pol III promoters has also been linked to cellular toxicities as they produce saturating concentrations of shRNAs that compete with endogenous pre-miRNAs for access to Exp5, Dicer or RISC [(Grimm, Streetz et al. 2006) and section 5.3].

1.14.2 Long-hairpin RNAs

Early work exploring the requirements for Dicer-mediated RNA cleavage revealed that dsRNAs with 21 bp duplexes were not processed by recombinant human Dicer *in vitro* (Provost, Dishart et al. 2002). In contrast, *Drosphila melanogaster* Dicer could cleave dsRNAs of 30 bp (short) to 130 bp (long) with equal efficiency *in vivo* (Elbashir, Lendeckel et al. 2001; Zhang, Kolb et al. 2002). Substrates with 3' overhangs were shown to be more efficiently processed than blunt-ended dsRNAs by human recombinant Dicer *in vitro* (Zhang, Kolb et al. 2002; Zhang, Kolb et al. 2004; Vermeulen, Behlen et al. 2005). Furthermore, human recombinant Dicer was shown to preferentially cleave long dsRNAs at their termini as opposed to their loop ends *in vitro* (Zamore 2001; Zhang, Kolb et al. 2002). Multiple cleavage products, including ~22 bp siRNAs, were detected when purified Dicer was incubated with 130 bp dsRNAs (Zhang, Kolb et al. 2002).

While this was suggested to be a result of more than one Dicer molecule binding co-operatively to each long dsRNA, no supporting data was provided. An alternative explanation could be that a single Dicer molecule may be able to cleave a single long dsRNA multiple times. This hypothesis has formed the premise for the development of long-hairpin RNAs as substrates for Dicer-mediated cleavage (Akashi, Miyagishi et al. 2005; Nishitsuji, Kohara et al. 2006). Theoretically, if Dicer could produce multiple siRNAs from a single expressed lhRNA substrate this would provide a powerful combinatorial RNAi-based strategy for inhibiting numerous targets. This has been a controversial application as a major concern of working with longer (> 30 bp) dsRNAs in mammalian cells involves the potential activation of an innate immune response [(Ui-Tei, Zenno et al. 2000; Elbashir, Harborth et al. 2001) and section 5.3]. Synthetic siRNAs encoding particular sequence motifs induce cellular immune stimulation when delivered naked or in lipid vehicles, but the same sequences are non-stimulatory when expressed endogenously (Robbins, Li et al. 2006). Similarly, long (800 bp) dsRNAs expressed endogenously as head-to-head inverted repeats did not induce an immunostimulatory response in various mammalian cell lines (Diallo, Arenz et al. 2003). Clearly, long dsRNAs can be processed by Dicer in the absence of an unwanted innate immune response. However, prior to the completion of this thesis, limited studies exploring the potential for lhRNAs to simultaneously produce multiple effective siRNAs had been published. Expressed lhRNAs with duplex stem regions of 50 to 197 bp were shown to suppress viral replication of HCV and HIV (Akashi, Miyagishi et al. 2005; Watanabe, Sudoh et al. 2006). All three studies utilised lhRNAs with G:U mismatches in the duplex regions and suggested that these modifications prevented immunostimulation. Notably, these studies used lhRNAs targeted to single continuous viral sequences, making it very difficult to identify the relative inhibitory contributions of each potential siRNA. In addition, no analyses were completed to determine how many functional siRNAs were produced from each of the lhRNAs. While promising, more research is required to elucidate how Dicer processes these constructs, what sequences can be targeted using lhRNAs and if additional sites within HIV are amenable to these RNAi mimics.

1.14.3 Polycistronic pri-miRNA mimics

The polycistronic nature of many endogenous pri-miRNAs provides a naturally-occurring combinatorial RNAi-based system that could be readily mimicked for anti-viral purposes. It has been shown that the guide strand of endogenous single pri-miRNAs can be successfully substituted with an artificial sequence capable of inhibiting its cognate target. In addition, individual Pol II promoters were shown to drive expression of up to eight tandem copies of one type of pri-miRNA scaffold (Zhou, Xia et al. 2005; Chung, Hart et al. 2006; Sun, Melegari et al.

2006). These studies provided important proof-of-concept data showing that single pri-miRNA scaffolds could be combined in one system capable of generating multiple exogenous guide sequences. Some data has also been published to show that endogenous polycistronic pri-miRNA clusters could produce exogenous siRNAs although not all of these were functional (Aagaard, Zhang et al. 2008; Liu, Haasnoot et al. 2008). These studies offer valuable information regarding the potential for polycistronic pri-miRNA mimics as combinatorial RNAi effectors but further investigation is needed to address important questions relating to the design of these mimics. For example, which endogenous pri-miRNA backbone represents the best scaffold for guide sequence substitution? Different endogenous pri-miRNAs vary significantly in their overall secondary structures and thus may exhibit discrete abilities to process exogenous guide sequences. Furthermore, a recent study has shown that pri-miRNA-like hairpins regulate the Microprocessor complex (Drosha/DGC28) suggesting that substrates other than canonical pri-miRNAs feed into the RNAi pathway at this level. As pri-miRNAs may produce functional guide sequences from either arm of the duplex, this could have an effect on the production of artificial guide sequences. Related to this, studies that compare the efficacies of exogenous guides produced from the 5' arm versus the 3' arm of the same pri-miRNA scaffold would be valuable. In addition to natural polycistronic pri-miRNAs, discrete scaffolds could be combined in a single expression system thereby providing a modular construct capable of targeting a combination of genes of interest. Data demonstrating the inhibitory efficacies of various pri-miRNA backbones in differing combinations is required and presents another interesting avenue of research. Elucidation of all these areas is of particular value in the context of HIV as the answers may provide another combinatorial RNAi tool for limiting escape mutants.

1.15 Thesis objectives

Monotherapies against HIV are ineffective, leaving combinatorial approaches as the only viable option. The ability of RNAi mimics to be multimerised provides a significant advantage over other gene therapy strategies especially when targeted to a highly mutable virus such as HIV. Multiply expressed RNAi mimics can be designed in a number of ways although lhRNAs and pri-miRNA mimics present the best two options due to their modularities, limited promoter requirements and abilities to remain highly effective at low doses. Concurrent to the initiation of this thesis, only limited data outside of the field of HIV was available and no research had been published on co-expressed anti-HIV pri-miRNA mimics thereby leaving many unanswered questions. Furthermore, little data was available on the HIV LTR as a target for RNAi.

Both lhRNAs and pri-miRNA mimics provide a post-transcriptional mechanism of gene suppression that can theoretically be targeted to any transcribed portion of the HIV genome. In

addition, the use of anti-LTR asRNAs are useful in transcriptional gene silencing. Thus the aims of this thesis were to explore two very different combinatorial RNAi strategies, namely lhRNAs and pri-miRNA mimics, using HIV as a target, as well as to investigate the possibility of targeting the viral LTR promoter with the intention of inducing transcriptional and post-transcriptional inhibition. Importantly, while preventing viral escape is a long-term goal of any anti-HIV therapeutic strategy, the focus of this thesis was to generate various combinatorial RNAi effectors that could produce multiple highly efficacious anti-HIV guide sequences capable of specific cognate target silencing. The results generated from this work significantly improved the understanding of modular RNAi mimics, provided new information on design approaches that yielded compelling suppression of HIV, and uncovered vital considerations for RNAi-based targeting of the viral promoter. This thesis encompassed the following aims:

1. The transcribed region of the HIV 5' LTR was investigated as a target for RNAi-mediated inhibition using multiple Pol III-expressed long-hairpin RNAs with 63 bp duplex stems. In addition, the un-transcribed region of the LTR was assessed as a target for antisense mediated inhibition by siRNAs.
2. The efficacies of siRNAs targeted to the U3 region of the 5' LTR of HIV were assessed for their ability to induce transcriptional gene silencing of the viral promoter. Histone modifications in the siRNA targeted regions were analysed to uncover the state of epigenetic silencing and assess if the silencing effect had migrated up- or downstream of the target sites.
3. The potential of the anti-LTR lhRNAs to inhibit processive and non-processive LTR-expressed transcripts was analysed using various HIV subtypes. Of particular interest was the use of a potent primary isolate of HIV to determine the abilities of anti-HIV lhRNAs to suppress varied target sequences present during wild type infection.
4. The ability of Dicer to process anti-HIV long-hairpins RNAs was investigated to establish if multiple functional siRNAs were produced from the duplex stem regions. Furthermore, as long dsRNAs had been shown to induce unwanted innate immune-stimulation, cellular immune responses following the introduction of anti-HIV lhRNAs were analysed in mammalian cells.
5. The inhibitory efficacies of Pol II-expressed anti-HIV polycistronic pri-miRNA mimics were assessed in order to elucidate the potential of these mimics as combinatorial RNAi effectors. Each polycistronic pri-miRNA comprised four discrete endogenous scaffolds designed to inhibit multiple HIV target mRNAs. A modular approach was used to investigate if the different endogenous pri-miRNA backbones affected the efficacies of anti-HIV guide sequences targeted to reporter vectors as well as wild type HIV.

Chapter two: expressed long hairpin RNAs targeted to the HIV LTR as a combinatorial RNAi strategy

2.1 Introduction

Traditionally expressed shRNAs have been designed with stem lengths of ~23 bp that are processed by Dicer to form a single guide strand capable of targeting one mRNA sequence. The use of RNA Pol III promoters to drive expression of shRNAs has been favoured as this allows for constructs with defined 5' and 3' ends required for Dicer recognition to be generated (Paul, Good*, et al.*, 2002; Tuschl and Borkhardt, 2002; Yu, DeRuiter*, et al.*, 2002; Cheng and Chang, 2007). RNA Pol III promoters are required *in vivo* for the constitutive and ubiquitous synthesis of ribosomal 5S rRNA, tRNA and other small RNAs in most cell types and therefore modifications for translation are not included in Pol III generated transcripts [for more detail on eukaryotic RNA polymerases see section 3.1 and (Cramer, Armache*, et al.*, 2008)]. Dicer has been shown to cleave longer dsRNAs (> 30 bp) into functional siRNAs capable of inhibiting their cognate targets (Elbashir, Lendeckel*, et al.*, 2001; Provost, Dishart*, et al.*, 2002; Zhang, Kolb*, et al.*, 2002). This paved the way for the development of long-hairpin RNAs (lhRNAs) that function as pre-miRNA mimics and thus serve as Dicer substrates. The approach has been used against HCV with expressed lhRNAs that contained stem regions of 50 to 197 bp (Akashi, Miyagishi*, et al.*, 2005; Watanabe, Sudoh*, et al.*, 2006). Akashi *et al* utilised U6- or $tRNA^{VAL}$-expressed lhRNAs with 50 or 100 bp stem regions that had been modified to contain G:U wobbles (Akashi, Miyagishi*, et al.*, 2005). The shorter of the two lhRNAs showed HCV target-specific inhibition in the absence of a cellular immunostimulatory response as assessed by the induction of members of the interferon (IFN) suite of genes. In addition, Northern blot analysis detected siRNAs from two adjacent positions along the stem of the duplex although the sequences detected nearest the loop of the hairpin were at a significantly lower concentration. Watanabe *et al* obtained similar inhibitory data with U6-expressed lhRNAs comprising 197 bp but no work was completed to detect processed products from different regions of the lhRNAs (Watanabe, Sudoh*, et al.*, 2006). The success of targeting HCV using expressed lhRNAs suggested that these mimics could be efficacious against other viruses such as Hepatitis B Virus (HBV) or HIV. A single study using expressed shRNAs targeted to a conserved region of HIV *integrase* demonstrated RNAi-mediated suppression but viral escape occurred within a number of days rendering the shRNAs ineffective (Nishitsuji, Kohara*, et al.*, 2006). U6-expressed lhRNAs comprising 50 bp stem regions were designed to target the same site. While these constructs effectively inhibited shIN resistant HIV, no data was provided on the specific siRNAs generated from these lhRNAs nor

was there any analysis of the target sites that were susceptible to such inhibition. Taken together, these studies suggest that lhRNAs are recognised by Dicer and cleaved into functional siRNAs capable of suppressing viral replication. Importantly, the extended duplex region of lhRNAs discriminates these pre-miRNA mimics from shRNAs and theoretically provides Dicer with a substrate that could be processed into multiple siRNAs. This would provide a unique combinatorial RNAi strategy to be investigated against HIV. However, additional data describing Dicer processing of long dsRNAs is required and the choice of HIV-specific target remains largely unexplored. This latter point is of particular relevance to HIV where a diverse set of viral mRNAs are present during wild-type infection.

Previous RNAi-mediated suppression of HIV has largely focused on viral transcripts and while one study using synthetic siRNAs targeted the primer binding site immediately downstream of the 5' LTR (Han, Wind-Rotolo*, et al.*, 2004), little is known about the efficacies of expressed RNAi mimics specific to the transcribed region of the LTR. The nature of the 5' LTR promoter means it is present within viral genomic RNA as well as sub-genomic RNA species implying that knockdown of the LTR would reduce transcripts produced at both early and late stages of the viral life cycle. A previous study noted unique LTR enhancer and promoter configurations for each HIV subtype and moreover, while all LTR sequences responded equally to a subtype B transcriptional transactivator Tat protein, differences were noted in the basal levels of transcription that occurred in the absence of Tat (Jeeninga, Hoogenkamp*, et al.*, 2000). This suggested that the mechanism of Tat activation was similar for all LTR subtypes, and the changes in basal transcription were correlated to the number of NF-κβ sites within each viral promoter. A subtype C promoter contains three of these sites while a subtype B sequence has only two, and both subtypes each encode three Sp1 binding sites. The implications of these variations are unclear but given the significance of Tat in the viral life cycle, assessing RNAi-mediated inhibition in the presence and absence of Tat provides an interesting avenue for exploration. Furthermore, the sequence diversity present between subtype B and C viral isolates remains an important consideration for any combinatorial approach and provides a mechanism to explore whether or not a combinatorial construct such as a lhRNA would be effective at inhibiting diverse targets.

In this chapter, three distinct RNA Pol III-expressed lhRNAs targeted to the R and U5 transcribed regions of the 5' LTR of HIV were investigated for the potential to generate functional siRNAs along the duplex of the stem that were capable of inhibiting subtype B and C viral sequences. Both episomal and integrated forms of the LTR were chosen as targets and importantly, a wild type HIV isolate that included a pool of non-isogenic target sequences was included in the analysis.

2.2 Materials and methods

2.2.1 Cloning of expressed anti-HIV long hairpin RNAs targeted to the 5' LTR

As the anti-HIV lhRNAs used here were designed to be pre-miRNA mimics, they were expressed from an RNA Pol III promoter, specifically the mammalian U6 small nuclear RNA promoter (GenBank® accession number X59362). The generation of a Pol III U6 shRNA cassette has been previously described (Castanotto, Li*, et al.*, 2002; Lee, Dohjima*, et al.*, 2002) and a similar approach was used to produce 3 distinct lhRNAs targeted to a 190 bp contiguous region of the HIV 5' LTR [co-ordinates 454-512 (lhRNA LTR 1), 507-565 (HIV lhRNA 2) and 566-624 (HIV lhRNA 3) numbered according to the HXB2 molecular clone sequence, GenBank® accession number K03455 and Figure 8].

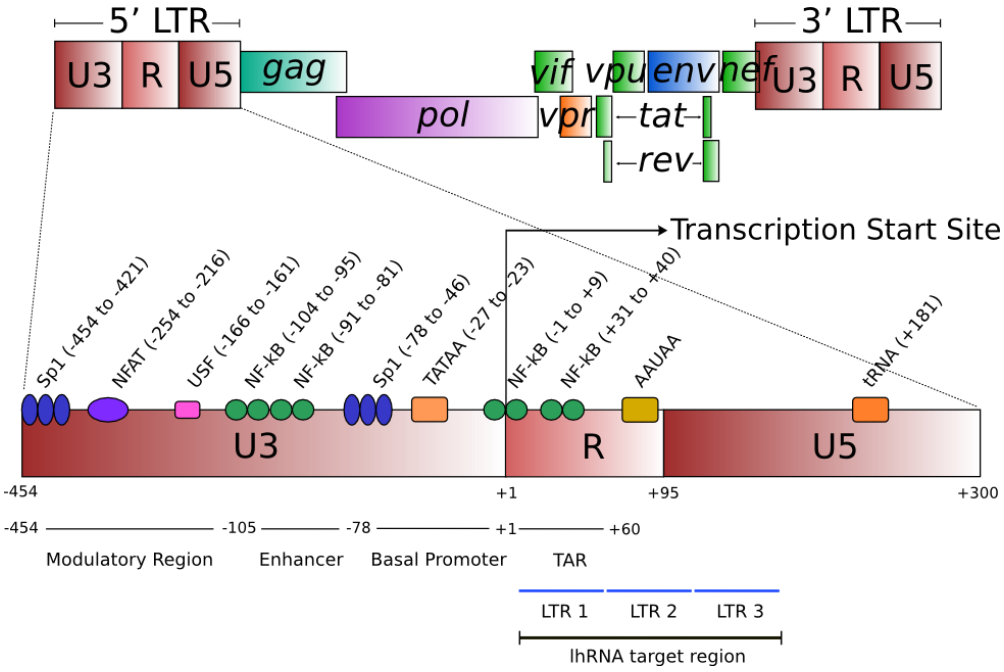


Figure 8. Long hairpin RNA target sites within the HIV 5' LTR. Diagrammatic representation of a subtype C HIV-1 proviral genome depicting all viral ORFs flanked by 5' and 3' LTR sequences (upper panel). A more detailed map of the 5' LTR is highlighted, with important *cis* and *trans* regulatory elements shown (see 2.1.). The conserved target regions for each of the 63 bp lhRNAs are shown within the R and U5 regions of the LTR (blue lines). The TAR loop corresponds completely with lhRNA LTR 1 (co-ordinates 454 to 512) while lhRNAs LTR 2 (co-ordinates 507 to 565) and LTR 3 (co-ordinates 566 to 624) are immediately downstream (numbers are according to the HXB2 HIV sequence GenBank® accession number K03455). Drawing is not to scale.

A two-step PCR approach was used to generate the anti-LTR lhRNAs (Figure 9) with oligonucleotide primers obtained from Integrated DNA Technologies, IA, USA.

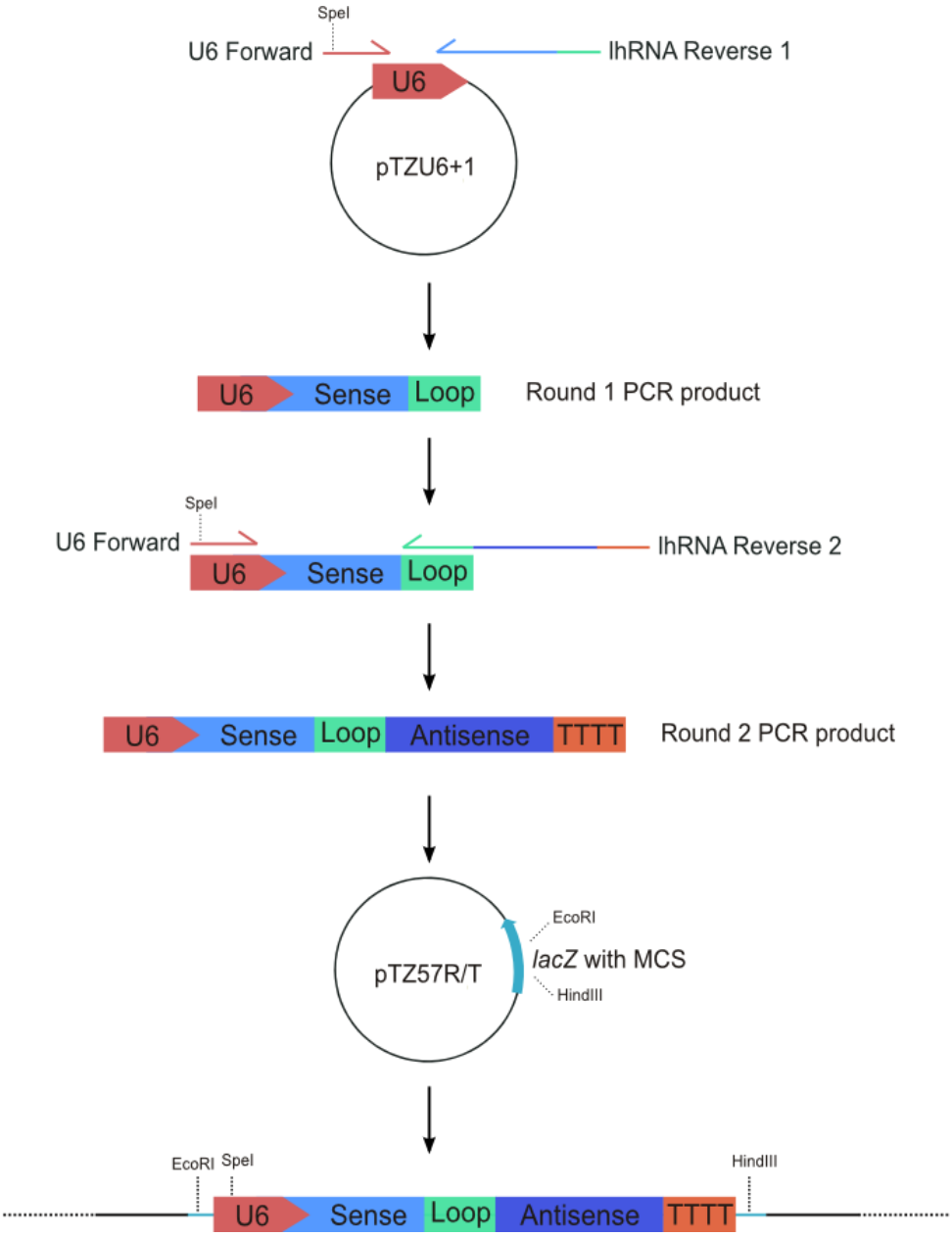


Figure 9. Two-step PCR cloning strategy used to generate anti-HIV lhRNAs. Round 1 PCR amplification used universal U6 forward and lhRNA LTR Reverse 1 primers with pTZU6+1 plasmid as template to generate a Round 1 PCR product that comprised the U6 promoter, sense strand and loop region of the lhRNA. This product was diluted and used as template in the second PCR reaction with universal U6 forward and lhRNA LTR Reverse 2 primers to generate a lhRNA LTR comprising the U6 promoter, sense, loop, antisense and polyT tail. The lhRNA LTR was ligated into the TA vector pTZ57R/T and candidate clones were screened using *Spe*I and *Hind*III digestion.

For lhRNA LTR 1, the first round of PCR amplification was carried out with a universal U6 forward primer containing a *Spe*I restriction site (5' CTA ACT AGT GGC GCG CCA AGG TCG GGC AGG AAG AGG G 3'; the *Spe*I site is underlined) and lhRNA LTR 1 Reverse 1 primer (5' CCT CTC TTG AAG AGT CCC CTA AAT AAC CAG AGA ACC CCC GGA CTC AGA TCC GGT

CCA CCC AGA AAG AAC CGG TGT TTC GTC CTT TCC ACA A 3') with pTZU6+1 plasmid DNA as template (Carmona, Ely*, et al.*, 2006). A PCR was carried out using Expand High Fidelity^{PLUS} PCR System reagents (Roche, Basel, Switzerland) on a Mastercycler® (BioRad, CA, USA). The PCR mixture included 0.4 µM each of forward and reverse primers, 10 ng of template, 1.5 mM $MgCl_2$, 200 µM dNTP mix, 2.5 U of Expand HiFi^{PLUS} Taq polymerase and 1 x final volume Expand HiFi^{PLUS} Reaction Buffer (proprietary information) in a final volume of 50 µL. PCR conditions comprised 3 minutes at 95 ^{o}C followed by 35 cycles of 30 seconds each at 95 ^{o}C, 58 ^{o}C and 72 ^{o}C followed by 10 minutes at 72 ^{o}C. The amplified product was excised and eluted from a 1 % agarose gel (see Appendices A1.2.3 and 1.2.4). Following extraction, the PCR product was diluted to 10 ng/µL in distilled water and 1 µL of the dilution was used as template for a second PCR amplification step using the universal U6 forward primer and lhRNA LTR 1 Reverse 2 primer containing an *Xba*I site (5' GAT CTC TAG AAA AAA GGG TCT CTC TAG GTA GAC CAG ATC TGA GCC CGG GAG CTC TCT GGC TAT CTA GGG AAC CCT CTC TTG AAG AGT CCC C 3'; the *Xba*I site is underlined). The PCR mixture and cycling conditions were the same as those described above. The amplified product was excised and eluted from a 1 % agarose gel. Following extraction the PCR product was diluted to 100 ng/µL in distilled water and ligated into the TA cloning vector pTZ57R/T (Fermentas, WI, USA). A ligation reaction containing a 3:1 molar ratio of insert to vector (0.09 pmol insert to 0.03 pmol vector) was incubated at room temperature for 1 hour in a 20 µL reaction volume containing 20 U T4 DNA Ligase (New England Biolabs® Inc, MA, USA). Aliquots (10 µL) of the ligation reaction were used to transform competent *Escherichia coli* DH5α (Invitrogen, CA, USA). Transformed cells were plated on Luria Bertani agar plates (see Appendix A1.1.1) containing Ampicillin (Amp, Gibco, BRL, UK), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; Sigma-Aldrich, MO, USA) and IPTG (Roche, Basel, Switzerland). To identify correctly cloned lhRNA LTR 1-encoding plasmids, individual white colonies (± 10) were cultured in 3 ml of Ampicillin positive medium (see Appendix A1.1.1). Plasmids were prepared using a commercially available kit (three methods are described in Appendix A1.2.1) and 10 µL was digested for 60 minutes at 37 ^{o}C with *Spe*I and *Hind*III in a 20 µL final reaction volume using 0.5 U of each enzyme, 33 mM Tris-acetate (pH 7.9 at 37 ^{o}C), 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/ml BSA (Tango Buffer, Fermentas, WI, USA). Digested plasmid DNA was electrophoresed on a 1 % agarose gel and compared to a molecular weight marker (O'GeneRuler^{TM} Mix, Fermentas, WI, USA). Correct candidate clones were sequenced to determine their fidelity (InqabaBiotech, Pretoria, South Africa). The same two step PCR cloning strategy was used to generate lhRNA LTR 2, lhRNA LTR 3 and a control lhRNA targeted to a site within the Hepatitis B Virus (HBV) genome (1581 to 1640 of subgenotype A1, GenBank® accession number J02203) with the

following changes: lhRNA LTR 2 was generated using the universal U6 forward primer, lhRNA LTR 2 Reverse 1 primer (5' TGC CCT CTC TTG AAA GAC ACA CAC CAC TCA GAA CAC CCA AGA CAA ACT TCA TTA AGG CCT AAA CAG TAG GCT CGG TGT TTC GTC CTT TCC ACA A 3') and lhRNA LTR 2 Reverse 2 primer containing a *Hind*III site (5' AAA AAA GAA CCC ACT GCT TAA GCC TCA ATA AAG CTT GCC TTG AGT GCT CTA AGT AGT GTG TGC CCT CTC TTG AAA GAC 3'; the *Hind*III site is underlined). LhRNA LTR 3 was generated using the universal U6 forward primer, lhRNA LTR 3 Reverse 1 primer (5' GGA AAT CTC TTG AAC TCC CAC ACC AAC ACA AAG AGT CTA AGA GAC CTC CAG TCA CCA GAA TCA CAC AAC AAA CGG TGT TTC GTC CTT TCC ACA A 3') and lhRNA LTR 3 Reverse 2 primer (5' AAA AAA GTC TGT TGT ATG ACT CTG GTA ACT AGA GAT CCC TCA GAC CCT TTG TGT TAG TGT GGA AAT CTC TTG AAC TCC 3'). LhRNA HBV was generated using the universal U6 forward primer, lhRNA HBV Reverse 1 primer (5' ACT CTC TTG AAG CGC AAA GGC GAA GCA AAG TAC ACA CGA TCC GAC AGA CGA GAA GAC ACA AAC AGG AAG TCG GTG TTT CGT CCT TTC CAC AA 3') and lhRNA HBV Reverse 2 primer containing an *Xba*I site (5' GAT CTC TAG AAA AAA GAC TCC CCG TCT GTG CCT TCT CAT CTG CCG GAC CGT GTG CAC TTC GCT TCA CCT CTG CAC TCT CTT GAA GCG CAA AG 3'; the *Xba*I site is underlined). Candidate pTZ57R/T-lhRNA LTR 2 and 3 clones were screened by digestion using *Spe*I and *Hind*III and sequencing as described above.

2.2.2 Cloning of a subtype C 5' LTR-driven luciferase reporter target vector

To assess target knockdown induced by the three distinct expressed anti-HIV lhRNAs described above, two reporter vectors were constructed that utilised either a subtype B or a subtype C 5' LTR promoter to drive expression of humanised Firefly luciferase (hFluc) activity. The subtype B luciferase vector was previously described (Jeeninga, Hoogenkamp*, et al.*, 2000) and contained an LTR promoter sequence from the HIV-1 LAI molecular clone (Peden, Emerman*, et al.*, 1991). To generate a subtype C luciferase reporter, a two-step cloning procedure was used. Firstly, the 5' LTR of the HIV-1 isolate Du151 (GenBank® accession number AY043173) was amplified by PCR from proviral samples obtained from M. Papathanasopolous using Du151 LTR Forward primer containing *Sal*I, *Eco*RV and *Bgl*III sites (5' GAT CGT CGA CGA TAT C*AG ATC T*GG AAG GGT TAA TTT ACT CTA AG 3'; the *Sal*I site is underlined, the *Eco*RV site is shown by a dotted underline and the *Bgl*III site is italicised) and Du151 LTR Reverse primer containing *Eco*RI and *Hind*III sites (5' ATC GAA TTC AAG CTT GTT CGG GCG CCA CTG CTA GAG ATT TTC CA 3'; the *Eco*RI site is underlined and the *Hind*III site is shown by a dotted underline). PCR was carried out using the same reaction mixture components as described for the generation of the anti-HIV lhRNAs (see section 2.2.1) however the cycling conditions differed as

follows: 3 minutes at 95 ^{o}C followed by 30 cycles of 45 seconds each at 95 ^{o}C, 60 ^{o}C and 72 ^{o}C followed by 10 minutes at 72 ^{o}C. Agarose gel-purified PCR products were ligated into the TA cloning vector pTZ57R/T to generate pTZ-Du151-LTR and candidate clones were screened by digestion using *Eco*RV and *Hind*III and sequencing. Secondly, the Luc-SV40polyA sequence derived from the Firefly luciferase cassette of plasmid pGL3 (Promega, WI, USA) was amplified by PCR using Luc-SV40pA Forward primer (5' ATT AGA ATT CAT GGA AGA CGC CAA AAA 3') and Luc-SV40pA Reverse primer containing *Apa*I, *Spe*I and *Hind*III sites (5' ATT AGG GCC CAC TAG T*AA GCT T*AC CAC ATT TGT AGA GGT TTT AC 3'; the *Apa*I site is underlined, the *Spe*I site is shown by a dotted underline and the *Hind*III site is italicised) as described for the 5' LTR PCR amplification. Agarose gel-purified PCR products were ligated into the TA cloning vector pCR®2.1-TOPO® (Invitrogen, CA, USA) to generate pCR®2.1-Luc-SV40pA and candidate clones were screened by digestion using *Eco*RI and *Hind*III and sequencing. Positive orientation pTZ-Du151-LTR and pCR®2.1-Luc-SV40pA clones were digested using *Bam*HI and *Pst*I in a 50 µL final volume reaction. Agarose gel-purified digestion products were ligated as described (see section 2.2.1). A 3:1 molar ratio of annealed fragment to vector (0.09 pmol annealed fragment to 0.03 pmol vector) was used to generate pLTRc-luc-SV40pA. This construct served as an episomal LTR-driven luciferase target reporter. To generate a LTR-driven luciferase target vector capable of integration, the LTRc-luc-SV40pA region of pLTRc-luc-SV40pA was amplified by PCR using Du151 LTR Forward primer and Luc-SV40pA Reverse primer and the agarose gel-purified PCR product was inserted in the TA vector psiLentGene-Puro (Invitrogen, CA, USA) using the same ligation molar ratio as described for pLTRc-luc-SV40pA to generate pLTRc-luc-SV40pA-Puro. Candidate clones were screened by digestion using *Eco*RI and *Hind*III and sequencing. The pLTRc-luc-SV40pA-Puro vector was used to generate a stable HEK293 cell line following Puromycin selection (see section 2.2.5).

2.2.3 Cloning of a tiling array set of target reporter vectors

The HIV tiling array consisted of a set of dual luciferase-based reporter vectors that comprised a single target vector complementary to the complete lhRNA LTR 2 target sequence as well as 5 overlapping target vectors (‘tiles’) that together spanned the entire lhRNA LTR 2 target sequence. The psiCheck2 dual luciferase vector (Promega, WI, USA) provided a convenient method of generating a luciferase based target reporter and comprised an HSV TK-driven hFluc cassette as well as an SV40-driven humanised *Renilla* luciferase (hRluc) cassette with a multiple cloning site within the 3' UTR. The cloning strategy used here introduced the HIV LTR target sequences within the 3' UTR of the hRluc cassette (Figure 10). Inhibition of the target sequence was detected as a change in the ratio of hRluc/hFluc. To generate the ‘LTR 2

complete target', the lhRNA LTR 2 target sequence was amplified from pLTRc-luc-SV40pA by PCR using LTR lhRNA 2 Forward primer containing an *Xho*I site (5' GAT CTC GAG GAA CCC ACT GCT TAA GCC TC 3'; the *Xho*I site is underlined) and LTR lhRNA 2 Reverse primer containing a *Not*I site (5' GAT CGC GGC CGC TTT CCA CAC TAA CAC AAA GG 3'; the *Not*I site is underlined) and the same PCR mixture and thermal cycler conditions as described above (see section 2.2.2). The purified PCR product was digested using *Xho*I and *Not*I and directionally ligated into the similarly restricted psiCheck2 vector (Promega, WI, USA) that had been dephosphorylated as follows: 3 µg of psiCheck2 was incubated for 60 minutes at 37 ^{o}C with 1 x Antarctic Phosphatase Reaction Buffer [50 mM bis-tris-propane-HCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ pH 6.0, New England Biolabs® Inc, MA, USA] and 5 U of Antarctic Phosphatase (New England Biolabs® Inc, MA, USA) in a final reaction volume of 10 µL followed by inactivation of the enzyme for 5 minutes at 65 ^{o}C. Dephosphorylated psiCheck2 and the digested PCR product were ligated using a 3:1 molar ratio as described for pLTRc-luc-SV40pA to generate psi-LTR 2-complete. To generate the first 'tile' in the HIV tiling array (Tile A), 10 µM each of complementary oligonucleotides Target A Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CGA ACC CAC TGC TTA AGC CTC AAG C 3') and Target A Reverse containing a *Not*I site (5' GGC CGC TTG AGG CTT AAG CAG TGG GTT CGA TAT C 3') were treated with T4 Polynucleotide Kinase (T4 PNK, Fermentas, WI, USA) for 30 minutes at 37 ^{o}C with the following conditions: 1 x Reaction Buffer A (50 mM Tris-HCl (pH 7.6 at 25 ^{o}C), 10 mM MgCl2, 5 mM DTT, 0.1 mM spermidine and 0.1 mM EDTA) in a final volume of 20 µL. The oligonucleotides were annealed after heating to 95 ^{o}C for 5 minutes followed by slow cooling to room temperature. Double stranded DNA, with 5′ *Xho*I and 3′ *Not*I cohesive ends, was ligated to the equivalent restriction sites of dephosphorylated psiCheck2 vector and candidate clones were screened by digestion using *Eco*RV and sequencing. The same cloning strategy was used to generate Tile B to E target vectors with the following oligonucleotide pairs: Target B Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCG TCT GTG CCT TCT CAT CTG CCG GAG C 3') and Target B Reverse containing an *Eco*RV site (5' GGC CGC TCC GGC AGA TGA GAA GGC ACA GAC GGA TAT C 3') for Tile B, Target C Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CTC ATC TGC CGG ACC GTG TGC AGC 3') and Target C Reverse containing an *Eco*RV site (5' GGC CGC TGC ACA CGG TCC GGC AGA TGA GAT ATC 3') for Tile C, Target D Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCC GTG TGC ACT TCG CTT CAC CTC TGG C 3') and Target D Reverse containing an *Eco*RV site (5' GGC CGC CAG AGG TGA AGC GAA GTG CAC ACG GGA TAT C 3') for Tile D and Target E Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCT TCG CTT CAC CTC TGC ACG TGC 3') and Target E Reverse containing an *Eco*RV site (5' GGC CGC ACG TGC AGA GGT GAA GCG AAG GAT ATC 3') for Tile E.

A similar strategy was used to generate a set of dual luciferase-based reporter vectors that comprised a single target vector complementary to a HBV-specific target sequence as well as 5 overlapping target tile vectors that together spanned the entire HBV target sequence (Appendix A3.1 contains additional information related to this assay). To generate the 'HBV complete target', the HBV target sequence was amplified from a plasmid (pCH-9/3091) that contains a greater than genome length HBV sequence similar to the HBV A1 subgenotype (Nassal, 1992) using lhRNA HBV Forward primer containing an *Xho*I site (5' GAT CTC GAG GAC TCC CCG TCT GTG CCT TCT 3'; the *Xho*I site is underlined) and lhRNA HBV Reverse primer containing a *Not*I site (5' GAT CGC GGC CGC ACG TGC AGA GGT GAA GCG AAG TGC ACA CGG 3'; the *Not*I site is underlined) and cloned into psiCheck2 as described above for the 'LTR 2 complete target' vector. The same cloning strategy as described above for the HIV tiling array was used to generate HBV-specific Tile A to E target vectors with the following oligonucleotide pairs: Target A Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CGA CTC CCC GTC TGT GCC TTC TGC 3') and Target A Reverse containing an *Eco*RV site (5' GGC CGC AGA AGG CAC AGA CGG GGA GTC GAT ATC 3') for Tile A, Target B Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCG TCT GTG CCT TCT CAT CTG CCG GAG C 3') and Target B Reverse containing an *Eco*RV site (5' GGC CGC TCC GGC AGA TGA GAA GGC ACA GAC GGA TAT C 3';) for Tile B, Target C Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CTC ATC TGC CGG ACC GTG TGC AGC 3') and Target C Reverse containing an *Eco*RV site (5' GGC CGC TGC ACA CGG TCC GGC AGA TGA GAT ATC 3') for Tile C, Target D Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCC GTG TGC ACT TCG CTT CAC CTC TGG C 3') and Target D Reverse containing an *Eco*RV site (5' GGC CGC CAG AGG TGA AGC GAA GTG CAC ACG GGA TAT C 3') for Tile D and Target E Forward containing *Xho*I and *Eco*RV sites (5' TCG AGA TAT CCT TCG CTT CAC CTC TGC ACG TGC 3') and Target E Reverse containing an *Eco*RV site (5' GGC CGC ACG TGC AGA GGT GAA GCG AAG GAT ATC 3') for Tile E. For all forward primers used to generate 'tile targets' in the HIV and HBV tiling arrays, the *Xho*I sites are underlined and the *Eco*RV sites are shown by a dotted underline. For all corresponding reverse primers, the *Eco*RV sites are underlined.

2.2.4 Mammalian cell culture

The human embryonic kidney (HEK293) cell line was maintained using Dulbecco's Modified Eagle's Medium (DMEM, Lorenza Biosciences, Basel, Switzerland) supplemented with heat-inactivated 10 % fetal calf serum (FCS, Invitrogen, CA, USA) at 37 ^{o}C and 5 % CO_2 in a Forma Series II 3110 Water Jacketed CO_2 incubator (Thermo Fisher Scientific, Inc, MA, USA). The

human astrocyte glioblastoma U87.CD4.CCR5 cell line was maintained and transfected using DMEM supplemented with heat-inactivated 15 % FCS, 50 U/mL Penicillin/50 µg/mL Streptomycin antibiotic mix (Gibco, BRL, UK), 1 µg/mL Puromycin antibiotic (Merck, Darmstadt, Germany), 300 µg/mL G418 antibiotic (Sigma-Aldrich, MO, USA) and 1 % L-glutamine (Sigma-Aldrich, MO, USA) at 37 ^{o}C and 5 % CO_2.

2.2.5 Transfection to generate a stable cell line expressing hFluc from a subtype C HIV 5' LTR

To generate a stable cell line that constitutively expressed hFluc from a HIV-1 subtype C 5' LTR promoter, HEK293 cells were transfected with pLTRc-luc-SV40pA-Puro and placed under Puromycin selection. Prior to the generation of the stable cell line, a minimum inhibitory concentration of Puromycin was determined by kill curve analysis using untransfected HEK293 cells. Twenty four hours prior to the addition of Puromycin, HEK293 cells were seeded at 120 000 cells per well in a 24 well dish (Nunc^{TM}, Thermo Fisher Scientific, Inc, MA, USA). Cells were counted on a haemocytometer using a 1:1 mixture of Trypan Blue stain (Sigma-Aldrich, MO, USA) and cells that had been washed in phosphate buffered saline (PBS, Gibco, BRL, UK) and resuspended in DMEM. Seeded cells were left to settle overnight and the following day duplicate wells were supplemented with 0.25 µg/mL, 0.50 µg/mL, 0.75 µg/mL, 1 µg/mL, 1.25 µg/mL, 1.50 µg/mL or 1.75 µg/mL Puromycin. Cell viability was monitored for 1 week using light microscopy and the minimum Puromycin concentration at which untransfected HEK293 cells were viable was 0.50 µg/mL. A concentration of 1 µg/mL Puromycin was selected to generate stable cell lines.

To generate a stable cell line using pLTRc-luc-SV40pA-Puro, 24 hours prior to transfection HEK293 cells were counted as described above and seeded at 2.5 million cells per 10 cm dish (Nunc^{TM}, Thermo Fisher Scientific, Inc, MA, USA). The following day, 10 µg of pLTRc-luc-SV40pA-Puro was mixed with 500 µL of OptiMem (Gibco, BRL, UK) and left to incubate for 10 minutes at room temperature. Separately, 10 µL of Lipofectamine 2000 (Invitrogen, CA, USA) was mixed with 500 µL OptiMem and incubated concurrently. The DNA/OptiMem mixture was added to the Lipofectamine 2000/OptiMem mixture, vortexed briefly and incubated for 20 minutes at room temperature before being added to the seeded HEK293 cells. Media was changed 24 hours post-transfection and supplemented with 1 µg/mL Puromycin for selection of stable clones. Single colonies of the LTRc-hFluc HEK293 stable cell line were picked approximately four weeks following transfection using Perspex rings and maintained in DMEM supplemented with Puromycin until a luciferase assay had been performed (see the following paragraph). Thereafter stable cells were maintained in the presence of 1 µg/mL Puromycin.

LTRc-driven hFluc activity was measured using the Promega Dual Luciferase Reporter Assay® Kit (Promega, WI, USA). Following selection and expansion of LTRc-hFluc HEK293 stable clones, cells were seeded at 120 000 cells per well as described above and left to settle overnight. The following day, cells were washed in PBS and incubated for 20 minutes at room temperature in 100 µL of 1 x Passive Lysis Buffer (proprietary information) per well. Fifty µL of 1 x Luciferase Assay II Reagent (LARII - proprietary information) was prepared per sample by transferring the contents of one bottle of LAR II into one vial of Luciferase Assay Substrate. Following lysis, 50 µL of each sample was aliquoted per well into a solid white luminometer plate (Turner Biosystems, CA, USA) before being loaded into a Veritas™ Dual-Injection Microplate Luminometer (Turner Biosystems, CA, USA) along with the pre-prepared LARII solution. Following priming of a single injector within the luminometer, samples were automatically exposed to LARII and hFluc expression was analysed using Veritas™ software.

2.2.6 Transfection and detection of luciferase activity in cultured mammalian cells

To assess the inhibitory effects of the three anti-LTR lhRNAs against the episomal subtype B and C LTR-driven luciferase reporter targets, HEK293 cells were seeded at 120 000 cells per well in 24 well dishes and transfected in triplicate as described (see section 2.2.5) with the following changes: cells were co-transfected with a combination of 90 ng of pLTRb-luc (LAI) or pLTRc-luc-SV40pA and 900 ng of either pTZ-lhRNA LTR 1, 2, 3 or pTZ-lhRNA HBV as well as 90 ng of either pCMV-Tat (a gift from J. van Harmelen) or pCI-Neo mock (Promega, WI, USA). In addition, 5 ng of an RNA Pol II-expressed (CMV) *Renilla* luciferase plasmid (pRSV-Rluc – a gift from J. J. Rossi) was included per sample, as well as 10 ng of a plasmid vector that constitutively produced enhanced green fluorescent protein [pCI-eGFP, (Passman, Weinberg*, et al.*, 2000)] to monitor transfection efficiency using fluorescence microscopy. The total amount of co-transfected DNA per well (1.10 µg) was incubated in 100 µL of OptiMem as described (see section 2.2.5). Separately, 1 µL of Lipofectamine 2000 was incubated in 100 µL of OptiMem and the transfection was completed as described (see section 2.2.5). Both hFluc (generated from pLTRb-luc or pLTRc-luc-SV40pA) and hRluc (generated from pRSV-Rluc) were measured using a luciferase assay as described above with the following changes: 48 hours post-transfection cells were washed in PBS, incubated with 1 x Lysis Buffer and 50 µL of each sample was aliquoted per well into a luminometer plate. Fifty µL each of 1 x LARII and 1 x Stop & Glo® solution (proprietary information) were prepared per sample. Both injectors within the luminometer were primed and samples were automatically exposed to LARII followed by Stop & Glo®. Notably hFluc activity was assayed using the LARII solution and hRluc was assayed using the Stop and Glo® solution.

To investigate the inhibitory effects of the three anti-LTR lhRNAs against the integrated subtype C LTR-driven luciferase reporter target, LTRc-hFluc HEK293 stable cells were seeded and co-transfected in triplicate as described above with 90 ng of either pCMV-Tat or pCI-Neo mock and 900 ng of either pTZ-lhRNA-LTR 1, 2 or 3 or pTZ-lhRNA HBV or pTZU6+1 mock (Carmona, Ely*, et al.*, 2006). The expression of hFluc was assessed by luciferase activity as described (see section 2.2.5).

To evaluate the inhibitory effects of the three anti-LTR lhRNAs against the HIV Tiling Array target reporter vectors, HEK293 cells were seeded and co-transfected in triplicate as described above with 90 ng of either psi-LTR 2 complete or psi-Tile A, B, C, D or E and 900 ng of pTZ-lhRNA LTR 2 or pTZ-lhRNA HBV or pTZU6+1 mock. To assess the inhibitory effects of lhRNA HBV against the HBV Tiling Array target reporter vectors, HEK293 cells were seeded and co-transfected in triplicate as described above with 90 ng of either psi-HBV complete or psi-Tile A, B, C, D or E and 900 ng of pTZ-lhRNA HBV or pTZU6+1 mock. Dual luciferase activity was assessed as described above.

To investigate the inhibitory effects of the three anti-LTR lhRNAs against a subtype B infectious molecular clone of HIV, HEK293 cells were seeded and co-transfected in triplicate as described above with 40 ng of pNL4-3.Luc.R-E- (NIH AIDS Research and Reference Reagents Program, MD, USA), 20 ng of pRSV-Rluc, 80 ng of pBluescript SK (as a carrier, Invitrogen, CA, USA) and 40 ng of either pTZ-lhRNA LTR 1, 2 or 3 of pTZ-lhRNA HBV or pTZ-shRev (a gift from J. J. Rossi). The pNL4-3.Luc.R-E- clone contained a hFluc gene within the *nef* ORF and two frameshift events had rendered the clone Vpr and Env deficient (Connor, Chen*, et al.*, 1995; He, Choe*, et al.*, 1995). The clone was capable of a single round of replication and the hFluc gene served as a reporter. Dual luciferase activity was assessed as described above.

2.2.7 Northern blot analysis to detect processed guide sequences from anti-LTR lhRNAs

TriReagent™ (Sigma-Aldrich, MO, USA) was used to extract total RNA (see Appendix A1.3.1) from HEK293 cells that had been seeded in 10 cm dishes and transfected with 20 µg of either pTZ-lhRNA HBV, pTZ-shHBV or pTZU6+1 as described (see section 2.2.5). A radioactively labeled RNA molecular weight marker (see following paragraph, Ambion Decade Marker™, Applied Biosystems, CA, USA) as well as 30 µg of RNA per sample were mixed with an equal volume of RNA loading dye (Appendix A1.3.3) and resolved on a 15 % polyacrylamide (Sigma-Aldrich, MO, USA) denaturing (7 M urea, Sigma-Aldrich, MO, USA) gel using TBE buffer (890 mM Tris base, 890 mM Boric acid and 32 mM EDTA, Merck, Darmstadt, Germany) and stained with 10 mg/mL ethidium bromide (Sigma-Aldrich, MO, USA) to ensure equivalent quantities of RNA were loaded in each lane, prior to being transferred to a nitrocellulose

membrane (Hybond, Amersham Biosciences, CT, USA) using a semi-dry blotter (Hybaid, Amersham Biosciences, CT, USA) for 60 minutes at 100 volts. RNA was cross-linked to the blots with 1000J UV light (Stratalinker, Stratagene, CA, USA) followed by baking for 60 minutes at 80 ^{o}C (Thermo Fisher Scientific, Inc, MA, USA).

DNA oligonucleotide probes A (5' GAC TCC CCG TCT GTG CCT TCT CA 3'), B (5' TCA TCT GCC GGA CCG TGT GCA CT 3'), C (5' TGC ACT TCG CTT CAC CTC TGC ACT C 3') were hybridized to baked blots to detect guide sequence production from lhRNA HBV and shHBV while a U6 snRNA probe (5' TAG TAT ATG TGC TGC CGA AGC GAG CA 3') was used to detect small nuclear RNA as a marker of equal RNA loading between samples. Probes were labeled at the 5' end with γ-32P-ATP (6000 ci/mmol, Perkin Elmer, MA, USA) and T4 PNK for 30 minutes at 37 ^{o}C with the following components: 1 x Reaction Buffer A (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 0.1 mM spermidine and 0.1 mM EDTA), 20 pmol of oligonucleotide probe, 20 pmol of γ p-32 ATP and 10 U of T4 PNK in a final volume of 20 µL. The addition of 1 µl of 0.5 M EDTA pH 8.0 followed by 10 minutes of incubation at 75 ^{o}C terminated the reaction. Probes were purified using Sephadex® G-25 columns (Sigma-Aldrich, MO, USA) and centrifugation for 5 minutes at 59.5 x g before being added to the blot that had been incubated for 30 minutes with 10 mL pre-warmed hybridization solution (Rapid Hyb, Amersham Biosciences, CT, USA) per 100 cm^2 of blot in a rotating hybridisation oven (Thermo Fisher Scientific, Inc, MA, USA). Hybridisation occurred overnight at 42 ^{o}C and the following day, blots were washed twice using 1 % SSC (NaCl/Na-citrate) and 0.1 % SDS (sodium dodecylsulphate, Sigma-Aldrich, MO, USA) in a total volume of 50 mL per wash before being exposed for 2 days to x-ray film at -80 ^{o}C. Probed blots were stripped using a high stringency wash solution (0.1 % SSC/1 % SDS) for 30 minutes at 50 ^{o}C.

2.2.8 Quantitative real-time reverse transcription PCR (qRT-PCR) of interferon-related genes in cultured mammalian cells transfected with anti-LTR lhRNAs

To analyse immunostimulation in HEK293 cells following transfection with the anti-LTR lhRNAs, interferon-β (IFN-β) mRNA levels were assayed by qRT-PCR. Total RNA was extracted from HEK293 cells that had been transfected with 100 ng pCi-eGFP and 20 µg of either pTZ-lhRNA LTR 1, 2 or 3 or pTZU6+1 mock or 2 µg of polyinosinic: polycytidylic acid [poly (I:C), Sigma-Aldrich, MO, USA]. Poly (I:C) is a dsRNA construct that is known to induce an IFN response in HEK293 cells (Loseke, Grage-Griebenow*, et al.*, 2006; Reynolds, Anderson*, et al.*, 2006). RNA pellets were resuspended at a concentration of 2 µg/µL in RNAse-free water and DNAse treated for 30 minutes at 37 ^{o}C using 10 mM Tris-HCl pH 7.5, 2.5 mM $MgCl_2$, 0.1 mM $CaCl_2$ and 1 U of DNAse I (Fermentas, WI, USA). Two µg of DNAse-treated RNA was reverse-transcribed for 60

minutes at 37 ^{o}C using Omniscript RT (Qiagen, Hilden, Germany) with the following reaction components: 0.5 mM dNTP mix, 0.1 µM Oligo-dT primer (12 nt), 10 U RNAse inhibitor and 4 U of Omniscript RT in a total volume of 20 µL. Complementary DNA (cDNA) was amplified by PCR with the Quantitect® SYBR® Green PCR Kit (Qiagen, Hilden, Germany) on a Roche Lightcycler V.2. (Roche, Basel, Switzerland). To allow for absolute quantification of RNA, two primer sets were used, one for IFN-β (Forward: 5' TCC AAA TTG CTC TCC TGT TGT GCT 3' and Reverse: 5' CCA CAG GAG CTT CTG ACA CTG AAA A 3') and one for the housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (GAPDH Forward: 5' AGG GGT CAT TGA TGG CAA CAA TAT CCA 3' and Reverse: 5' TTT ACC AGA GTT AAA AGC AGC CCT GGT G 3'). PCR reaction mixtures included 0.5 µM of each forward and reverse primer set, 1 µg of cDNA, and 1 x Quantitect® SYBR® Green PCR Master Mix (HotStar Taq DNA polymerase, Tris-Cl, KCl, $(NH_4)_2SO_4$, dNTP mix, SYBR Green and $MgCl_2$, concentrations not provided) in a final volume of 25 µL. PCR cycling conditions included: hot start for 30 seconds at 95 ^{o}C followed by 50 cycles of 95 ^{o}C for 10 seconds, 58 ^{o}C for 10 seconds and 72 ^{o}C for 10 seconds. PCR product specificity was determined by melting curve analysis. Controls included water blanks (no template added), and RNA template that had not been subjected to reverse transcription.

2.2.9 Viral propagation and challenge assays using p24 ELISA

To investigate the inhibitory efficacies of the three anti-LTR lhRNAs against a wild type strain of HIV-1, a viral challenge assay was conducted in U87.CD4.CCR5 cells. A South African R5-tropic subtype C HIV isolate that had been isolated from an HIV positive AIDS patient admitted to the Johannesburg General Hospital in 1998 was used. The strain was designated FV5 (GenBank® accession number DQ382363) and propagated by standard peripheral blood mononuclear cell (PBMC) co-culture techniques (M. Papathanasopolous). The co-receptor tropism of FV5 was established genotypically by automated sequencing of the V3 loop of the viral *env* gene (GenBank® accession number 05ZAFV5) and confirmed phenotypically by MT-2 assay (personal communication with M. Papathanasopolous).

For the challenge assay U87.CD4.CCR5 cells were seeded at 20 000 cells per well in a 96 well dish and transfected in triplicate as described (section 2.2.5) using 0.25 µg of either pTZ-lhRNA LTR 1, 2 or 3 or pTZ-lhRNA HBV or pTZ-U6+1 mock or pTZ-shRev in 25 µL of OptiMem and 0.25 µL of Lipofectamine 2000 per well in 25 µL of OptiMem. 24 hours later the cells were infected with FV5 using a $TCID_{50}$ of 1000/mL (tissue culture infectious dose as measured by p24 ELISA, see Appendix A2.1) and cultured for a further 24 hours. Cells were washed three times using PBS and fresh media was added to a final volume of 250 µL per well. On days zero (day

of washing), four and seven a maximum of 200 µL of supernatant was collected and stored at -80 ^{o}C for p24 ELISA analysis and 200 µL of fresh media was replaced.

Stored supernatants collected from the various days of a challenge assay were simultaneously thawed at room temperature and analysed by ELISA for p24 antigen production (Murex Biotech LTD, Kent, UK). One hundred µL of Conjugate Working Solution I (biotinylated murine anti-p24 monoclonal antibodies, Proclin® preservative, bovine albumin, heat inactivated mouse serum, aggregated human IgG in phosphate buffer, concentrations not supplied) was aliquoted per well into pre-coated ELISA strips. At least one positive and two negative samples (supplied in the kit) were included in each assay. One hundred µL per sample of thawed viral supernatant was aliquoted per well and aspirated to mix followed by incubation for 60 minutes at 37 ^{o}C. ELISA strips were washed five times with Wash Buffer (phosphate buffer and Proclin® preservative, concentrations not supplied) using an automated plate washer (BioRad, CA, USA). This was followed by the addition of 200 µL per well of Conjugate Working Solution II (peroxidase conjugated streptavidin, phosphate buffer, Proclin® preservative, bovine casein and bovine aprotinin, concentrations not supplied) and incubation for 30 minutes at 37 ^{o}C. A second wash step was completed followed by the addition of 200 µL of Substrate Solution (tetramethylbenzidine, dimethyl sulphoxide, Thiomersal preservative, phosphate citrate buffer and hydrogen peroxide, concentrations not provided) per well. Following 30 minutes of incubation at room temperature, 50 µL of 2M H_2SO_4 (Merck, Darmstadt, Germany) was added per well. Fluorescence was analysed at 450 and 690 nm using an automated plate reader (BioRad, CA, USA). A standard curve of serially diluted positive sample (supplied in the kit) was assessed in parallel to the experimental samples and values at 450/690 nm were converted to ng/mL using the standard curve.

2.2.10 Statistical analysis

Data are expressed as the mean ± the standard error of the mean (SEM) and where appropriate are normalised to a control sample within the experiment. Statistical difference was considered to be significant when $p < 0.05$ and highly significant when $p < 0.01$. In all cases $\alpha = 0.99$ and data was analysed using a one-way ANOVA followed by a Dunnett's Multiple Comparison post-test calculated using GraphPad Prism (GraphPad Software Inc, CA, USA).

2.3 Results

2.3.1 Inhibitory efficacies of anti-HIV lhRNAs targeted to subtype B and C LTR-driven reporters

Three discrete U6-driven lhRNAs (LTR 1 to 3) were expressed as sense-loop-antisense constructs with 63 bp stem regions and two bp 3' overhangs required for Dicer recognition. G:U mismatches were included in the sense strand approximately every four to eight bp (Figure 10) as this had been shown to facilitate the cloning steps required to generate such long constructs (Akashi, Miyagishi*, et al.*, 2005). The limitation of these mismatches to the sense strand ensured that the antisense (guide) strand was still 100 % complementary to its cognate target. Each lhRNA stem region was 63 bp in length as it was hypothesised this may provide Dicer with a substrate that could be processed to form three siRNAs of 21 bp each if Dicer cleaved from the two bp overhang towards the loop of the pre-miRNA mimic.

The HIV LTR provided a novel target for RNAi-mediated inhibition and two distinct LTR-based reporter vectors were used to assess the efficacies of the lhRNAs against a subtype B and C target sequence (Figure 11). As both vectors used the LTR promoter to drive expression of hFluc, changes in the hFluc activity in the presence of the hairpins was indicative of lhRNA-mediated LTR suppression. The addition of a separate plasmid driving hRluc expression controlled for transfection efficiencies between samples and provided a background luciferase reading against which changes in hFluc could be normalised. To explore the possible implications of basal vs. Tat-transactivated levels of LTR transcription, the inhibitory efficacies of the lhRNAs were determined in the presence and absence of exogenously supplied Tat. Moreover, as the wild type HIV-1 provirus only generated processive transcripts following the activation of TAR and RNA Pol II by Tat (Ramanathan, Rajpara*, et al.*, 2001; Richter, Cao*, et al.*, 2002), the presence and absence of exogenous Tat within cell culture assays served to mimic active and inactive forms of the LTR target respectively.

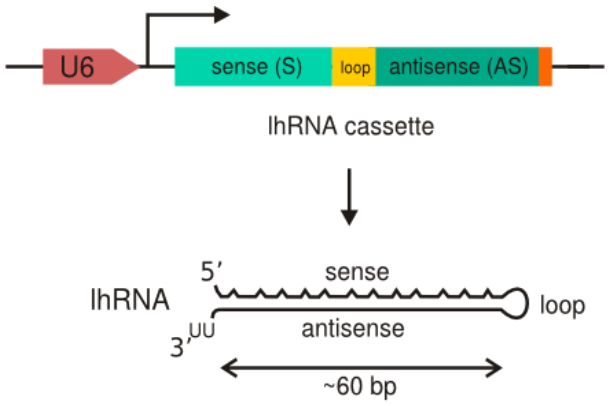


Figure 10. Design of expressed anti-HIV lhRNA cassettes. Pol III (U6) cassettes were designed to express three individual anti-HIV lhRNAs that each folded into a sense-loop-antisense structure with a 63 bp stem and characteristic 2 bp 3' overhangs. G:U wobble bases (arrows) were included within the sense strand of each lhRNA duplex to facilitate cloning. No mismatches were incorporated into the antisense strand thereby ensuring the guide sequence remained complementary to its cognate target.

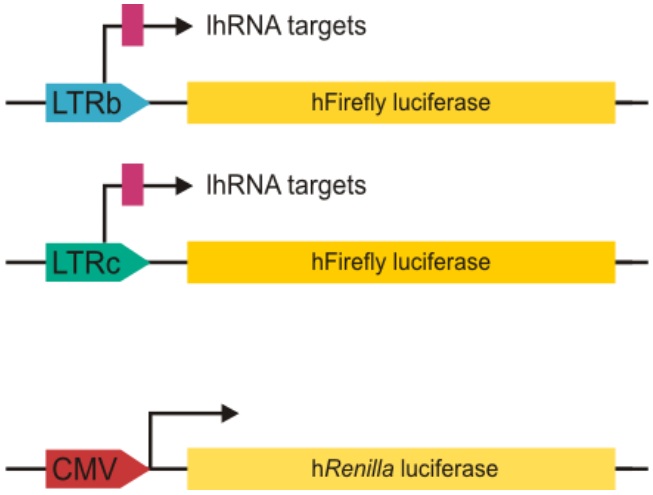


Figure 11. Design of LTR subtype B and C luciferase-based targets. Two separate hFluc reporter constructs were generated by cloning the subtype B (LAI) or C (Du151) LTR promoter upstream of a hFluc ORF. A CMV cassette expressing hRluc was co-transfected as a background luciferase control.

When targeted to a subtype B LTR-driven reporter vector (LAI), all three lhRNAs showed highly significant inhibition of the target ($\alpha = 0.99$; $p < 0.01$) in the absence and presence of exogenous Tat (Figure 12). As previous publications had revealed the TAR loop to be refractory to siRNA-mediated inhibition (Jacque, Triques*, et al.*, 2002; Yoshinari, Miyagishi*, et al.*, 2004), it was surprising to observe knockdown by lhRNA LTR 1 both in the absence (23.57 % ± 3.21 %) and presence (33.57 % ± 2.15 %) of Tat. The remaining two lhRNAs exhibited even greater target knockdown with lhRNA LTR 2 showing 52.30 % ± 4.81 % (no Tat) and 65.23 % ± 5.03 % (with

Tat) inhibition, while lhRNA LTR 3 achieved 30.17 % ± 3.11 % (no Tat) and 40.30 % ± 5.69 % (with Tat) target suppression. A slightly different set of data emerged using the subtype C LTR-driven reporter vector (Du151). In the absence of Tat, all three lhRNAs showed highly significant knockdown of the target vector with lhRNA LTR 1 and LTR 2 mediating 19.77 % ± 7.18 % and 39.43 % ± 8.80 % respectively ($\alpha = 0.99$; $p < 0.01$) while lhRNA LTR 3 achieved the greatest target suppression with 78.90 % ± 1.57 %. In the presence of exogenous Tat, lhRNA LTR 1 did not achieve any significant inhibition ($\alpha = 0.99$; $p > 0.05$) with 18.20 % ± 1.50 % knockdown. Both lhRNAs LTR 2 and 3 showed highly significant suppression ($\alpha = 0.99$; $p < 0.01$) with 49.40 % ± 6.62 % and 81.57 % ± 1.82 % inhibition respectively.

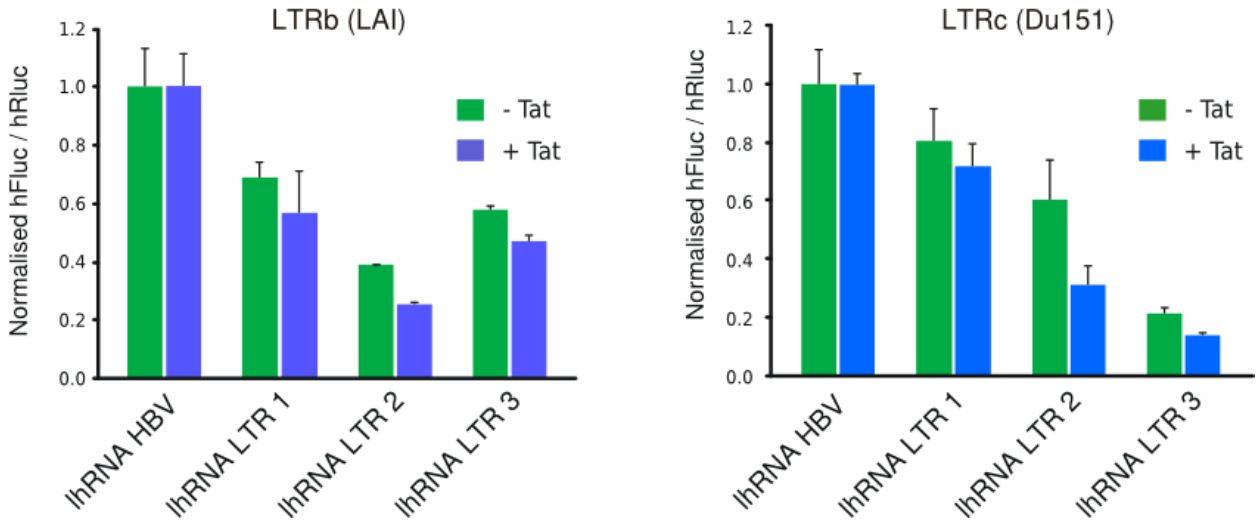


Figure 12. Inhibitory efficacies of three anti-HIV lhRNAs targeted to LTRb and LTRc luciferase reporter vectors. LTR1, 2 and 3 were co-transfected with target reporter plasmids in the presence (blue bars) and absence (green bars) of exogenous Tat in HEK293 cells. Data are expressed as a ratio of hFluc to hRluc normalised to a non-specific lhRNA control (lhRNA HBV) determined 48 hours post-transfection (n = 3, ± SEM). The left hand panel depicts lhRNA-mediated inhibition of a subtype B reporter vector (LTRb LAI) and the right hand panel shows inhibition of a subtype C based reporter (LTRc Du151).

Taken together, these data suggested that at least two of the three anti-LTR lhRNAs were capable of mediating significant inhibition of both subtype B and C target reporter vectors regardless of the presence of exogenous Tat. Although lhRNA LTR 1 mediated some inhibition of the subtype B target, the degree of suppression (33.57 %) was demonstrated against a reporter vector and may not be sufficient in the context of a wild type HIV infection. The inhibition mediated by lhRNAs LTR 2 and 3 in the presence of exogenous Tat suggested that these pri-miRNA mimics were efficacious despite numerous copies of target RNA. This would provide an important advantage in the context of wild type HIV infection where high viral loads are present.

Of greater interest was the inhibition mediated by lhRNAs LTR 2 and 3 in the absence of exogenous Tat as this suggested that non-processive transcripts were suppressed with implications for the completion of the viral life cycle (discussed further in section 2.4). The observation that lhRNAs LTR 2 and 3 mediated inhibition of both LTRb and LTRc reporter vectors suggested that minor sequence variations may be tolerated by lhRNAs. An additional indication that sequence variation may have influenced the inhibitory data was the observation that lhRNA LTR 2 mediated greater knockdown of a subtype B target while lhRNA LTR 3 showed greater inhibition of a subtype C based vector, regardless of the presence of Tat.

2.3.2 Alignment of LTRb and LTRc promoter sequences

The observed variations between lhRNA-mediated inhibition of the LTRb and LTRc-driven luciferase reporters may have been as a result of sequence differences between the two promoter regions. RNAi is well-known to be negatively influenced by mismatches between the guide strand of a siRNA and its target mRNA, particularly within the seed region (Yekta, Shih*, et al.*, 2004; Brennecke, Stark*, et al.*, 2005). Alignment of the LTRb and LTRc target sequences revealed minor base changes that may have been responsible for the differences in lhRNA efficacies (Figure 13). When aligned with lhRNA LTR 1, four mismatches that spanned the length of the stem duplex were detected within the subtype B (LAI) target sequence. Specifically these were A to G at position 464, G to T at position 466, C to T at position 484 and T to A at position 501. However, the significance of these specific mismatches was difficult to interpret because lhRNA LTR 1 was targeted to the TAR loop that was known to be refractory to RNAi-mediated inhibition. Alignment of LTRb and LTRc regions targeted by lhRNA LTR 2 revealed two mismatches within the LTRb sequence, both located close to the loop side of the stem duplex. These included a C to T followed by a T to C at positions 548 and 549. The mismatches could perhaps have affected siRNAs that were produced from the loop side of the stem while the remaining sequence of both target subtypes was a perfect match to lhRNA LTR 2. Three subtype B mismatches were revealed within the lhRNA LTR 3 target alignments one of which was located within the initial 23 bp of the duplex that may have corresponded with the first Dicer cleavage product if the lhRNA was processed from the 3' overhang end. This specific base change was A to G at position 575, followed by a G to T at position 610 and a T to C at position 614. Thus while the degree of inhibition varied between the two LTR sequences, knockdown still occurred with two of the three lhRNAs and this suggested that minor sequence variations as detected by sequence alignments may be tolerated by lhRNAs. This would be favourable if these specific types of RNAi mimics were to be used as a combinatorial anti-HIV strategy.

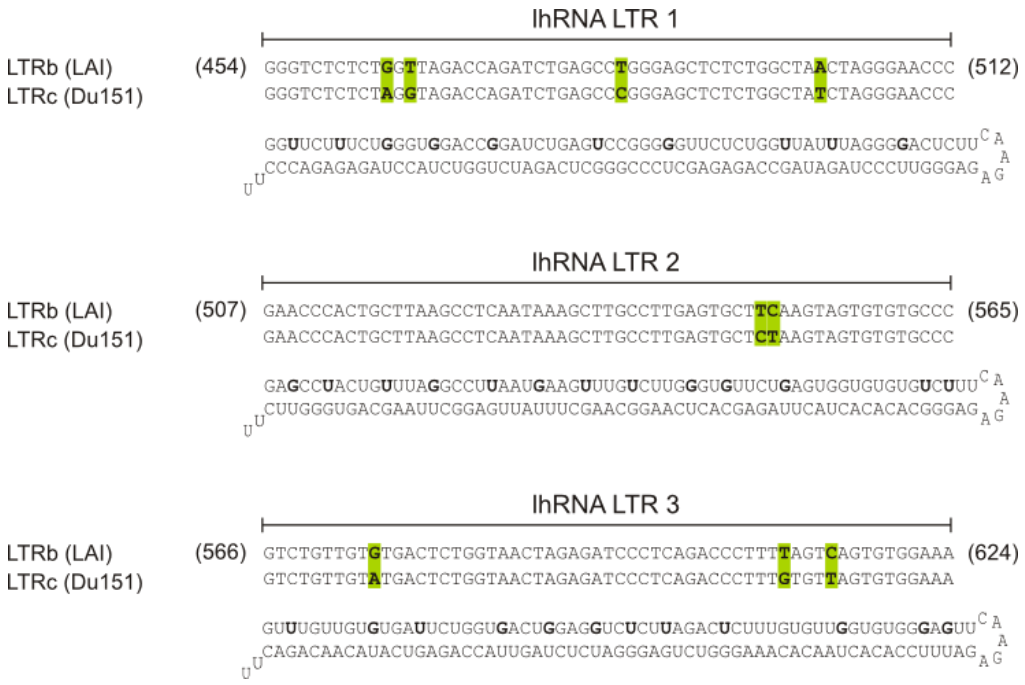


Figure 13. Alignment and comparison of LTRb (LAI) and LTRc (Du151) target sequences. All three lhRNAs targeted the LTRc sequence (bottom line of sequence in each panel). Mismatches between the LTRb and LTRc sequences are highlighted in bold with green boxes. Long hairpins are shown in with respect to their cognate target sequence, and G:U mismatches are shown in bold. Numbering is according to the molecular clone HXB2.

2.3.3 Inhibitory efficacies of anti-HIV lhRNAs against an integrated LTR target

The data described in sections 2.3.1 and 2.3.2 was generated following transient transfections of episomal target plasmids with the anti-LTR lhRNAs shown to be highly effective at post-transcriptional silencing of the cognate mRNAs. To assess the inhibitory efficacies of the lhRNAs against mRNAs expressed from an integrated target, a stable cell line that expressed hFluc from a subtype C LTR was generated.

Similarly to the transient transfections described in section 2.3.1, each of the anti-LTR lhRNAs, an empty vector control (mock) as well as a negative control lhRNA targeted to HBV were transfected into the LTRc-hFluc HEK293 stable cell line in the presence or absence of exogenous Tat. The pattern of inhibition was similar to that observed following transient transfection, with all 3 lhRNAs showing highly significant inhibition ($\alpha = 0.99$; $p < 0.01$) of the target sequence in the presence and absence of Tat (Figure 14). Intriguingly, the non-specific lhRNA HBV also mediated significant inhibition ($\alpha = 0.99$; $p < 0.01$) with 33.07 % ± 6.07 % (no Tat) and 31.37 % ± 4.86 % (with Tat) knockdown observed. In the absence of Tat, lhRNA LTR 1 mediated less significant suppression of the target ($\alpha = 0.99$; $p < 0.05$) with 14.93 % ± 1.19 % knockdown, while lhRNAs LTR 2 and 3 mediated 46.77 % ± 1.58 % and 58.83 % ± 1.28 %

inhibition respectively. In the presence of Tat, lhRNAs LTR 2 and 3 showed 61.80 % ± 3.29 % and 74.57 % ± 2.10 % suppression respectively.

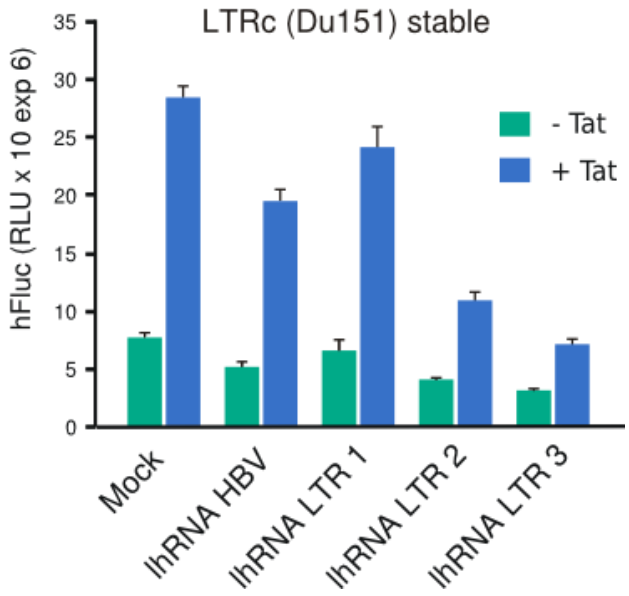


Figure 14. Anti-HIV lhRNA inhibition of cognate RNAs expressed from an integrated LTRc (Du151) target. Three lhRNAs and controls were transfected in the absence (turquoise bars) and presence (blue bars) of exogenous Tat in HEK293 cells that stably expressed subtype C LTR-driven hFluc. Data are expressed as the average relative light units (RLU) of hFluc normalised to an empty vector control (mock) determined 48 hours post-transfection (n = 3, ± SEM).

Overall these data supported the previous observation that lhRNA LTR 3 was most effective when targeted to a subtype C LTR reporter sequence, and suggested that an integrated target was amenable to lhRNA-mediated inhibition. Recent data has shown that HIV integration is a highly co-ordinated concert of interactions (Bera, Pandey*, et al.*, 2009) and that the transcription factor binding sites located within the enhancer region of the 5' LTR synergise with IN to ensure that the pre-integration complex is tethered to transcriptionally active host chromatin (Felice, Cattoglio*, et al.*, 2009). This ensures that HIV proviral DNA is preferentially integrated within promoters, regulatory sequences or other euchromatin. The number of integrated HIV copies per infected cell has an important bearing on the success of any anti-retroviral therapy including RNAi-based effectors such as lhRNAs. In one study, integrated HIV was found in > 90 % of the active genes in CD4+ memory T cells taken from HAART-suppressed patients (Duverger, Jones*, et al.*, 2009). Thus any therapeutic agent targeted to integrated HIV needs to be present at concentrations capable of controlling transcription of nascent viral RNA from numerous target sites. The lhRNAs used in this study were expressed from U6 (Pol II) promoters that transcribe

high quantities of mRNA. While this has been linked to cellular toxicity and even fatality in mice (see sections 2.3.6 and 5.3), one advantage may be that U6-expressed lhRNAs are transcribed at levels capable of suppressing highly active HIV.

2.3.4 Inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA LTR 2

Dicer recognises a 2 bp 3' overhang within dsRNA and cleaves to form a siRNA that can then be loaded into RISC [section 1.8.1 and (Elbashir, Harborth*, et al.*, 2001; Elbashir, Lendeckel*, et al.*, 2001)]. Should the guide strand be correctly processed and loaded, silencing of the cognate mRNA can then occur. However, should Dicer cleave *within* the intended guide strand, or should the correct guide strand be processed by Dicer but incorrectly loaded into RISC, silencing will not occur. The data generated from the experiments in sections 2.3.1 and 2.3.3 demonstrated that anti-LTR lhRNAs effectively suppressed their target sequences. However the data did not reveal which region(s) of the duplex stem was actively inhibiting the targets. To better understand this, a 'tiling array' was designed whereby the entire 63 bp target sequence of lhRNA LTR 2 was cropped into 5 tiled segments that overlapped each other by 11 bp (Figure 15). Each tile was 22 bp in length thereby representing a potential siRNA target along the duplex. These were individually cloned into the 3' UTR of hRluc within a dual luciferase vector and assessed in transient transfection assays with either an empty vector control (mock), lhRNA LTR 2 or a non-specific lhRNA targeted to HBV.

Luciferase activity assessed 48 hours post-transfection revealed highly significant suppression ($\alpha = 0.99$; $p < 0.01$) only when the complete LTR 2 sequence or the first 22 bp that corresponded to target 'tile A' were targeted (Figure 16). None of the remaining target tiles (B to E) were significantly inhibited either by lhRNA LTR 2 or the non-specific control lhRNA targeted to HBV ($p > 0.05$). The observation that tile A was the only sequence to be inhibited by lhRNA LTR 2 (70.53 % ± 1.69 %) suggested that the suppression of the complete target sequence (62.10 % ± 6.14 %) was probably due to inhibition of the initial 22 bp that corresponded to tile A.

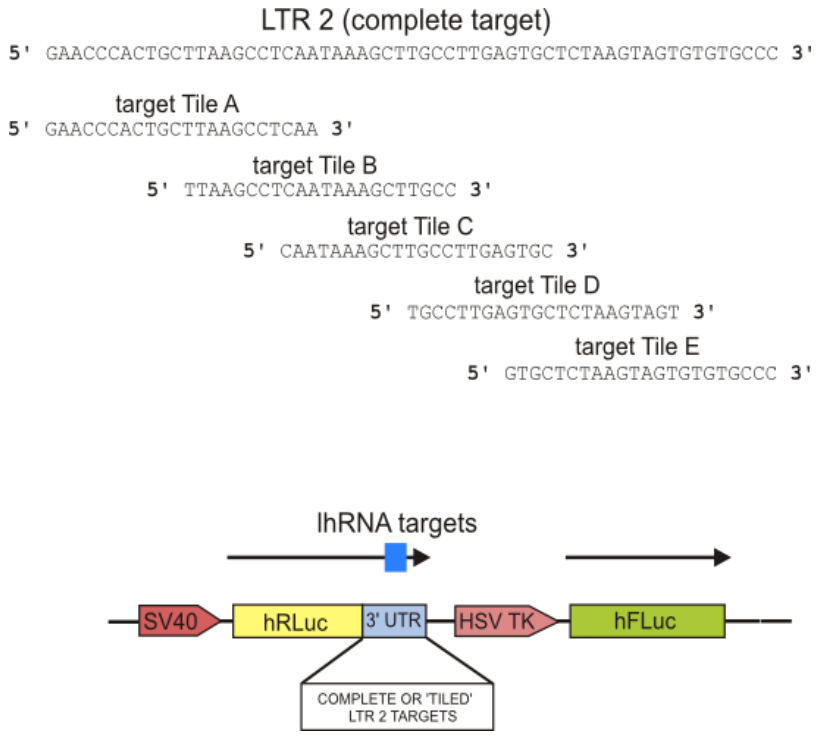


Figure 15. Tiling sequences used to generate a set of dual luciferase target reporter vectors. In the upper panel, the complete lhRNA LTR 2 target sequence, as well as 5 overlapping tiled segments (target Tiles A to E), are depicted and the latter are aligned relative to the complete sequence. All 6 sequences were cloned into the 3' UTR of hRluc within the dual luciferase psiCheck2 vector that expressed hFluc from a separate HSK TK promoter within the same plasmid (bottom panel).

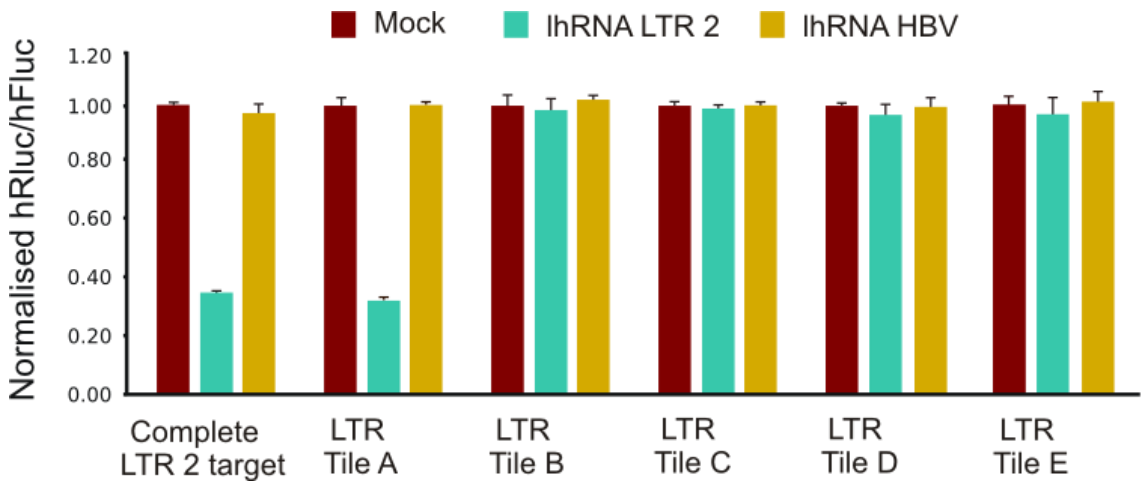


Figure 16. Tiling array to assess the inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA LTR 2. A dual luciferase reporter system was used to assess inhibition of the complete LTR 2 sequence as well as 5 tiled segments of the same target that had been individually cloned into the 3' UTR of hRluc. Forty eight hours post-transfection lhRNA LTR 2-mediated target suppression (turquoise bars) was determined in HEK293 cells. Data are expressed as a ratio of hRluc to hFluc normalised to an empty vector (mock – red bars) control (n = 3, ± SEM). A non-specific lhRNA was included in each experiment (lhRNA HBV – yellow bars).

The lhRNA LTR 2 tiling array data was interesting as it suggested functional siRNAs may not have been produced across the entire stem of the duplex and in the context of lhRNA LTR 2, were predominantly processed from the first Dicer cleavage position. To ensure that the pattern of functional siRNA production from a lhRNA construct was not specific to anti-LTR lhRNAs, a similar dual luciferase-based tiling array was conducted using the lhRNA HBV cassette targeted to an HBV-specific set of target tiles (see Appendix A3.1 for a full description). Using HEK293 cells, each target tile vector as well as the complete target was co-transfected with either an empty vector control (mock) or lhRNA HBV and luciferase activity was assessed 48 hours later (Figure 17).

Similarly to the previous tiling array, lhRNA HBV-mediated suppression was highly significant ($\alpha = 0.99$; $p < 0.01$) when targeted to the complete sequence (89.60 % ± 1.01 %) and target tile A (88.80 % ± 1.41 %). However unlike the previous tiling array, highly significant suppression of tiles B (32.90 % ± 1.70 %) and C (59.90 % ± 1.33 %) were also observed although with lower efficacy than against tile A. The target tile vectors overlapped each other by 14 bp thus target tile B straddled the putative initial Dicer cleavage site of lhRNA HBV that corresponded to tile A. Thus siRNAs processed from the first Dicer reaction would have only partial complementarity to tile B and mediate poor silencing efficacy of this sequence. More effective lhRNA HBV silencing of tile C compared to tile B may have been due to improved hybridisation of a putative second Dicer-generated siRNA from lhRNA HBV. Target tiles D and E were both poorly silenced by lhRNA HBV.

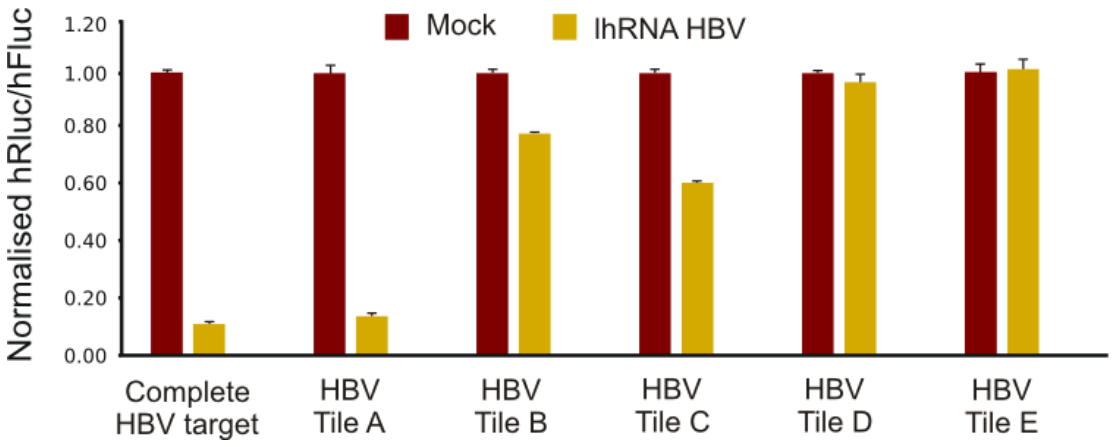


Figure 17. Tiling array to assess inhibitory efficacies of guide sequences produced along the duplex stem of lhRNA HBV. A dual luciferase reporter system was used to assess inhibition of the complete HBV X ORF sequence as well as 5 tiled segments of the same target that had been individually cloned into the 3' UTR of hRluc. Forty eight hours post-transfection lhRNA HBV-mediated target suppression (yellow bars) was determined in HEK293 cells. Data are expressed as a ratio of hRluc to hFluc normalised to an empty vector (mock – red bars) control (n = 3, ± SEM).

The data generated from the two tiling arrays led to several working hypotheses that may have explained the results. Firstly, Dicer may have recognised the 3' overhang and cleaved to generate one siRNA only from the first position (that corresponded to tile A). The HIV tiling array data supported this hypothesis (Figure 16). However, as the HBV specific target tiles A, B *and* C were inhibited by lhRNA HBV, this suggested that siRNAs were produced from more than just the initial 22 bp of the lhRNA stem. Another hypothesis proposed that Dicer was generating up to three siRNAs from the lhRNA stem region but a greater number of functional siRNAs were being produced from the first cleavage reaction. Both the HIV and HBV tiling array data supported this hypothesis. The observation that lhRNA HBV mediated greater silencing of target tile C as compared to tile B led to a third hypothesis that some of the siRNAs being produced were mismatched with respect to their target sequences. Alternatively, functional siRNAs that corresponded to each of the target tiles within the arrays may have been produced by Dicer but the guide strands were being degraded thereby preventing silencing of target mRNAs. A final hypothesis related to strand bias within the lhRNA duplex, such that siRNAs corresponding to target tiles B or C may have been produced from the sense strand. However, this was unlikely as the G:U mismatches within the sense strand of the lhRNAs ensured that siRNAs produced from this strand were not complementary to the target tiles. To dissect out more data regarding which region(s) of the lhRNAs were generating active siRNAs, a Northern blot assay was completed.

2.3.5 Northern blot analysis of lhRNA HBV

The differences observed between the HIV and HBV tiling arrays in section 2.3.4 suggested that Dicer processing of lhRNAs was not restricted to the first 22 bp only but that siRNAs produced by second and third Dicer-mediated cleavage reactions may have been mismatched with respect to their target sequences. As lhRNA HBV had mediated inhibition of tile A as well as target tiles that corresponded with sequences located closer to the loop side of the hairpin, Northern blot analysis was completed on this construct. In addition, a shRNA that targeted HBV tile D had previously been generated within the same research laboratory and was included as a control.

Three separate probes which together spanned the length of lhRNA HBV were used to detect guide sequence RNA in HEK293 cells that had been transfected with either an empty vector control (mock), lhRNA HBV or the shHBV control. Precursor unprocessed shHBV was detected in all 3 blots as a band at ~60 nt and a U6 snRNA probe that was used to detect small nuclear RNAs showed equal loading of RNA across the blots (Figure 18). Probe A (23 nt), located at the 3' overhang end of the lhRNA and corresponding to tile A, detected a putative

guide sequence of ~23 nt only in cells transfected with lhRNA HBV. Probe B (23 nt) detected putative ~23 nt guide sequences in lhRNA HBV- and shHBV-transfected cells. Probe C (25 nt) detected a ~23 nt guide sequence as expected for cells transfected with shHBV but no guide was detected for cells transfected with lhRNA HBV.

Importantly, previous data had shown shHBV to silence target tile D effectively (Carmona, Ely*, et al.*, 2006) suggesting that the target sequence was amenable to RNAi. Probes B and C overlapped by six nt and were designed to be complementary to lhRNA HBV and shHBV. However, the observation that both probes only detected guide sequences in shHBV-transfected cells suggested that no siRNAs corresponding to target tile D (or E) were produced from that portion of the lhRNA duplex stem. Alternatively, siRNAs may have been produced from these regions but the lack of inhibition noted in the tiling array did not support this.

Overall these data suggested that Dicer efficiently cleaved functional siRNAs from the 3' stem base end of lhRNAs and while subsequent cleavage reactions yielded additional functional siRNAs, processing was less efficient as Dicer moved towards the loop of the duplex.

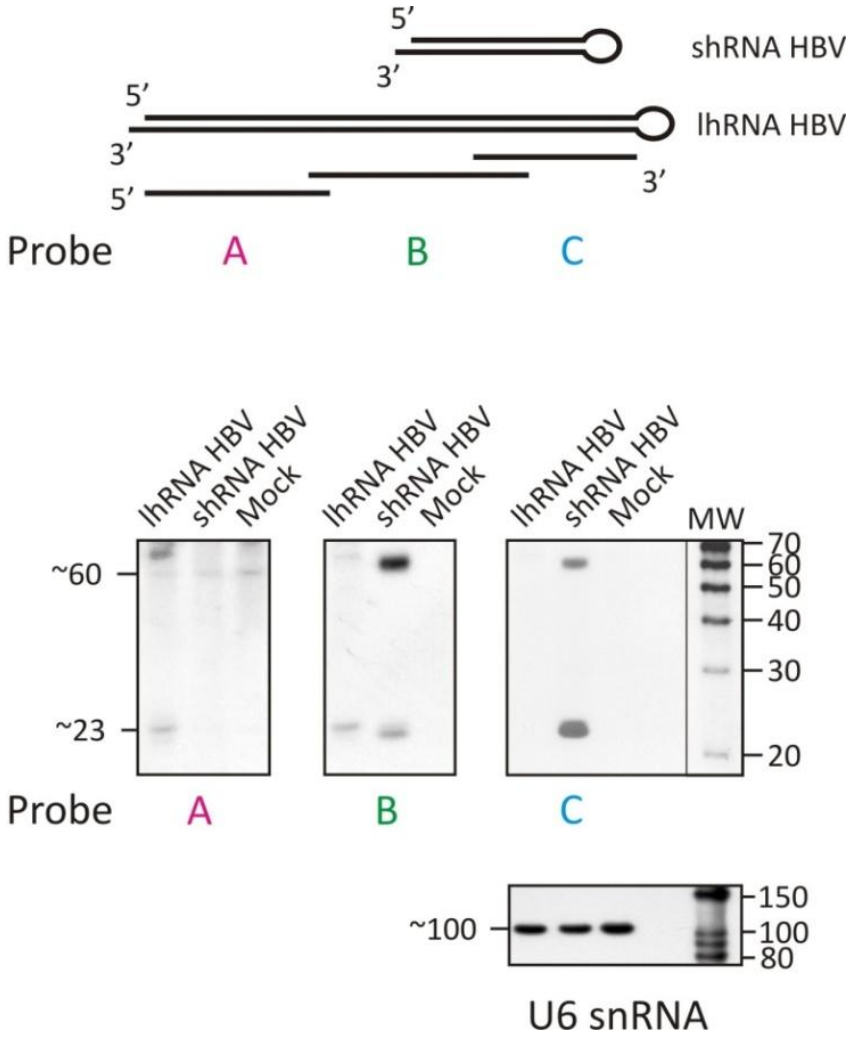


Figure 18. Northern blot analysis to detect putative guide sequences produced from lhRNA HBV. The upper panel depicts a schematic illustration of lhRNA HBV as well as a shHBV control with 3 ssDNA oligonucleotide probes (A, B and C) shown relative to the RNAi mimics. The bottom panel depicts a Northern blot following sequential individual hybridisations of all 3 probes as well as U6 snRNA oligonucleotides to RNA from HEK293 cells transfected with shRNA HBV, lhRNA HBV or an empty vector control (mock). Approximate sizes (nt) of hybridised bands and molecular weight markers (MW) are indicated. Repeated hybridisation reactions showed the data to be reproducible.

2.3.6 Assessment of possible cellular toxicity induced by anti-HIV lhRNAs

The introduction of exogenous RNA into mammalian cells, particularly dsRNA > 30 bp in length (Caplen, Parrish*, et al.*, 2001), had been linked to the production of an interferon (IFN) response that may ultimately lead to mRNA degradation and cell death (see section 5.3). Synthetic siRNAs and expressed shRNAs < 30 bp had been shown to evade the IFN response but the effect of lhRNAs was largely unknown. As an index of activation of the IFN response, IFN-β mRNA concentrations were measured using real-time RT-PCR on RNA extracted from HEK293 cells that had been transfected with either one of the LTR-specific lhRNAs, an empty vector

control (mock) or a dsRNA construct called poly (I:C) that was a known inducer of IFN in this cell line (Loseke, Grage-Griebenow*, et al.*, 2006; Reynolds, Anderson*, et al.*, 2006). Absolute quantification of the IFN-β mRNA levels within each sample was determined by normalisation to the housekeeping gene GAPDH.

Relative to poly (I:C) none of the lhRNAs or the mock treated samples elicited a statistically significant increase in IFN-β mRNA ($\alpha = 0.99$; $p < 0.01$), both in the absence of presence of Tat (Figure 19). In addition, cell morphology as measured by light field microscopy and transfection efficiencies as determined by GFP expression were the same for control treated cultures as well as those that received anti-LTR lhRNAs (Figure 20). Taken together, these data showed that the anti-LTR lhRNAs could be efficiently transfected and were non-immunostimulatory as measured by IFN-β expression when introduced into HEK293 cells.

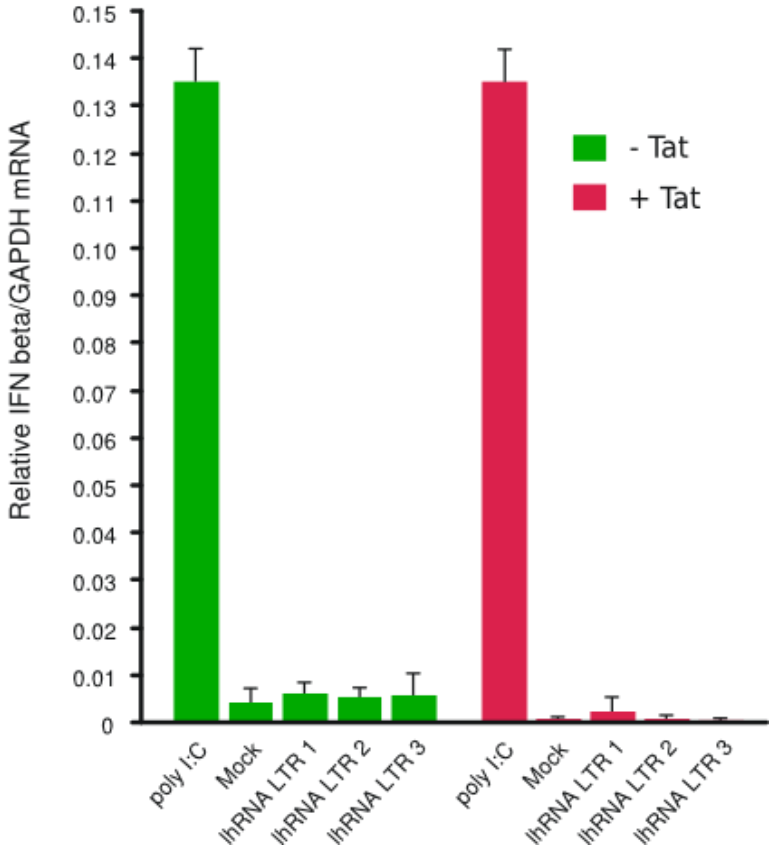


Figure 19. Quantitative real-time RT-PCR analysis to determine IFN-β levels in HEK293 cells following transfection with anti-HIV lhRNAs. IFN-β mRNA expression relative to the housekeeping gene GAPDH was determined 48 hours post-transfection in cells treated with one of the anti-LTR lhRNAs, an empty vector control (mock) or a positive control [poly (I:C)] in the absence (green bars) or presence (pink bars) of exogenous Tat. Data are represented as a ratio of IFN-β and GAPDH mRNA expression (n = 3, ± SEM).

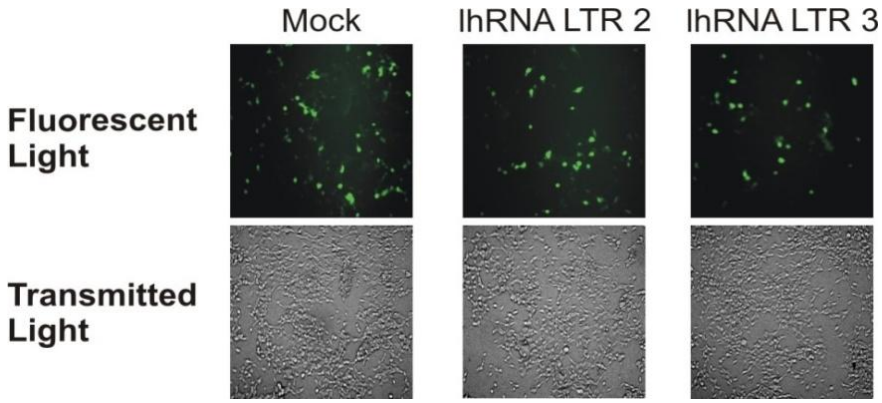


Figure 20. Anti-HIV lhRNAs do not affect cell morphology or transfection efficiency. HEK293 cells were co-transfected with two of the three anti-HIV lhRNAs and a GFP expressing plasmid, to verify equivalent transfection efficiencies between samples 48 hours post-transfection. The upper panel shows GFP expression under fluorescent light for empty vector control (mock) and lhRNA treated cultures and the lower panel represents transmitted light images of the same cultures. Data is represented at 100 x original magnification.

2.3.7 Suppression of HIV replication using LTR-specific lhRNAs

Although inhibitory efficacies of the LTR-specific lhRNAs against both episomal and integrated target reporter vectors (sections 2.3.1 and 2.3.3) showed these RNAi mimics to be effective suppressors, the target sequences used were expressed as single transcripts outside of their natural genomic context. To assess the ability of the anti-LTR lhRNAs to inhibit their cognate targets expressed from the context of an HIV genome, two separate assays were used. Initially the LTR-specific lhRNAs were tested against a subtype B infectious molecular clone of HIV (pNL4-3.Luc.R-E-) that expressed hFluc in place of the *nef* ORF and was capable of a single round of infection. HEK293 cells were co-transfected with pNL4-3.Luc.R-E-, a hRluc expressing plasmid (that served as a background luciferase control) and either one of the LTR-specific lhRNAs or the non-specific lhRNA HBV or a shRNA complementary to HIV Rev that had been shown to effective against the molecular clone (Lee, Dohjima*, et al.*, 2002).

Measurement of luciferase activity 48 hours post-transfection revealed that all three lhRNAs mediated highly significant inhibition of pNL4-3.Luc.R-E- ($\alpha = 0.99$, $p < 0.01$) as compared to lhRNA HBV (Figure 21). Overall, lhRNA LTR 2 showed the most inhibition (80.00 % ± 2.64 %) although none of the lhRNAs were as effective as the shRev (94.83 % ± 0.08 %). Notably, the pattern of inhibition was similar to that previously observed for the episomal subtype B LTR target vectors (Figure 12) suggesting that the ability of lhRNAs to suppress their cognate targets was not significantly affected by the context from which that target sequence was expressed.

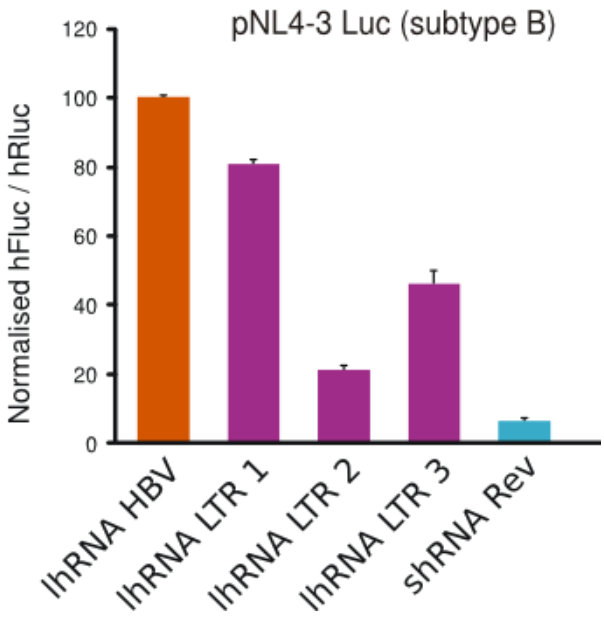


Figure 21. Anti-HIV lhRNAs inhibit replication of a subtype B HIV molecular clone. All three LTR-specific lhRNAs (purple bars), a non-specific lhRNA (HBV – orange bar) as well as a shRNA (Rev – blue bar) were co-transfected with the subtype B infectious clone pNL4-3.Luc.R-E- in HEK293 cells. Luciferase activity was determined 48 hours post-transfection and data are expressed as a ratio of hFluc to hRluc activity normalised to lhRNA HBV (n = 3, ± SEM).

The second assay used to score the ability of anti-HIV lhRNAs to inhibit their targets in the context of a genome utilised a South African R5-tropic subtype C HIV isolate (FV5). As an infectious replicating HIV strain, FV5 had continually evolved during the course of infection to develop into a pool of similar but probably non-identical subtype C strains. The use of FV5 in a challenge assay therefore provided a unique opportunity to assess the inhibitory efficacies of the anti-LTR lhRNAs against a non-isogenic population of target sequences expressed from within their native genomic context. The fibroblast cell line U87.CD4.CCR5 bearing CD4 and CCR5 receptors required for HIV infection was transfected with the anti-HIV lhRNAs, the non-specific control lhRNA HBV and positive control shRev before being challenged with FV5 using a $TCID_{50}$ of 1000/mL. Inhibition of HIV was assessed four days post-infection by determining Gag p24 antigen concentrations (Figure 22).

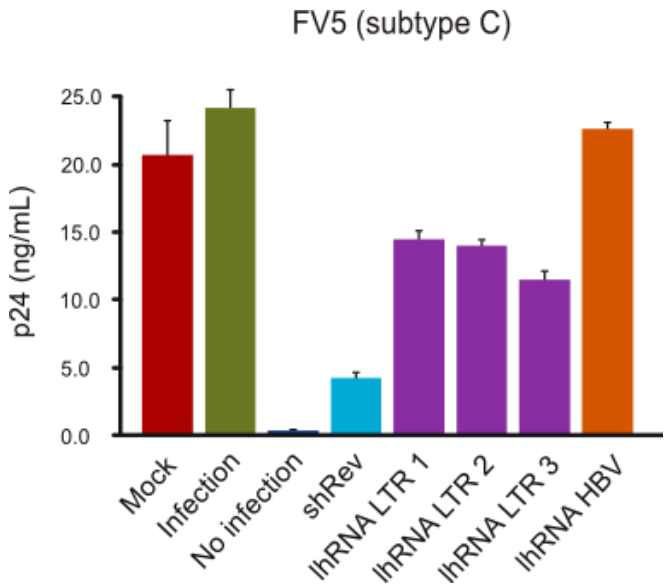


Figure 22. Challenge assay to assess anti-HIV lhRNAs against a subtype C viral isolate of HIV. All three LTR-specific lhRNAs (purple bars), a non-specific lhRNA (HBV – orange bar) as well as a shRNA (Rev – blue bar) were transfected into U87.CD4.R5 cells prior to challenge with viral isolate FV5 at a $TCID_{50}$ of 1000/mL. Viral Gag p24 antigen levels were determined four days post-infection by ELISA (450 nm). Values were converted to ng/mL using a standard curve supplied with the ELISA kit and data are expressed as p24 antigen levels in ng/mL (n = 3, ± SEM).

Cells that received only FV5, empty vector (mock) treated cells or those transfected with lhRNA HBV prior to FV5 infection showed no significant differences ($\alpha = 0.99$; $p > 0.05$) indicating that the viral isolate was indeed infectious and the non-specific lhRNA control was unable to inhibit wild type HIV replication as expected. All three anti-LTR lhRNAs showed highly significant inhibition of FV5 ($\alpha = 0.99$; $p < 0.01$) with lhRNA LTR 3 mediating 57.98 % ± 9.28 % suppression. One notable observation was that the highest level of inhibition overall was mediated by shRev (84.52 % ± 9.54 %). This was significantly greater ($p < 0.05$) than the suppression mediated by lhRNA LTR 3 and the implications of these differences are discussed elsewhere (section 2.4). Interestingly, the overall trend in lhRNA-mediated inhibition of FV5 was similar to that previously observed for the episomal and integrated subtype C LTR target vectors (Figures 12 and 14). This suggested that either the majority of targets within the FV5 pool shared the same sequence, or alternatively that slight variations between target sequences within the pool did not significantly impair the inhibitory abilities of the lhRNAs. This latter argument was more likely given that FV5 had been isolated from a drug-naive AIDS patient who had been infected for a number of years thereby providing time for sequence mutations to occur and accumulate. Thus these data provided a very promising proof-of-principle result that lhRNAs were able to inhibit a wild type infection using a natural HIV isolate.

2.4 Discussion and conclusions

The approach used in this study has been to design three separate RNA Pol III-expressed lhRNAs targeted to the viral 5' LTR and to assess the inhibitory efficacies of these RNAi mimics against three distinct target sequences in both episomal and integrated forms and, importantly, against wild type HIV. In addition, Dicer processing of lhRNAs has been investigated as a means of understanding how these constructs are cleaved to form functional siRNAs.

The observation that U6-expressed lhRNAs can target the HIV LTR whether in an episomal or integrated context suggests that these RNAi mimics would be effective against active and inactive forms of the proviral promoter. However, the suppression levels did not exceed 50 % regardless of the type of reporter used and this level would not be acceptable in a therapeutic strategy. The inhibition mediated by lhRNA LTR 1 despite being targeted to the TAR loop was interesting given that previous studies had shown the secondary structure of the TAR loop to interfere with siRNA-mediated inhibition. Notably, the repeated trend of inhibition whereby lhRNA LTR 2 was more efficacious against a subtype B target and lhRNA LTR 3 against a subtype C LTR suggested that sequence variation influenced these mimics in a manner similar to shRNAs. Mismatches within the seed region of the guide strand(s) probably decreased the overall inhibitory efficacy of the lhRNA and sequence alignment of the LTRb and LTRc promoters revealed mismatches that may have been responsible for the differences observed. The presence and absence of exogenously supplied Tat generated processive and non-processive LTR-driven transcripts respectively and the lhRNAs were efficacious under both conditions. Notably though, the lhRNA-mediated inhibition of the LTR in the absence of Tat suggested a novel strategy to impair the completion of the HIV life cycle. Following integration of the virus and in the absence of Tat, small amounts of truncated viral RNA are transcribed from the LTR as RNA Pol II remains de-phosphorylated. This low level transcription is sufficient to generate Tat which phosphorylates RNA Pol II and creates a positive feedback resulting in increased transcription (see section 1.4.4). As Tat is required to ensure that processive transcription occurs and all viral RNAs are transcribed, inhibition of the *non-processive* transcripts as mediated by lhRNAs LTR 2 and 3 may block the virus at a crucial point in the life cycle. The inhibition mediated by the non-specific lhRNA HBV was only observed when targeted against the integrated subtype C LTR (Du151; Figure 14) both in the presence and absence of exogenous Tat. The suppression was observed despite repeated preparation of the lhRNA HBV construct and numerous technical and biological replicates of the experiment. Analysis of the Du151 LTR sequence and alignment with the seed region of the non-specific lhRNA did not reveal any base pairing. Thus the observed suppression was difficult to explain but was most likely due to an unintended off-target effect.

The design of a tiling array to investigate the activity of siRNAs processed from various regions along the duplex stem of a lhRNA was novel and generated an intriguing set of data that suggested Dicer efficiently processes the first two putative siRNAs along the lhRNA stem only. This has implications for the design of future lhRNAs as modifications may be required to ensure that functional and highly potent siRNAs are produced from all potential Dicer cleavage reactions and additional data related to this is discussed further below.

The potential toxicities of the expressed lhRNAs were assessed using poly (I:C) as a positive marker of IFN within HEK293 cells. However, poly I:C traverses the endosome and induces the IFN response via toll-like receptors (TLRs) whereas expressed RNAi mimics do not traverse endosomes and activate the pathway via PKR or OAS1 (see section 5.3). Thus a more appropriate positive control would be an expressed construct that activates IFN similarly to shRNAs and lhRNAs. This remains to be developed and should be an important consideration for future lhRNA studies. In addition, the ability of Dicer to process the lhRNA duplex into multiple siRNAs could generate guides capable of targeting unintended sequences such as host mRNAs. These are known as off-target effects (OTEs) and have been linked to death in mice [(Grimm, Streetz*, et al.*, 2006) and refer to section 5.3 for an account of the mammalian IFN response]. While the lhRNA sequences designed here were subjected to a BLAST search to minimise the chances of OTEs, the possibility remains. The data presented in Figures 19 and 20 suggested that OTEs as characterised by the IFN response were not elicited in the presence of the lhRNAs although a more exhaustive investigation could include a genome-wide microarray assay.

The use of a wild type HIV isolate provided a unique and extremely important piece of data as previous studies to assess RNAi mimics against HIV utilised infectious molecular clones such as pNL4-3.Luc.R-E- or LAI. While these assays are reliable, technically undemanding and robust, the target sequences are isogenic and thus do not mimic the diversity present in natural HIV infections. FV5 was a mixed population of actively replicating HIV and the inhibition mediated by the anti-LTR lhRNAs, while only ~ 50%, was a vital proof-of-principle argument for the use of lhRNAs as a combinatorial strategy against HIV.

Concurrent to this work, data was published describing expressed lhRNAs targeted to HIV *nef*, *tat* and *rev* that were designed with 300 bp stem and 46 bp loop regions (Konstantinova, de Vries*, et al.*, 2006). EF1α-driven lhRNAs displayed poor efficacy against an HIV molecular clone (LAI) as well as a luciferase-based reporter target although this was not surprising given that EF1α is an RNA Pol II promoter and thus the lhRNAs would not have been transcribed with defined 5' and 3' ends required for Dicer recognition. Similarly, LTR-driven lhRNAs with a 1000 bp spacer sequence included in the loop showed poor inhibition of LAI and substitution of the promoter with a TetO system did not improve lhRNA-mediated inhibition. The

inclusion of a leader sequence between the LTR promoter and *nef* target of one of the lhRNAs dramatically increased the suppressive efficacy of this construct against LAI as well as two LAI-based clones that were previously shown by the same group to be resistant to a *nef*-specific shRNA (Das, Brummelkamp*, et al.*, 2004). While these data showed impressive inhibition of *nef*-resistant HIV, the use of the LTR as a promoter (transcribed by RNA Pol II) in conjunction with a leader sequence suggests that the observed inhibition was not due to an RNAi-mediated effect. Finally, as no data were provided to show how these lhRNAs were processed into functional siRNAs or the specific target sequences inhibited, these data do not clearly extend the knowledge base regarding anti-HIV lhRNAs.

The tiling array and Northern blot data described in sections 2.3.4 and 2.3.5 suggested that the 'molecular ruler mechanism' inherent to the PAZ domain of Dicer influences the processing of lhRNA duplex regions into functional siRNAs (refer to the section entitled 'Nuclear export and cytoplasmic processing' in section 1.8.1). Based on these findings, the spacing between putative guide sequences in a set of anti-HIV lhRNAs was manipulated and while improvements could be made to first and second position siRNAs by adjusting spacing arrangements at the junction site, silencing of third position siRNAs (closest to the loop) remained largely ineffective (Saayman, Barichievy*, et al.*, 2008). This was supported by data from two other groups, one of which showed the positioning of two siRNAs within the lhRNA stem was crucial for their functioning (Liu, Haasnoot*, et al.*, 2007) while the other also suggested that Dicer seems to process the first two siRNAs from a lhRNA duplex encoding three putative guides (Sano, Li*, et al.*, 2008). Taken together, these data suggest that a stem region of ~50 bp is long enough to generate two functional siRNAs if the guide sequences are not immediately adjacent. Obviously, lhRNAs with less than one guide sequence are simply shRNAs but an upper limit to lhRNA stem length has not been clarified. In one set of experiments, a lhRNA of 80 bp was efficacious against a reporter vector but did not induce long-term suppression of HIV and did not produce functional siRNAs from the entire length of the stem (Sano, Li*, et al.*, 2008). Indeed, the most effective lhRNA from this work was a 53 bp lhRNA that produced two distinct siRNAs in a manner similar to Saayman *et al* (Saayman, Barichievy*, et al.*, 2008). Recently published data suggested that a minimum of 43 bp are needed to obtain 2 functional siRNAs (Liu, von Eije*, et al.*, 2009). The authors further suggested that lhRNAs with stem duplex regions of 66 bp could produce 3 functional siRNAs but the inhibition mediated by the siRNA generated closest to the loop was not as pronounced as the other two. Taken together, these studies have revealed conditions required for the design of lhRNAs that generate potent siRNAs from more than one position in the duplex. Importantly, only one long-term study to date has been conducted to see if a combination of two or more siRNAs generated from a single lhRNA is sufficient to inhibit HIV escape mutants. A single lhRNA targeted to *nef*, *pol* and the *tat*/*rev*

overlapping sequence was shown to suppress viral replication for 49 days although escape mutants were detected particularly in the presence of high viral load (Liu, von Eije*, et al.*, 2009). Additional long-term studies to assess viral escape in the presence of expressed lhRNAs are required to support this data.

In conclusion, the lhRNAs described in this chapter are functional in suppressing LTR-driven luciferase activity across subtype B (HIV-1_{LAI}) and C (HIV-1_{Du151}) sequences, both in the presence and absence of Tat and do not induce a detectable IFN response. Importantly, the anti-LTR lhRNAs showed suppression of a primary HIV-1 isolate comprising a varied pool of target sequences. New data have been provided on how these mimics may be processed by Dicer with implications for the future design of potent lhRNAs. Yet despite these advances, modifications and improvements are required as none of the lhRNAs reduced HIV to the same levels as an expressed shRNA. Additional data has been published to show how these mimics can be designed to ensure their overall potency but long-term studies need to be completed to ensure that lhRNAs can indeed function as a combinatorial RNAi mimic capable of suppressing viral escape. This suggests that alternative combinatorial RNAi approaches need to be explored and pri-miRNA mimics represent a novel strategy to be investigated against HIV.

Chapter three: expressed polycistronic pri-miRNAs targeted to discrete HIV mRNAs as a combinatorial RNAi strategy

3.1 Introduction

The data presented in the previous chapter demonstrated that anti-HIV lhRNAs may be an effective combinatorial modality but modifications are required for these constructs to generate numerous functional siRNAs. In addition, the HIV 5' LTR was shown to be amenable to inhibition using expressed combinatorial RNAi mimics. Other combinatorial RNAi strategies against HIV have utilised variations in the number and arrangement of expressed shRNAs that enter the RNAi pathway as pre-miRNA mimics at the level of Dicer (see section 1.10.2). An alternative approach that has not been attempted against HIV is the use of 'pri-miRNA mimics' (endogenous pri-miRNAs encoding exogenous guide strands) that enter the RNAi pathway as Drosha substrates (see section 1.10.2). These present an exciting avenue particularly since it has been shown that the endogenous guide strand of pri-miRNA 30 (pri-miR30) can be replaced with an exogenous guide sequence capable of binding its own cognate target (Zhou, Xia*, et al.*, 2005; Sun, D, Melegari*, et al.*, 2006). Endogenous pri-miRNAs are generally located within protein-coding genes and are thus expressed by RNA Pol II activity (Lagos-Quintana, Rauhut*, et al.*, 2003; Lee, Y, Kim*, et al.*, 2004) although some RNA Pol III expression has been described (Borchert, Lanier*, et al.*, 2006). The use of RNA Pol II expression would allow for tissue-specific production as these promoters are naturally required for the synthesis of protein-coding mRNAs (Cramer, Armache*, et al.*, 2008). More importantly, the use of Pol II-driven RNAi mimics would dramatically reduce the saturating toxicities observed with Pol III-expressed effectors (see section 5.3). U6-expressed shRNAs caused fatalities in mice as a result of Exp5 saturation and while the concentrations of shRNAs used in this study were high, over-expression of these RNAi mimics was clearly deleterious (Grimm, Streetz*, et al.*, 2006). As RNA Pol II promoters naturally transcribe lower levels of transcripts *and* can function in a tissue-specific manner, they provide an enticing option for pri-miRNA mimic transcription.

For a combinatorial RNAi approach against HIV to be effective, it has been estimated that a minimum of four target mRNAs need to be suppressed (ter Brake, Konstantinova*, et al.*, 2006; ter Brake, Hooft*, et al.*, 2008). Incorporating this into a pri-miRNA-based combinatorial strategy would require the expression of four separate guide sequences either from four individual pri-miRNA backbones expressed in tandem, or four repeats of the same pri-miRNA backbone each expressing one anti-HIV guide sequence. Two studies have been published that focused on the latter option. In one instance, two identical *Homo sapien* (hsa) pri-miR30a-based

mimics transcribed in tandem from a single Pol II promoter were modified to express guide sequences targeted to *Skp2* mRNA [S-phase kinase-associated protein 2; (Sun, D, Melegari*, et al.*, 2006)]. Interestingly, two copies of pri-miR30-Skp2 were more efficacious than one when targeted against a GFP-based reporter, but the addition of a third similar scaffold only resulted in marginal improvement in GFP suppression. In the second study, two, four or eight copies of a hsa pri-miR155 backbone that had been modified to express a guide sequence targeted to mouse *neuroD1* mRNA were expressed from a single RNA Pol II promoter (Chung, Hart*, et al.*, 2006). Using a luciferase-based reporter assay, progressive enhancement of the guide sequence efficacy was observed as the number of pri-miRNA scaffolds increased. This was in direct contrast to the data published by Sun *et al* but the reasons for the discrepancies remain unclear. While the use of different pri-miRNA backbones or the lengths of the ss flanking regions may have contributed to the inconsistent results, the feasibility of tandemly repeating pri-miRNA mimics under the control of a single Pol II promoter was clearly demonstrated in both publications. This provided a foundation from which to investigate the use of pri-miRNA scaffolds in a combinatorial RNAi system targeted to various HIV mRNAs.

The design of a combinatorial RNAi system that encompasses multiple, distinct pri-miRNA scaffolds designed to express discrete exogenous guide sequences had not previously been investigated. However, for this approach to be feasible, the efficacies of exogenous guide sequences processed from various pri-miRNA backbones needed to be assessed. Based on the studies by Sun *et al* and Chung *et al*, hsa pri-miRNAs 30a and 155 were included in the current body of work and two additional pri-miRNAs (hsa pri-miR122 and hsa pri-miR31) were chosen based on data showing that these pri-miRNAs yield functional exogenous guide sequences (Ely, Naidoo*, et al.*, 2008). Importantly, as mature miRNAs can be excised from either arm of a pre-miRNA stem or even from both arms (Khvorova, Reynolds*, et al.*, 2003), the dominant guide strand of each of the four pri-miRNAs was designated according to published data downloaded from the Sanger miRNA Registry database (http://microrna.sanger.ac.uk/sequences/). Notably, this revealed that 3 of the pri-miRNAs (hsa pri-miRNAs 122, 31 and 155) generated guide strands predominantly from their 5' arms while pri-miR30a predominantly generated 3' arm guides. This provided the opportunity to explore an additional characteristic that may have affected guide sequence production.

In this chapter, the relationship between pri-miRNA scaffold and guide sequence efficacy was investigated using four individual pri-miRNA backbones (hsa pri-miRNAs 122, 30a, 31 and 155) each modified to produce one of four separate anti-HIV guide sequences. This approach was used in order to tease out the most efficacious scaffold/guide pairs with the aim of generating a combinatorial system ideally comprised of four different pri-miRNA backbones expressing four discrete anti-HIV guides. The requirement for four individual guide sequences

was based on the estimation provided by ter Brake *et al* (ter Brake, Konstantinova*, et al.*, 2006; ter Brake, Hooft*, et al.*, 2008) and the use of four discrete pri-miRNA backbones was hypothesised to decrease the chances of a recombination event that could excise one or more of the scaffolds.

3.2 Materials and methods

3.2.1 Cloning of individual expressed pri-miRNA mimics targeted to four separate HIV mRNAs

To construct different pri-miRNA scaffolds with anti-HIV guide sequences that were expressed from an RNA Pol II promoter, each pri-miRNA backbone in a group of four was modified to encode each of four separate anti-HIV guide sequences and cloned into the pCi-Neo mammalian expression vector (Promega, WI, USA) such that the cassettes were expressed from the Cytomegalovirus (CMV) promoter as follows: the sequences of four endogenous pri-miRNAs (hsa pri-miR122, GenBank® accession number Mi0000442 and co-ordinates 54269286 to 54269370 on chromosome 18; hsa pri-miR30a, GenBank® accession number Mi0000088 and co-ordinates 72169975 to 72170045 on chromosome 6; hsa pri-miR31, GenBank® accession number Mi0000089 and co-ordinates 21502114 to 21502184 on chromosome 9; hsa pri-miR155, GenBank® accession number Mi0000681 and co-ordinates 25868163 to 25868227 on chromosome 21) were downloaded from the Sanger miRNA Registry database (http://microrna.sanger.ac.uk/sequences/). The sequences included the 5' and 3' ss flanking regions of each pri-miRNA required for Drosha recognition and processing [refer to *'Expression and nuclear processing'* of section 1.8.1 and (Zeng and Cullen, 2005; Han, Lee*, et al.*, 2006)].

The endogenous guide strand of each pri-miRNA was substituted with each of four guide sequences taken from shRNAs that had been shown to be highly effective against HIV [(Lee, N S, Dohjima*, et al.*, 2002; Cave, Weinberg*, et al.*, 2006; Nishitsuji, Kohara*, et al.*, 2006; ter Brake, Konstantinova*, et al.*, 2006; Westerhout, Vink*, et al.*, 2006) and see Appendix A3.2]. These included shRNAs targeted to *int*, *gag*, LTR and *tat* mRNAs and guide sequence substitution generated a panel of sixteen anti-HIV pri-miRNA mimics (four anti-HIV guides in each of four pri-miRNA backbones).

The predicted secondary structure of each endogenous pri-miRNA as well as the anti-HIV mimics was analysed using the RNA folding program Mfold (http://mfold.burnet.edu.au/) as previously described (Mathews, Sabina*, et al.*, 1999; Zuker, 2003). Based on a 'nearest neighbour fit', Mfold is ideally suited to generating a predicted secondary structure for single hairpin sequences. As the ss flanking regions of the pri-miRNAs did not naturally form duplex structures, they were omitted from the Mfold analysis. The predicted secondary structure of the

pre-miRNA portion of each anti-HIV pri-miRNA mimic was compared to the endogenous counterpart to determine the shape and thermodynamic stability of each (Figure 26). Related to the former characteristic, the duplex regions, loop arrangements and bulge sequences of each mimic were analysed to determine if the guide sequence replacement altered the overall shape of each mimic as compared to their endogenous backbone. To determine if the predicted secondary structure of each mimic was energetically equivalent to the endogenous backbone, the free energy values for each predicted structure were compared to the wild type.

The entire sequence (~ 235 bp) of each anti-HIV pri-miRNA mimic was divided into four overlapping oligonucleotides and a two step approach (Figure 23) was used to generate the panel of modified pri-miRNA cassettes as follows: for pri-miR122 targeted to HIV *int* (pri-miR122 IN), an extension step was carried out using two oligonucleotide primers (IDT, USA) that overlapped by 38 nt called Pri-miR122-IN Forward (5' GAG TTT CCT TAG CAG AGC TGA CTA GCC ATT GCT CTC CGG CTT TGT CTA AAC TAT AAG CCG GAG 3') and Pri-miR122-IN Reverse (5' GGA TTG CCT AGC AGT AGC TAA AGA ACC ATT TCT CTC CGG CTT ATA GTT TAG ACA AAG CCG GA 3'). A PCR was carried out using Expand High Fidelity^{PLUS} PCR System reagents (Roche, Basel, Switzerland) on a Mastercycler® (BioRad, CA, USA). The PCR mixture included 0.5 µM of each primer, 1.5 mM $MgCl_2$, 200 µM dNTP mix, 2.5 U of Expand HiFi^{PLUS} Taq polymerase and 1 x final volume Expand HiFi^{PLUS} Reaction Buffer (proprietary information) in a final volume of 50 µL. PCR conditions comprised 3 minutes at 95 ^{o}C followed by 35 cycles of 30 seconds each at 95 ^{o}C, 58 ^{o}C and 72 ^{o}C followed by 10 minutes at 72 ^{o}C. The product formed the complete ds portion of each pri-miRNA mimic (5' arm, loop and 3' arm, Figure 23) and was excised and eluted from a 1 % agarose gel (see Appendices A1.2.3 and 1.2.4). Following extraction, the product was diluted to 10 ng/µL in distilled water and 2 µL of the dilution was used as template for a second PCR step using primers Pri-miR122 Flank Left containing *Xho*I and *Spe*I sites (5' GAT CCT CGA GAC TAG TCA GGT AAG GTG GAG GTG AAG TTA ACA CCT TCG TGG CTA CAG AGT TTC CTT AGC AGA GCT G 3', the *Xho*I site is underlined and the *Spe*I site is shown with a dotted underline) and Pri-miR122 Flank Right containing an *Nhe*I site (5' CGC TAG CGG ACA CCT GTG GAG AGA AAG AGC AAA CGA TGC CAA GAC ATT TAT CGA GGG AAG GAT TGC CTA GCA GTA GCT A 3', the *Nhe*I site is underlined). The second round primers overlapped the first round template by 20 nt and generated the 5' and 3' flanking regions of the pri-miRNA mimic respectively. The same reaction and cycling conditions were used during the second PCR step. The second round PCR product was cloned into the TA cloning vector pTZ57R/T (Fermentas, WI, USA) and candidate clones were screened using *EcoR*I and *Xho*I as described in section 2.2.1.

The same two step cloning strategy was used to generate the other anti-HIV pri-miRNA mimics with the following primer combinations: for pri-miR122 targeted to HIV *gag* (pri-miR122

Gag), the first round extension was completed using Pri-miR122 Gag Forward (5' GAG TTT CCT TAG CAG AGC TGA TAT TGA CGC TCT CGC ACC CTT TGT CTA AAC TAT AAG GGT GCG 3') and Pri-miR122 Gag Reverse (5' GGA TTG CCT AGC AGT AGC TAA CCT TGA CGC CCT CGC ACC CTT ATA GTT TAG ACA AAG GGT GC 3'), and the second round PCR was completed using the Pri-miR122 Flank Left and Right primers. For pri-miR122 targeted to HIV LTR (pri-miR122 LTR), the first round extension was completed using Pri-miR122 LTR Forward (5' GAG TTT CCT TAG CAG AGC TGT CTG AGG GAT CTC TAG TTA CTT TGT CTA AAC TAT AAG TAA CTA 3') and Pri-miR122 LTR Reverse (5' GGA TTG CCT AGC AGT AGC TAA CCT TGA CGC CCT CGC ACC CTT ATA GTT TAG ACA AAG GGT GC 3'), and the second round PCR was completed using the Pri-miR122 Flank Left and Right primers. For pri-miR122 targeted to HIV *tat* (pri-miR122 Tat), the first round extension was completed using Pri-miR122 Tat Forward (5' GAG TTT CCT TAG CAG AGC TGC TCT TCG TCG CTG TCT CCG CTT TGT CTA AAC TAT AAG CGG AGA 3') and Pri-miR122 Tat Reverse (5' GGA TTG CCT AGC AGT AGC TAC GAT TCA TCG GTG TCT CCG CTT ATA GTT TAG ACA AAG CGG AG 3'), and the second round PCR was completed using the Pri-miR122 Flank Left and Right primers.

For pri-miR30a targeted to HIV *int* (pri-miR30a IN), the first round extension was completed using Pri-miR30a IN Forward (5' CTG TTG ACA GTG AGC GCA GCC GGA GAG CAT CAT GGT TAG TCT GTG AAG CCA CAG ATG GGA CTA GCC AT 3') and Pri-miR30a IN Reverse (5' TCC GAG GCA GTA GGC AAA GCC GGA GAG CAA TGG CTA GTC CCA TCT GTG GCT TCA CAG ACT AAC CAT 3'), and the second round PCR was completed using Pri-miR30a Flank Left containing *Xho*I and *Spe*I sites (5' GAT CCT CGA GAC TAG TCA GGT GAG AGG TTA ACC CAA CAG AAG GCT AAA GAA GGT ATA TTG CTG TTG ACA GTG AGC G 3', the *Xho*I site is underlined and the *Spe*I site is shown with a dotted underline) and Pri-miR30a Flank Right containing an *Nhe*I site (5' CGC TAG CGA GGC CTC TAG AGC CGC CGG AAA CAA GAT AAT TGC TCC TAA AGT AGC CCC TTG AAG TCC GAG GCA GTA GGC A 3', the *Nhe*I site is underlined). For pri-miR30a targeted to HIV *gag* (pri-miR30a Gag), the first round extension was completed using Pri-miR30a Gag Forward (5' CTG TTG ACA GTG AGC GGA GGG TGC GAG AGT CCG TCA ATA TCT GTG AAG CCA CAG ATG GGA TAT TGA CG 3') and Pri-miR30a Gag Reverse (5' TCC GAG GCA GTA GGC AAA GGG TGC GAG AGC GTC AAT ATC CCA TCT GTG GCT TCA CAG ATA TTG ACG 3'), and the second round PCR was completed using the Pri-miR30a Flank Left and Right primers. For pri-miR30a targeted to HIV LTR (pri-miR30a LTR), the first round extension was completed using Pri-miR30a LTR Forward (5' CTG TTG ACA GTG AGC GCA GTA ACT AGA GAT CTC TCT CAG ACT GTG AAG CCA CAG ATG GGT CTG AGG GA 3') and Pri-miR30a LTR Reverse (5' TCC GAG GCA GTA GGC AAA GTA ACT AGA GAT CCC TCA GAC CCA TCT GTG GCT TCA CAG TCT GAG AGA 3'), and the second round PCR was completed using the Pri-miR30a Flank Left and Right primers.

For pri-miR30a targeted to HIV *tat* (pri-miR30a Tat), the first round extension was completed using Pri-miR30a Tat Forward (5' CTG TTG ACA GTG AGC GCA GCG GAG ACA GCT CGA TGA AGA GCT GTG AAG CCA CAG ATG GGC TCT TCG TC 3') and Pri-miR30a Tat Reverse (5' TCC GAG GCA GTA GGC AAA GCG GAG ACA GCG ACG AAG AGC CCA TCT GTG GCT TCA CAG CTC TTC ATC 3'), and the second round PCR was completed using the Pri-miR30a Flank Left and Right primers.

For pri-miR31 targeted to HIV *int* (pri-miR31 IN), the first round extension was completed using Pri-miR31 IN Forward (5' TAA CTT GGA ACT GGA GAG GAA CTA GCC ATT GCT CTC CGG CTT TGA ACT GGG AAA CTC CGG AG 3') and Pri-miR31 IN Reverse (5' TGC TGT CAG ACA GGA AAG ATA CTA ATC ATG GCT CTC CGG AGT TTC CCA GTT CAA AGC CGG AGA 3'), and the second round PCR was completed using Pri-miR31 Flank Left containing *Xho*I and *Spe*I sites (5' GAT CCT CGA GAC TAG TAA GGT AAG GTA AAA CAC TGA AGA GTC ATA GTA TTC TCC TGT AAC TTG GAA CTG GAG AGG A 3', the *Xho*I site is underlined and the *Spe*I site is shown with a dotted underline) and Pri-miR31 Flank Right containing an *Nhe*I site (5' CGC TAG CGC ACA CCT ATG GAG AGA GGA CAC GAA GGA CTG GCA TGC AGG TGG CCA TGG CTG CTG TCA GAC AGG AAA GAT 3', the *Nhe*I site is underlined). For pri-miR31 targeted to HIV *gag* (pri-miR31 Gag), the first round extension was completed using Pri-miR31 Gag Forward (5' TAA CTT GGA ACT GGA GAG GAA TAT TGA CGC TCT CGC ACC CTT TGA ACT GGG AAA CTG GTG CG 3') and Pri-miR31 Gag Reverse (5' TGC TGT CAG ACA GGA AAG ATA TAT TTA CGG TCT CGC ACC AGT TTC CCA GTT CAA AGG GTG CGA 3'), and the second round PCR was completed using the Pri-miR31 Flank Left and Right primers. For pri-miR31 targeted to HIV LTR (pri-miR31 LTR), the first round extension was completed using Pri-miR31 LTR Forward (5' TAA CTT GGA ACT GGA GAG GAT CTG AGG GAT CTC TAG TTA CTT TGA ACT GGG AAA CTT AAC TA 3') and Pri-miR31 LTR Reverse (5' TGC TGT CAG ACA GGA AAG ATT CTG ATA GAG CTC TAG TTA AGT TTC CCA GTT CAA AGT AAC TAG 3'), and the second round PCR was completed using the Pri-miR31 Flank Left and Right primers. For pri-miR31 targeted to HIV *tat* (pri-miR31 Tat), the first round extension was completed using Pri-miR31 Tat Forward (5' TAA CTT GGA ACT GGA GAG GAC TCT TCG TCG CTG TCT CCG CTT TGA ACT GGG AAA CTC GGA GA 3') and Pri-miR31 Tat Reverse (5' TGC TGT CAG ACA GGA AAG ATC TCT TTA TCT CTG TCT CCG AGT TTC CCA GTT CAA AGC GGA GAC 3'), and the second round PCR was completed using the Pri-miR31 Flank Left and Right primers.

For pri-miR155 targeted to HIV *int* (pri-miR155 IN), the first round extension was completed using Pri-miR155 IN Forward (5' TTG CTG TAG GCT GTA TGC TGA CTA GCC ATT GCT CTC CGG CTG TTT TTG CCT CCA ACT GAC AGC CGA G 3') and Pri-miR155 IN Reverse (5' GTA ACA GGC ATC ATA CAC TGA CTA ACC ATT GTC TCG GCT GTC AGT TGG AGG CAA AAA CAG CCG GA 3'), and the second round PCR was completed using Pri-miR155

Flank Left containing *Xho*I and *Spe*I sites (5' GAT CCT CGA GAC TAG TAA GGT GAG TCT TTA TGC CTC ATC CTC TGA GTG CTG AAG GCT TGC TGT AGG CTG TAT GCT G 3', the *Xho*I site is underlined and the *Spe*I site is shown with a dotted underline) and Pri-miR155 Flank Right containing an *Nhe*I site (5' CGC TAG CGA GGC CTC TGG AGC CGC GCG CCA CGG CAG CAA TTT GTT CCA TGT GAA TGC TAG TAA CAG GCA TCA TAC ACT G 3', the *Nhe*I site is underlined). For pri-miR155 targeted to HIV *gag* (pri-miR155 Gag), the first round extension was completed using Pri-miR155 Gag Forward (5' TTG CTG TAG GCT GTA TGC TGA TAT TGA CGC TCT CGC ACC CTT TTT TTG CCT CCA ACT GAA AGG GTC G 3') and Pri-miR155 Gag Reverse (5' GTA ACA GGC ATC ATA CAC TGA TAT TGA CGC TTC GAC CCT TTC AGT TGG AGG CAA AAA AAG GGT GC 3'), and the second round PCR was completed using the Pri-miR155 Flank Left and Right primers. For pri-miR155 targeted to HIV LTR (pri-miR155 LTR), the first round extension was completed using Pri-miR155 LTR Forward (5' TTG CTG TAG GCT GTA TGC TGT CTG AGG GAT CTC TAG TTA CTG TTT TTG CCT CCA ACT GAC AGT AAT A 3') and Pri-miR155 LTR Reverse (5' GTA ACA GGC ATC ATA CAC TGT CTG AGA GAT CCT ATT ACT GTC AGT TGG AGG CAA AAA CAG TAA CT 3'), and the second round PCR was completed using the Pri-miR155 Flank Left and Right primers. For pri-miR155 targeted to HIV *tat* (pri-miR155 Tat), the first round extension was completed using Pri-miR155 Tat Forward (5' TTG CTG TAG GCT GTA TGC TGC TCT TCG TCG CTG TCT CCG CTG TTT TTG CCT CCA ACT GAC AGC GGG A 3') and Pri-miR155 Tat Reverse (5' GTA ACA GGC ATC ATA CAC TGC TCT TCA TCG CGT CCC GCT GTC AGT TGG AGG CAA AAA CAG CGG AG 3'), and the second round PCR was completed using the Pri-miR155 Flank Left and Right primers.

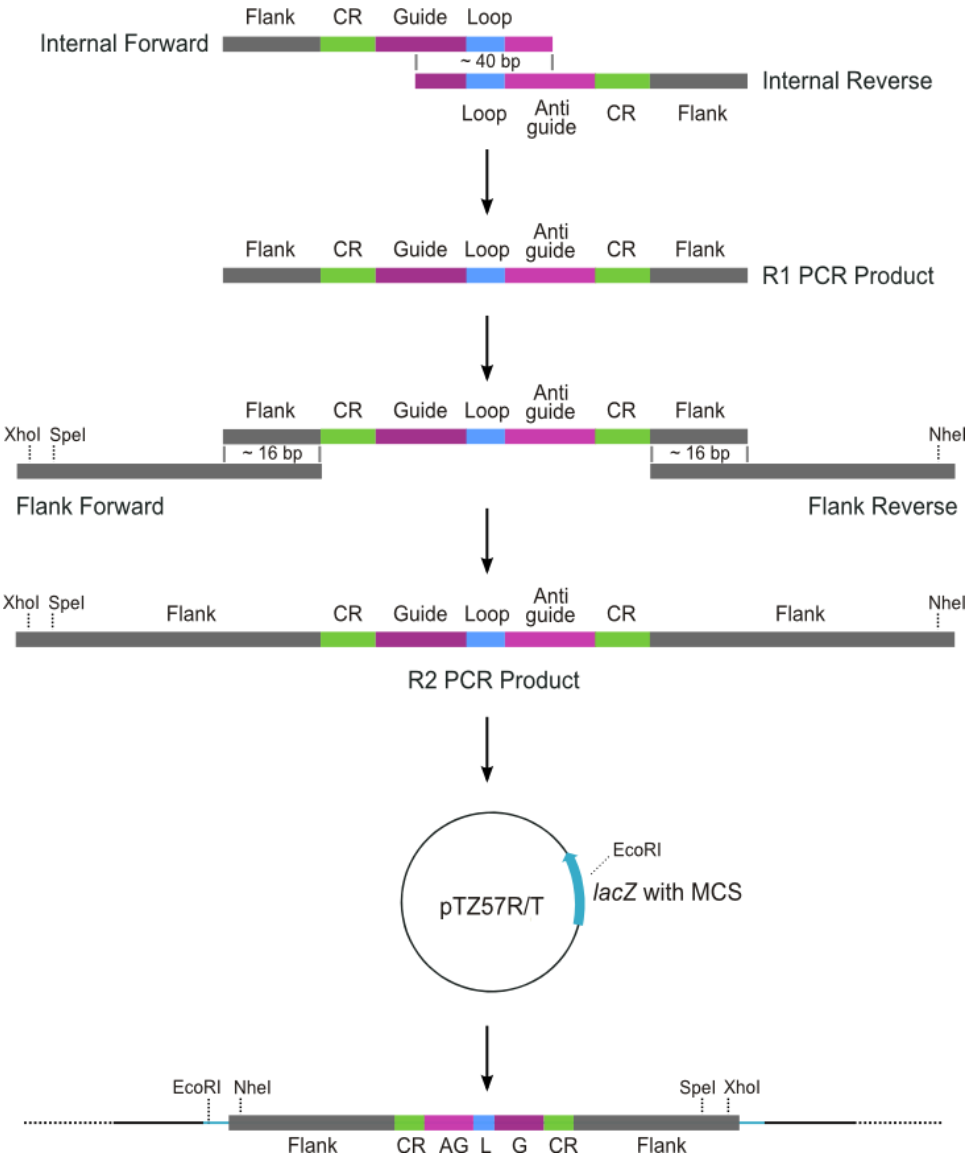


Figure 23. Two-step cloning strategy used to generate anti-HIV pri-miRNA mimics. Round 1 extension used internal forward and reverse primers to generate a Round 1 (R1) product comprising a ~16 bp region of each ss pri-miRNA flank (grey boxes), two regions common to all pri-miRNA mimics (CR, green boxes), the anti-HIV guide (G, purple box), loop (L, blue box) and anti-guide (AG, pink box) sequences. This product was diluted and used as template in the second PCR reaction with flank forward and reverse primers to generate a Round 2 (R2) PCR product that included *Xho*I, *Spe*I and *Nhe*I sites within the flanking regions. The entire pri-miRNA mimic was ligated into the TA vector pTZ57R/T and candidate clones were screened using *EcoR*I and *Xho*I digestion.

Positive orientation pri-miR122 IN clones were digested using *Xho*I and *Xba*I, electrophoresed and agarose-gel purified digestion product was directionally ligated into the similarly restricted pCi-Neo vector [(Promega, WI, USA) and Figure 24) that had been dephosphorylated as follows: 3 µg of pCi-Neo was incubated for 60 minutes at 37 ^{o}C with 1 x

Antarctic Phosphatase Reaction Buffer [50 mM bis-tris-propane-HCl, 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$ pH 6.0, New England Biolabs® Inc, MA, USA] and 5 U of Antarctic Phosphatase (New England Biolabs® Inc, MA, USA) in a final reaction volume of 10 µL followed by inactivation of the enzyme for 5 minutes at 65 ^{o}C. Dephosphorylated pCi-Neo and the digested PCR product were ligated using a 3:1 molar ratio as described for the psiCheck2-based tiling array vectors (see section 2.2.3). The same cloning strategy was used to generate pCi-pri-miR122 Gag, pCi-pri-miR122 LTR, pCi-pri-miR122 Tat, pCi-pri-miR31 IN, pCi-pri-miR31 Gag, pCi-pri-miR31 LTR, pCi-pri-miR31 Tat, pCi-pri-miR155 IN, pCi-pri-miR155 Gag, pCi-pri-miR155 LTR and pCi-pri-miR155 Tat. An *Xba*I site was present within the 5' flanking sequence of pri-miR30a, thus positive orientation pri-miR30a IN clones were digested using *Xho*I and *EcoR*I, electrophoresed and the agarose-gel purified digestion product was directionally ligated into the similarly restricted pCi-Neo vector (Promega, WI, USA) to generate pCi-pri-miR30a IN as described above. The same cloning strategy was used to generate pCi-pri-miR30a Gag, pCi-pri-miR30a LTR, pCi-pri-miR30a Tat and candidate pCi-Neo based clones were screened by digestion using *Spe*I and *Not*I and sequencing.

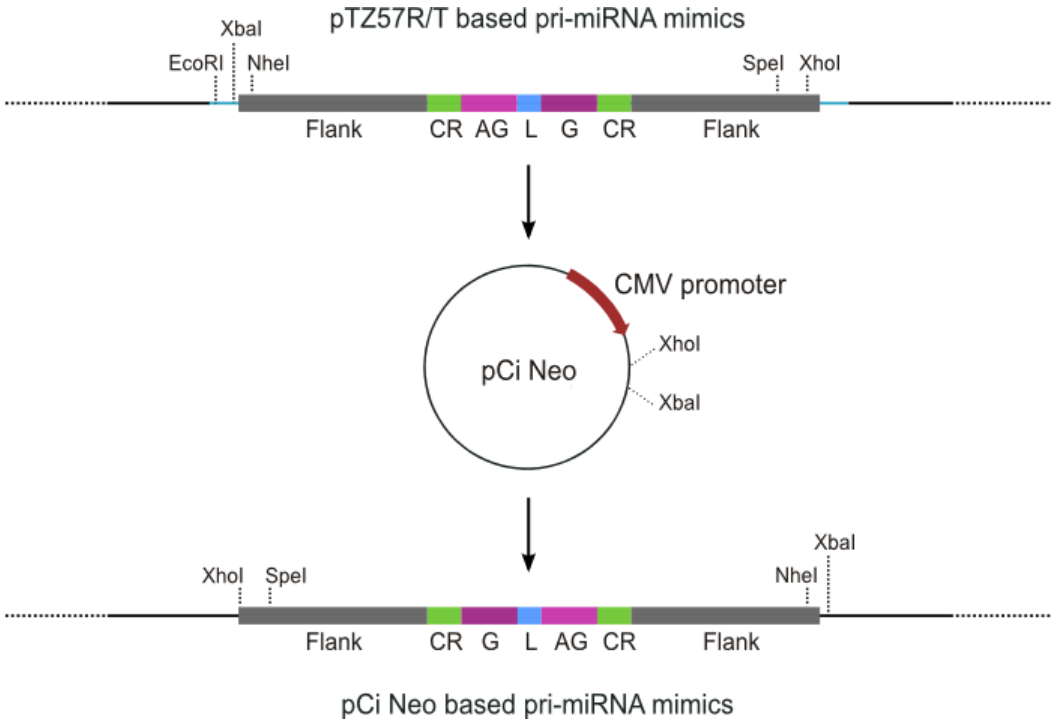


Figure 24. Cloning steps used to generate RNA Pol II-driven anti-HIV pri-miRNA mimics. Following the 2nd round of PCR amplification, anti-HIV pri-miRNA mimics were cloned into the TA vector pTZ57R/T. Positive orientation clones were digested using XbaI and XhoI and directionally ligated into the similarly digested pCiNeo vector containing a CMV promoter (red arrow). Common region (CR), G (guide), L (loop), AG (anti-guide).

3.2.2 Cloning of HIV target reporter vectors

To assess the inhibitory efficacies of the anti-HIV pri-miRNA mimics, a set of psiCheck2 based target vectors were generated as described in section 2.2.3 although the oligonucleotides used here were digested with *Xho*I and *Spe*I as opposed to *Xho*I and *Not*I. Candidate clones were screened by digestion using *Eco*RV and sequencing as described for the psiCheck2-based tiling array vectors (see section 2.2.1). To generate the psi-IN (HIV *int*) target vector, IN Forward containing a *Xho*I and *Eco*RV sites (5' TCG AGA TAT CAG CCG GAG AGC AAT GGC TAG TA 3', the *Xho*I site is underlined and the *Eco*RV site is shown with a dotted underline) and IN Reverse containing a *Spe*I site (5' CTA GTA CTA GCC ATT GCT CTC CGG CTG ATA TC 3', the *Spe*I site is underlined) were used. To generate the psi-Gag (HIV *gag*) target vector, Gag Forward containing a *Xho*I and *Eco*RV sites (5' TCG AGA TAT CGG TGC GAG AGC GTC AAT ATT AA 3', the *Xho*I site is underlined and the *Eco*RV site is shown with a dotted underline) and Gag Reverse containing a *Spe*I site (5' CTA GTT AAT ATT GAC GCT CTC GCA CCG ATA TC 3', the *Spe*I site is underlined) were used. To generate the psi-LTR (HIV LTR) target vector, LTR Forward containing a *Xho*I and *Eco*RV sites (5' TCG AGA TAT CAG TAA CTA GAG ATC CCT CAG CAC 3', the *Xho*I site is underlined and the *Eco*RV site is shown with a dotted underline) and LTR Reverse containing a *Spe*I site (5' CTA GTG TCT GAG GGA TCT CTA GTT ACT GAT ATC 3', the *Spe*I site is underlined) were used. To generate the psi-Tat (HIV *tat*) target vector, Tat Forward containing a *Xho*I and *Eco*RV sites (5' TCG AGA TAT CAG CGG AGA CAG CGA CGA AGA GA 3', the *Xho*I site is underlined and the *Eco*RV site is shown with a dotted underline) and Tat Reverse containing a *Spe*I site (5' CTA GTC TCT TCG TCG CTG TCT CCG CTG ATA TC 3', the *Spe*I site is underlined) were used. To generate the psi-Complete targets vector that comprised targets sites for *int*, *gag*, LTR and *tat*, Complete Forward containing a *Xho*I and *Eco*RV sites (5' TCG AGA TAT CAG TAA CTA GAG ATC CCT CAG ACG AAG AGC CGG AGA GCA ATG GCT AGT AGG CAG CGG AGA CAG CGA CGA AGA TGA CGG GTG CGA GAG CGT CAA TAT TAA 3', the *Xho*I site is underlined and the *Eco*RV site is shown with a dotted underline) and Complete Reverse containing a *Spe*I site (5' CTA GTT AAT ATT GAC GCT CTC GCA CCC GTC ATC TTC GTC GCT GTC TCC GCT GCC TAC TAG CCA TTG CTC TCC GGC CTC TTC GTC TGA GGG ATC TCT AGT TAC TGA TAT C 3', the *Spe*I site is underlined) were used.

3.2.3 Cloning of multimeric polycistonic anti-HIV pri-miRNA mimics

The generation of the multimeric polycistronic anti-HIV pri-miRNA mimic vectors was based on a cloning strategy that had been previously described (Sun, D, Melegari*, et al.*, 2006), whereby

single pri-miRNA mimics were amplified by PCR using primers that contained *Xba*I (5' primer) and *Spe*I sites (3' primer). To multimerise two (or more) single pri-miRNA mimics, each construct was digested using *Xba*I and *Spe*I and ligated using the 3' *Spe*I cohesive end of the first mimic and the 5' *Xba*I cohesive end of the second thereby joining the two and destroying both sites. A similar strategy was used in the current body of work as follows: from the panel of 16 individual pCi-pri-miRNA mimics, four were selected and cloned as multimeric vectors. These included pCi-pri-miR31 IN, pCi-pri-miR155 Gag, pCi-pri-miR30a LTR and pCi-pri-miR122 Tat. To generate multimeric vectors such that each pri-miRNA mimic was cloned within each possible position in the combination (one to four), four sets of '2mers' were cloned (IN-Gag, Gag-LTR, LTR-Tat and Tat-IN) and these were then used to generate four sets of '4mers' (IN-Gag-LTR-Tat, Gag-LTR-Tat-IN, LTR-Tat-IN-Gag and Tat-IN-Gag-LTR).

To generate IN-Gag, pTZ-pri-miR31 IN was digested using *Sca*I and *Nhe*I while pTZ-pri-miR155 Gag was digested using *Sca*I and *Spe*I as described in section 2.2.1. The two gel-purified pri-miRNA fragments were ligated such that the *Nhe*I cohesive end of pri-miR31 IN recombined with the *Spe*I cohesive end of pri-miR155 Gag destroying both sites to generate pTZ-IN-Gag (Figure 25). Positive orientation clones were digested using *Xho*I and *EcoR*I, electrophoresed and the agarose-gel purified digestion product was directionally ligated using a 1:1 ratio into the similarly restricted pCi-Neo vector (Promega, WI, USA) to generate pCi-pri-miR IN Gag as described in sections 2.2.1 and 2.2.2. The same strategy was used to clone the remaining 2mers and all four 4mers.

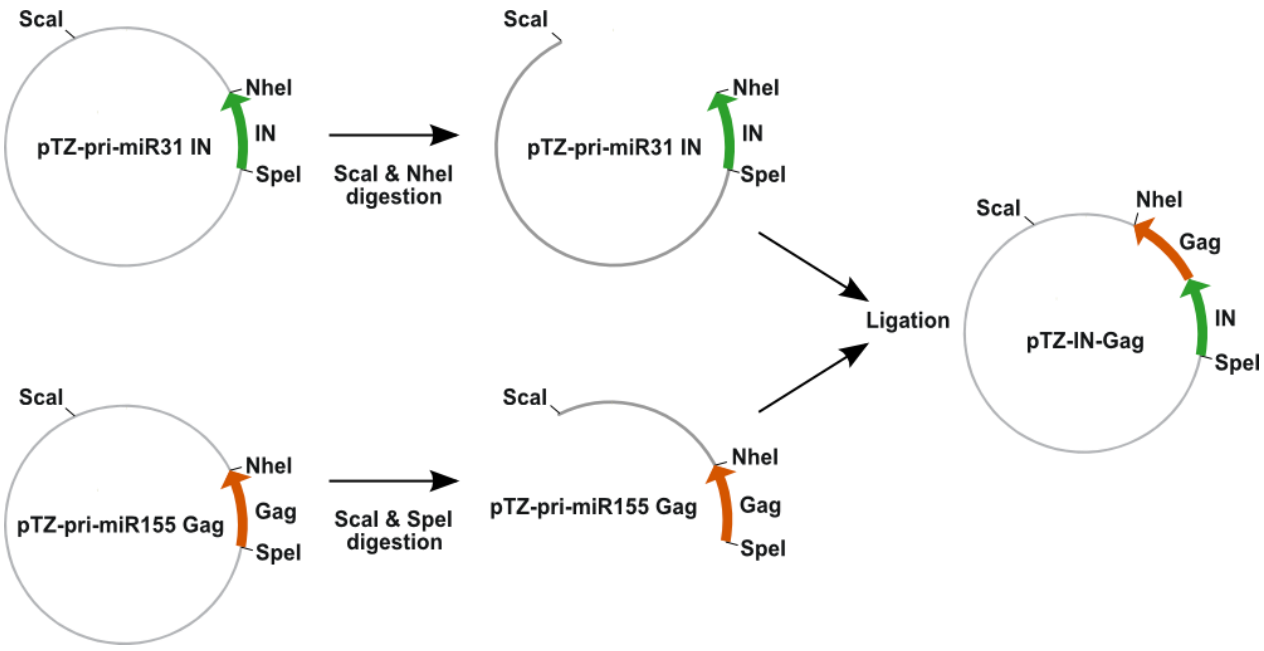


Figure 25. Cloning strategy used to generate multimeric polycistronic anti-HIV pri-miRNA constructs. The cloning of pTZ IN-Gag is depicted as an example of the steps used to generate all 2mer and 4mer multimeric constructs. Initially, pTZ-pri-miR31 IN (green arrow) was *Sca*I/*Nhe*I digested and pTZ-pri-miR155 Gag (orange arrow) was *Sca*I/*Spe*I digested. The two pri-miRNA fragments were ligated such that the *Nhe*I/*Spe*I sticky ends recombined but destroyed both sites to generate the 2mer pTZ-IN-Gag.

3.2.4 Transfection and detection of luciferase activity in cultured mammalian cells

To assess the inhibitory effects of the panel of sixteen individual anti-HIV pri-miRNA mimics against their cognate dual luciferase-based targets, HEK293 cells (maintained as described in section 2.2.4) were seeded at 120 000 cells per well in 24 well dishes and transfected as in triplicate as described in section 2.2.5 with the following changes: cells were co-transfected with a combination of 90 ng of psi-Complete or psi-IN, psi-Gag, psi-LTR, psi-Tat and 900 ng of either pCI-Neo mock (Promega, WI, USA) or one of the 16 individual pCi-based anti-HIV pri-miRNA mimics or one of the pCi-based anti-HIV pri-miRNA 2mer constructs or one of the pCi-based anti-HIV pri-miRNA 4mer constructs. In addition, 10 ng of pCI-eGFP (Passman, Weinberg*, et al.*, 2000) was added per well to monitor transfection efficiency by fluorescence microscopy. As a positive control to measure psiCheck2-based target inhibition, triplicate wells were co-transfected with a combination of 90 ng of psi-Complete or psi-IN, psi-Gag, psi-LTR, psi-Tat and 900 ng of either pTZU6+1 mock (Carmona, Ely*, et al.*, 2006), shIN, shGag, shLTR, shTat or shHBV [negative control, (Carmona, Ely*, et al.*, 2006)]. In addition 10 ng of pCi-eGFP was added per well. Dual luciferase activity was measured using the Promega Dual Luciferase Reporter Assay® Kit (Promega, WI, USA) as described in section 2.2.6.

3.2.5 Northern blot analysis to detect processed guide sequences from anti-HIV pri-miRNA mimics

Total RNA was extracted from HEK293 cells that had been seeded in 10 cm dishes and transfected with 20 µg of either one of the sixteen individual pCi-based anti-HIV pri-miRNA mimic constructs or one of the four pCi-based anti-HIV pri-miRNA 4mer constructs (see section 2.2.5). Northern Blot analysis was completed as previously described (see section 2.2.7) with the following changes: to detect guide sequences complementary to HIV *int*, *gag*, LTR and *tat* DNA oligonucleotide probes IN (5' AAG CCG GAG AGC AAT GGC TAG T 3'), Gag (5' AAG GGT GCG AGA GCG TCA ATA T 3'), LTR (5' AAG TAA CTA GAG ATC CCT CAG A 3') or Tat (5' GCG GAG ACA GCG ACG AAG AGC TT 3') were used respectively. Hybridised blots were exposed for 14 days to x-ray film at -80 ^{o}C.

3.2.6 Challenge assay using p24 ELISA to determine the inhibitory efficacies of multimeric polycistronic anti-HIV pri-miRNA mimics

To investigate the inhibitory efficacies of the multimeric polycistronic anti-HIV pri-miRNA mimics against a wild type strain of HIV-1, a viral challenge assay was conducted in U87.CD4.CCR5 cells using the South African R5-tropic subtype C HIV isolate FV5 as described in section 2.2.9 with the following changes: U87.CD4.CCR5 cells were seeded at 20 000 cells per well in a 96 well dish and transfected in triplicate using 0.25 µg of either pCi-pri-miR31 IN, pCi-pri-miR155 Gag, pCi-pri-miR30a LTR, pCi-pri-miR122 Tat or one of the four pCi-based anti-HIV pri-miRNA 4mer constructs or pCi-Neo mock.

3.2.7 Computational and statistical analysis

Computational analysis

Sequences of endogenous pri-miRNA mimics used in this chapter were downloaded from the Sanger miRNA Registry database (http://microrna.sanger.ac.uk/sequences/) using the 'miRNA name search' function with the prefix 'hsa' to denote *Homo sapien*. For example, a search for the pri-miR30a sequence was completed using 'hsa-miR30a' as a search keyword.

Predicted secondary structures of anti-HIV pri-miRNA mimics were analysed using the RNA folding program Mfold [http://mfold.burnet.edu.au/ and (Mathews, Sabina*, et al.*, 1999; Zuker, 2003)]. Sequences were uploaded in FASTA format as linear RNA with a folding temperature of 37 ^{o}C, 1 M NaCl ionic conditions with no divalent ions, 5 % suboptimality, an upper bound of 50 computed foldings, a maximum interior bulge size of 30 bp, a maximum

asymmetry of an interior bulge of 30 bp and no limit to the maximum distance between paired bases.

Statistical analysis

Data are expressed as the mean ± the standard error of the mean (SEM) and where appropriate are normalised to a control sample within the experiment. Statistical difference was considered to be significant when $p < 0.05$ and highly significant when $p < 0.01$. In all cases $\alpha = 0.99$ and data was analysed using a one-way ANOVA followed by a Dunnett's Multiple Comparison post-test calculated using GraphPad Prism (GraphPad Software Inc, USA).

3.3 Results

3.3.1 Analysis of the predicted secondary structures of anti-HIV pri-miRNA mimics

The sequences and predicted secondary structures of four endogenous pri-miRNAs were used as scaffolds for the introduction of exogenous anti-HIV guide sequences with the aim of developing a combinatorial construct capable of simultaneously inhibiting multiple HIV mRNAs. The pri-miRNAs included pri-miR122, pri-miR30a, pri-miR31 and pri-miR155 while the anti-HIV guide sequences included those targeted against *int*, *gag*, LTR and *tat*. Endogenous pri-miR122 is hepatocyte-specific and involved in hepatic circadian rhythms (Gatfield, Le Martelot*, et al.*, 2009) while pri-miR31 is expressed in mesenchymal cells and controls cellular proliferation and adipogenesis (Sun, F, Wang*, et al.*, 2009). Pri-miR155 is vital during haemostasis for the regulation of immune cell development and function (O'Connell, Chaudhuri*, et al.*, 2009). To date, no specific function had been ascribed to pri-miR30a but this endogenous regulator has been found to occur ubiquitously in vertebrate cells (Lagos-Quintana, Rauhut*, et al.*, 2001). Notably, as these pri-miRNAs function in a variety of different cells, it is unlikely that they require cell-specific post-transcriptional modifications to function. The sequence of each endogenous pri-miRNA as well as each of the sixteen anti-HIV pri-miRNA mimics was subjected to computational analysis using Mfold to generate their predicted secondary structures. In addition, as a measure of the thermodynamic stability of each predicted structure, the Gibbs free energy value (δG measured in kcal/mol) was recorded (Figure 26).

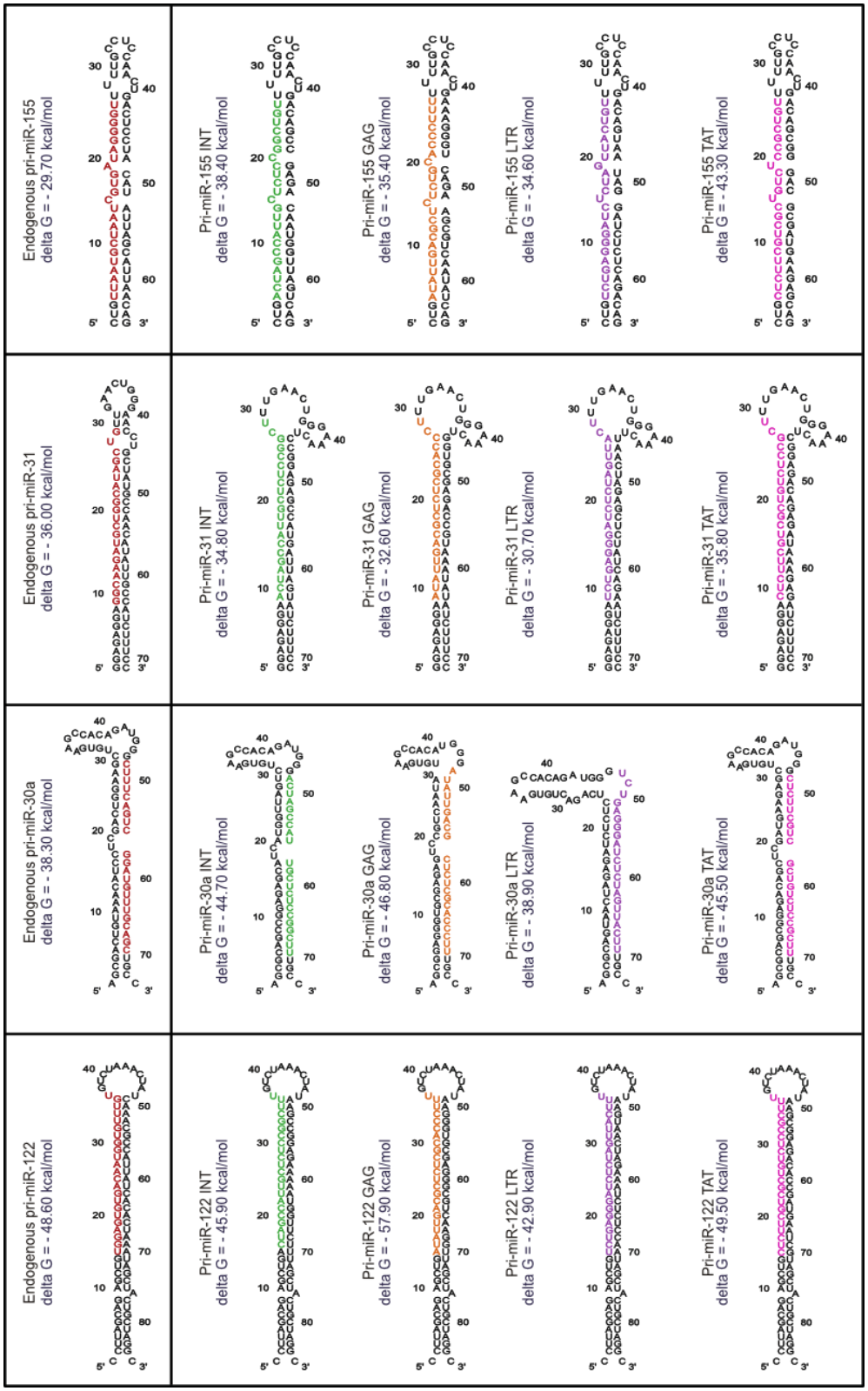


Figure 26. Predicted secondary structures and Gibbs free energy values (δG) for pri-miRNA mimics. The predicted secondary structure of each endogenous pri-miRNA, excluding the ss flanking regions, is depicted at the top of each column with guide sequences shown in red. Each endogenous guide sequence has been substituted with an anti-HIV guide (IN in green, Gag in orange, LTR in purple and Tat in pink) and the predicted secondary structures and δG values for each pri-miRNA mimic is displayed in columns based on the pri-miRNA backbone used.

The endogenous pri-miR122 backbone was predicted to form a stem-loop structure comprised of a 35 bp duplex with one unmatched base (position nine) but no bulges and a 13 bp loop. The endogenous guide sequence was predicted to span bases 15 to 37 from the 5' end of the duplex (Figure 26 top panel, far left column). Substitution of the endogenous guide strand with any of the four anti-HIV guides *int*, *gag*, LTR or *tat*, did not result in any changes to the predicted secondary structure of the pri-miR122 backbone. The duplex stem regions and loops of all four anti-HIV pri-miR122-based mimics closely resembled that of the endogenous backbone and the anti-HIV guide sequences were predicted to span the same bases as the endogenous guide. The δG values for two of the pri-miRNA122-based mimics (pri-miRNAs 122 IN and LTR) were slightly higher than that of the endogenous backbone (-45.90 kcal/mol and -42.90 kcal/mol respectively vs. -48.60 kcal/mol) suggesting that the predicted secondary structures of these two mimics were energetically less favourable than the endogenous scaffold.

A similar set of results was obtained for pri-miR155-based mimics. The endogenous pri-miR155 backbone was predicted to form a 30 bp duplex structure with two unmatched bases (positions 15 and 19), one bulge (positions 27 to 29) and a six bp loop. The endogenous guide sequence was predicted to span bases 4 to 26 from the 5' end of the duplex (Figure 26 top panel, second column from the left). Substitution of the endogenous guide strand with any of the four anti-HIV guides *int*, *gag*, LTR or *tat*, resulted in minor changes to the predicted secondary structure of the pri-miR155 backbone. The duplex stem regions and loops of all four anti-HIV pri-miR155-based mimics closely resembled that of the endogenous backbone and the anti-HIV guide sequences were predicted to span the same bases as the endogenous guide. However, the predicted structures of pri-miRNAs 155 IN and Gag showed a second unmatched base at position 20 as opposed to 19. Overall, the δG values for all four anti-HIV mimics were lower than that of the endogenous backbone suggesting that the predicted secondary structures of the mimics were energetically more favourable than the endogenous scaffold.

The endogenous pri-miR30a backbone was predicted to form a 30 bp duplex structure with two unmatched bases (positions 18 and 19) and an unusual loop-stem-loop structure (positions 31 to 47). Unlike the other three pri-miRNAs, the endogenous guide sequence of pri-miR30a was predicted to span bases 48 to 69 from the 5' end of the duplex such that the guide was produced from the 3' arm of the structure (Figure 26 top panel, third column from the left). Substitution of the endogenous guide strand with anti-HIV guides *int* or *tat*, did not result in any changes to the predicted secondary structure of the pri-miR30a backbone and the δG values for these two mimics were lower than that of the endogenous scaffold (-44.70 kcal/mol and -45.50 kcal/mol respectively vs. -38.30 kcal/mol). Substitution of the endogenous guide strand with the anti-HIV *gag* guide resulted in an unmatched base at position 30 that shifted the guide sequence into the loop. Interestingly, the δG value for this mimic was the lowest for all pri-miR30a-based

mimics (-46.80 kcal/mol) suggesting it was the most thermodynamically stable secondary structure for this set of scaffolds. Substitution of the endogenous guide strand with the anti-HIV LTR guide resulted in a δG value almost exactly the same as the endogenous scaffold (-38.90 kcal/mol vs. -38.30 kcal/mol respectively) but the overall predicted shape of pri-miR30a LTR resembled the predicted endogenous structure the least. The duplex region of pri-miR30a LTR contained no unmatched bases and the stem-loop-stem structure spanned bases 24 to 50 with an additional bulge present at bases 27 and 28. Furthermore, while the LTR guide was predicted to span the same bases as the endogenous sequence (positions 48 to 69) the rearranged stem-loop-stem structure caused the LTR guide to be present within the loop region.

The endogenous pri-miR31 backbone was predicted to form a 31 bp duplex structure with one bulge at position 28 and a nine bp loop (positions 31 to 40). The endogenous guide sequence was predicted to span bases 9 to 29 from the 5' end of the duplex (Figure 26 top panel, far right column). Substitution of the endogenous guide strand with any of the anti-HIV *int*, *gag*, LTR or *tat*, resulted in significant changes to the predicted secondary structure of the pri-miR31 backbone such that the mimics closely resembled each other in shape but not the endogenous scaffold. All four pri-miR31-based mimics were predicted to form a 27 bp duplex with no bulges and a loop-loop structure at positions 28 to 45. In addition, while the anti-HIV guides were predicted to span the same bases as the endogenous sequence (positions 9 to 29) the rearranged loop-loop structure caused the guides to be present within the loop region. Overall, the δG values for all four pri-miR31-based mimics were higher than that of the endogenous backbone suggesting that the predicted secondary structures of the mimics were energetically less favourable than the endogenous scaffold.

Importantly, as programs such as Mfold provide predicted secondary structures for RNA sequences, no functional data could be extrapolated. The data described in Figure 26 suggested that substitution of the endogenous guide sequences of four discrete pri-miRNA backbones with anti-HIV guides resulted in mimics whose predicted structures were generally a close fit to their endogenous counterparts. However, to assess the functional consequences of substituting the endogenous guide sequence of the various pri-miRNA backbones and the abilities of the different scaffolds to generate efficacious anti-HIV siRNAs, a luciferase-based assay was completed.

3.3.2 Inhibitory efficacies of individual anti-HIV pri-miRNA mimics targeted to a luciferase reporter to assess the ability of different pri-miRNA backbones to generate functional guide sequences

To determine the ability of each pri-miRNA scaffold to generate functional anti-HIV guide sequences, each of the anti-HIV pri-miRNA mimics was cloned into a mammalian expression vector under the control of a CMV promoter. The transcribed cassettes folded into pri-miRNAs with ds duplex regions comprising 5' and 3' arms connected by a loop and two ss flanking regions (Figure 27). To investigate the inhibitory efficacies of anti-HIV guide sequences when expressed from discrete pri-miRNA backbones, a dual luciferase-based target vector that encoded a complementary target sequence for each anti-HIV guide (*int*, *gag*, LTR and *tat*) was co-transfected with each of the individual modified pCi-pri-miRNA mimics. In addition, each of the shRNAs (IN, Gag, LTR and Tat), a non-specific shRNA targeted to HBV, as well as an empty vector were included as controls.

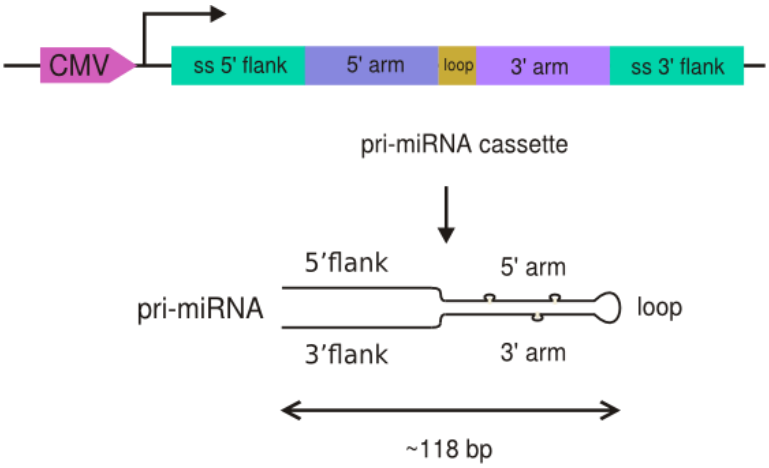


Figure 27. Design of expressed anti-HIV pri-miRNA cassettes. Pol II (CMV) cassettes were generated using a two step PCR approach such that transcribed pri-miRNA mimics folded to form a ds duplex structure comprising 5' and 3' arms connected by a loop and two ss flanking regions. The entire folded pri-miRNA was ~118 bp.

Luciferase data analysed 48 hours post-transfection revealed an interesting pattern of inhibition (Figure 28). Overall, all 16 individual anti-HIV pri-miRNA mimics showed highly significant inhibition of the target vector as compared to the mock ($\alpha = 0.99$; $p < 0.01$) although the total percentage suppression varied between the different pri-miRNA backbones. Interestingly, three

of the four sets of mimics (pri-miRNAs 122, 30a and 155) showed similar levels of target inhibition regardless of the guide sequence present. Pri-miR122-based mimics consistently showed the least overall target inhibition with an average suppression level of 45.00 % ± 2.18 %. In contrast, pri-miR30a yielded highly potent anti-HIV guide sequences from all 4 scaffolds with inhibitory levels comparable to the shRNA controls that had been chosen based on their potent suppressive efficacies. From the entire panel of 16 mimics, pCi-pri-miR30a LTR mediated the greatest overall inhibition of psi-Complete with 94.00 % ± 0.01 % suppression. Similarly to the pri-miR30a-based scaffolds, potent inhibition of psi-Complete was mediated by all 4 anti-HIV pri-miR155-based scaffolds with pCi-pri-miR155 IN and Gag suppressing their targets to the same degree as the corresponding pri-miR30a mimics. pCi-pri-miR155 LTR mediated the poorest suppression for this set but the level was highly significant at 73.33 % ± 1.16 %. In contrast, target inhibition mediated by the pri-miR31 scaffolds varied depending on the guide sequence present with pCi-pri-miR31 IN showing the most (64.33 % ± 0.01 %) and pCi-pri-miR31 Tat showing the least (16.67 % ± 3.51 %) suppression for this set of backbones. This latter construct mediated the worst target suppression out of the entire panel of 16 mimics.

Intriguingly, the suppression data seemed to correlate with the predicted secondary structures for three out of the four sets of mimics (refer to Figure 26). The anti-HIV pri-miR122 and 155-based mimics showed the ‘closest fit’ overall to their relative endogenous scaffolds and both sets of mimics generated comparable target inhibition levels regardless of the guide sequence present. Similarly, pri-miR30-based mimics mediated highly significant inhibition with all four guide sequences and were predicted to fold into secondary conformations that closely resembled the endogenous scaffold with the exception of pri-miR30a LTR. The predicted shape of this mimic was the ‘least likely fit’ as compared to the endogenous backbone and yet this mimic mediated the greatest percentage of luciferase inhibition out of the entire panel of 16 mimics. Pri-miR31-based mimics showed a different trend for both their predicted secondary structures and inhibitory levels. None of the four anti-HIV pri-miR31-based mimics were predicted to fold into a structure that closely resembled the endogenous scaffold and the inhibition mediated by each of the mimics differed depending on the guide sequence present.

Taken together, these data suggested that the pattern of inhibition was the same for each scaffold regardless of the guide sequence present *if* the folding parameters were as close to the endogenous backbone as possible. Pri-miR31 was the exception and the data for this scaffold suggested that the guide sequence itself was the main predictor of inhibitory efficacy.

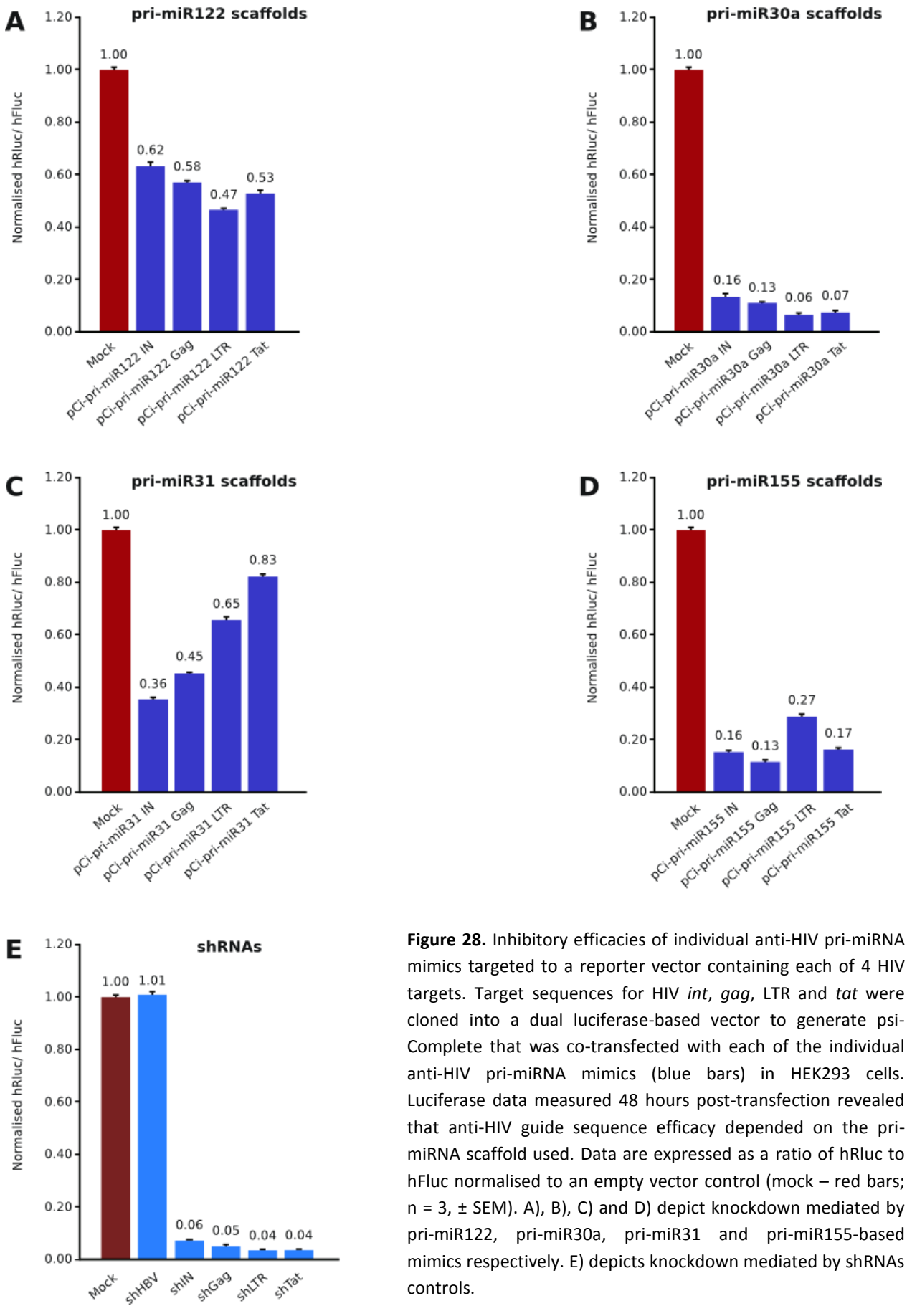


Figure 28. Inhibitory efficacies of individual anti-HIV pri-miRNA mimics targeted to a reporter vector containing each of 4 HIV targets. Target sequences for HIV *int*, *gag*, LTR and *tat* were cloned into a dual luciferase-based vector to generate psi-Complete that was co-transfected with each of the individual anti-HIV pri-miRNA mimics (blue bars) in HEK293 cells. Luciferase data measured 48 hours post-transfection revealed that anti-HIV guide sequence efficacy depended on the pri-miRNA scaffold used. Data are expressed as a ratio of hRluc to hFluc normalised to an empty vector control (mock – red bars; n = 3, ± SEM). A), B), C) and D) depict knockdown mediated by pri-miR122, pri-miR30a, pri-miR31 and pri-miR155-based mimics respectively. E) depicts knockdown mediated by shRNAs controls.

3.3.3 The design and construction of multimeric polycistronic anti-HIV pri-miRNA mimics

To develop a combinatorial RNAi system based on anti-HIV pri-miRNAs, the relationship between pri-miRNA scaffold and guide sequence efficacy was investigated. The data described in section 3.3.2 revealed that the same anti-HIV guide sequence could mediate varying levels of target inhibition depending on the pri-miRNA scaffold it was processed from. To avoid possible recombination events that could excise one or more of the pri-miRNA members within a combinatorial system, a representative scaffold from each of the four backbones was included. However, as there was no obvious trend for which guide sequence/pri-miRNA scaffold 'pair' induced the greatest target suppression, the following mimics were chosen to generate a Pol II-expressed multimeric system: pri-miR122 Tat, pri-miR30a LTR, pri-miR31 IN and pri-miR155 Gag.

To investigate if the position of an individual anti-HIV pri-miRNA mimic within the multimeric system affected the guide sequence efficacy, each individual pri-miRNA mimic was cloned such that it occupied either the first, second, third or fourth position within the multimer relative to the Pol II (CMV) promoter. To achieve this, each of the individual anti-HIV pri-miRNA mimics were cloned in tandem to generate a set of 2mers with each mimic cloned in both the first and second position within each vector (Figure 29A). The 2mers were then directionally cloned in tandem to generate a set of 4mers with each 2mer pair cloned in both the first and second position within each vector. This generated a total of four 2mers and four 4mers each expressed off a single CMV promoter (Figure 29B).

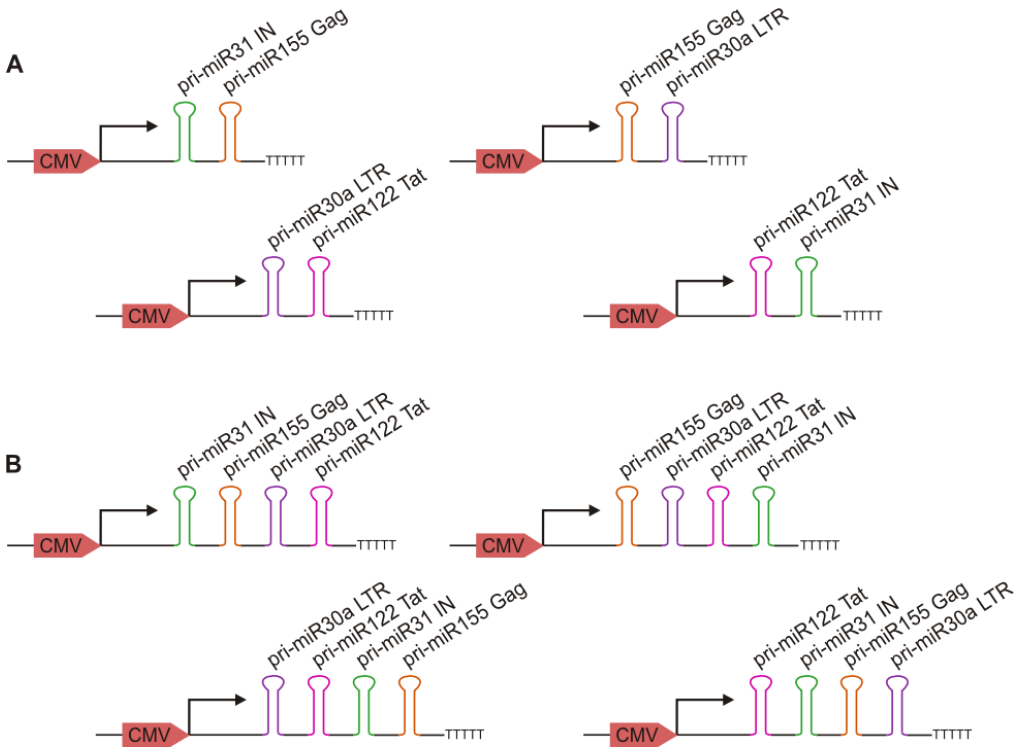


Figure 29. Generation of a panel of multimeric anti-HIV pri-miRNA mimics using pri-miR122 Tat, pri-miR30a LTR, pri-miR31 IN and pri-miR155 Gag. A) Single pri-miRNA mimics were directionally cloned in tandem to generate a set of 2mers with each mimic cloned in the 1st and 2nd position relative to the CMV promoter within each vector. B) The 2mers were directionally cloned in tandem to generate a set of 4mers with each 2mer pair cloned in the 1st and 2nd position relative to the CMV promoter within each vector. This generated a set of four 2mers and four 4mers each expressed off a single CMV promoter (red arrows).

3.3.4 Inhibitory efficacies of multimeric polycistronic anti-HIV pri-miRNA mimics targeted to individual luciferase-based HIV target reporter vectors

To assess target knockdown induced by each of the multimeric polycistronic anti-HIV pri-miRNA mimics, a set of dual luciferase target vectors was generated. The sequence for each of the HIV targets (*int*, *gag*, LTR and *tat*) was individually cloned within the 3' UTR of hRluc of psiCheck2 to generate psi-IN, psi-Gag, psi-LTR and psi-Tat. Each target vector was co-transfected with one of the multimeric pri-miRNA constructs or an empty vector control (mock) and luciferase data was collected 48 hours post-transfection.

When co-transfected with psi-IN, psi-Gag or psi-Tat, all four pri-miRNA mimic constructs mediated highly significant inhibition of each target vector ($\alpha = 0.99$; $p < 0.01$) with an average percentage inhibition of 91.33 % ± 0.03 % which was comparable to that of the shRNA controls (95.25 % ± 0.01 % and Figure 30). Furthermore, there was no significant difference between

each of the 4mers regardless of the target construct ($p > 0.05$) suggesting that the position of the individual pri-miRNA mimic within the multimer did not affect guide sequence efficacy.

The 2mer constructs that contained target-specific guide sequences mediated highly significant suppression ($\alpha = 0.99$; $p < 0.01$) with the exception of pCi-LTR-Tat. Sequencing analysis of this construct revealed 10 mismatches within the LTR guide sequence of this mimic that had been introduced during the cloning process (see Appendix A3.3). These data reinforced the need for complementarity between the guide sequence and its cognate target. Notably this clone had not been used to generate the 4mer constructs and so the mismatches were limited to the 2mer. In contrast to the other HIV targets, inhibition of psi-Tat was poor regardless of the pri-miRNA mimic used. One of the 2mers (pCi-Tat-IN) and three of the 4mers mediated highly significant inhibition of psi-Tat ($\alpha = 0.99$; $p < 0.01$) although the greatest suppression was only 55.33 % ± 1.53 % (pCi-LTR-Tat-IN-Gag). Sequence analysis revealed 11 mismatches within the Tat guide sequence of pCi-LTR-Tat that had been introduced during the cloning process (see Appendix A3.3) but this clone had not been used to generate the 4mers and thus the reasons for the poor knockdown mediated by the latter mimics were unclear.

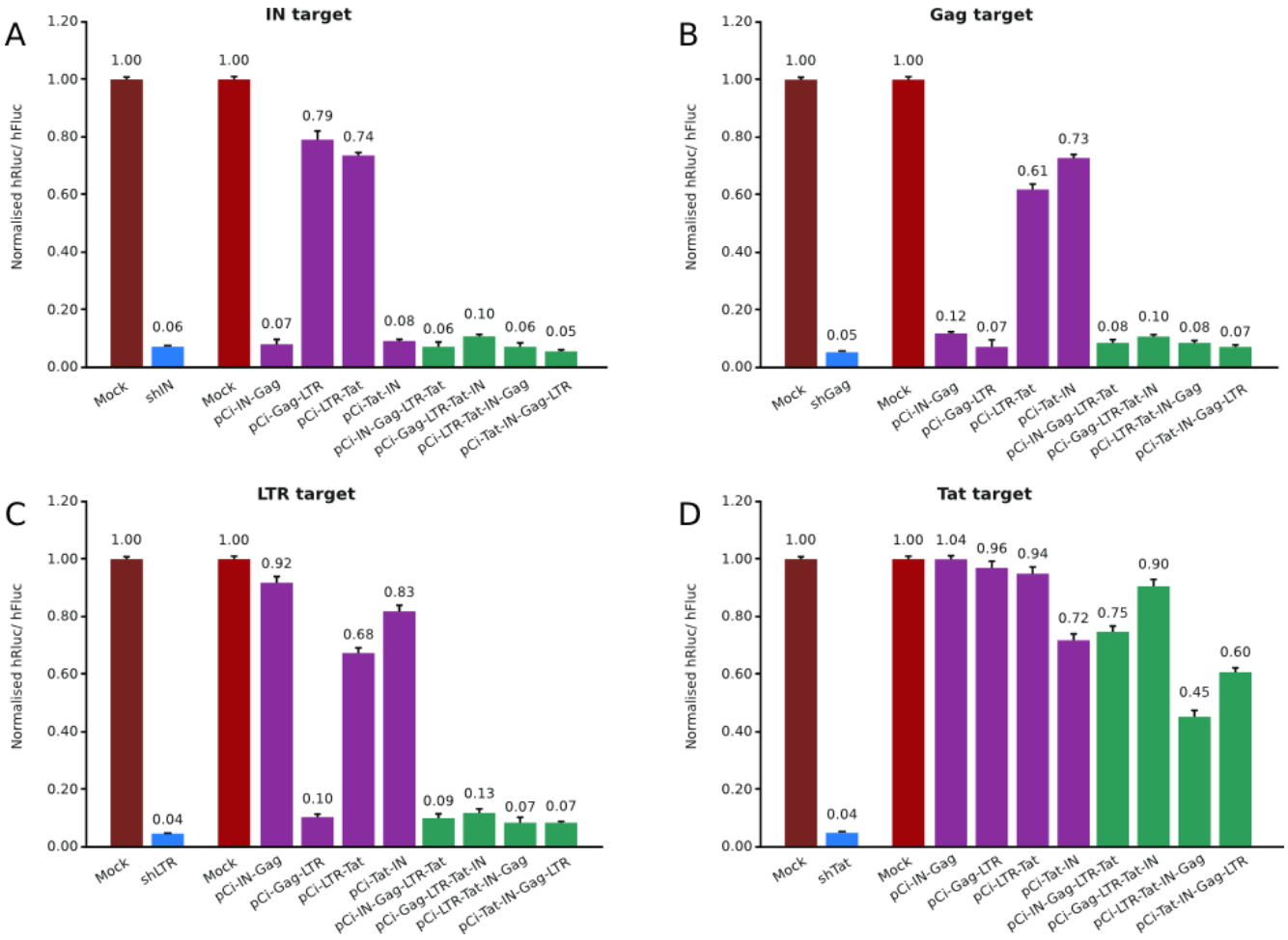


Figure 30. Multimeric anti-HIV pri-miRNA mimics targeted to individual HIV targets. Four HIV target sites, *int*, *gag*, LTR and *tat* were individually cloned within the 3' UTR of hRluc in the psiCheck2 dual luciferase vector to generate psi-IN, psi-Gag, psi-LTR and psi-Tat respectively. The target vectors were co-transfected into HEK293 cells with an empty vector control (mock – red bars), one of the 2mer (purple bars) or 4mer (green bars) multimeric anti-HIV pri-miRNA mimics. In addition, target-specific shRNAs (blue bars) were co-transfected along with a shRNA-specific mock (dark red bars). Data are expressed as a ratio of hRluc to hFluc normalised to the mock (n = 3, ± SEM). A), B), C) and D) depict knockdown of psi-IN, psi-Gag, psi-LTR and psi-Tat respectively.

3.3.5 Northern blot analysis of anti-HIV pri-miRNA mimics to determine guide sequence

The luciferase-based reporter assays described in sections 3.3.2 and 3.3.4 demonstrated that the inhibitory efficacies of four separate anti-HIV guide sequences depended on the pri-miRNA scaffold from which they were expressed. To determine if this variation was linked to guide sequence production, Northern blot analysis was completed on all single and polycistronic (4mer) pri-miRNA mimics. An empty vector control (mock) as well as a target-specific shRNA was included in each of four blots that were probed using radio-labelled oligonucleotides specific to the IN, Gag, LTR and Tat guide sequences.

In each blot, the shRNA controls (last lane of each blot) were detected indicating that the guide-specific probes were able to bind their target sequences and that the Northern blots had been successful (Figure 31). In addition, the mock-treated samples were not detected (first lane of each blot) suggesting that the probe sequences were specific to their complementary targets and not non-specific miRNAs or other small RNAs. When using the IN probe, a precursor and guide sequence was detected for the single pCi-pri-miR30a IN construct only (Figure 31A, 3rd lane). In addition, the guide sequence was very faint despite 14 days of exposure suggesting that a lower concentration of functional guide was produced as compared to the precursor. No other sequences were detected for any of the remaining IN-based single pri-miRNA mimics or the 4mer constructs. A similar pattern was observed for blots probed with LTR and Tat-specific oligonucleotides (Figure 31C and D, 3rd lanes). Precursor and guide sequences were detected for single pri-miR30a-based mimics only, with less guide sequence detected as compared to the precursors. In contrast, the Gag-specific probe detected processed precursor and guide sequences for pCi-pri-miR30a Gag as well as guide sequence only for pCi-pri-miR155 Gag (Figure 31B, 3rd and 5th lanes respectively). This latter construct showed a higher concentration of guide as compared to pCi-pri-miR30a Gag although no precursor was detected. Furthermore, precursor and guide sequences were observed for all four 4mers using the Gag probe.

Intriguingly, the Northern blot data seemed to correlate with the luciferase-based suppression data (refer back to Figures 28 and 30). All 4 individual pri-miR30a-based mimics showed the greatest luciferase reporter suppression regardless of the guide sequence present (Figure 28B) and both precursor and guide strands were detected for all 4 of these constructs using radio-labelled probes (3rd lane of all 4 Northern blots in Figure 31). Similarly, pCi-pri-miR155 Gag mediated the greatest percentage inhibition of psi-Complete for this set of mimics (Figure 28D) and a strong guide sequence signal was detected by Northern blot (5th lane of Figure 31B). No precursor or guide sequences were detected for any of the remaining individual anti-HIV pri-miRNA mimics and, with the exception of the pri-miR155-based mimic set, none of the others mediated highly suppressive target inhibition. This strongly suggests that the amount of guide produced by the pri-miR122- and pri-miR31-based mimics was below the limit of detection of the Northern blot assay. The reasons for the absence of detectable guide sequence from pCi-pri-miRNAs 155 IN, LTR and Tat was unclear as these three constructs had mediated highly significant target inhibition albeit at lower levels as compared to pCi-pri-miR155 Gag. This latter construct was a component of the 4mer constructs and interestingly, only the Gag-specific radio-labelled probe was able to detect guide sequences for any the 4mers (Figure 31B right hand panel). This observation did not correlate with the luciferase data (Figure 30) as all four 4mer constructs mediated highly significant inhibition of all four HIV target reporter vectors.

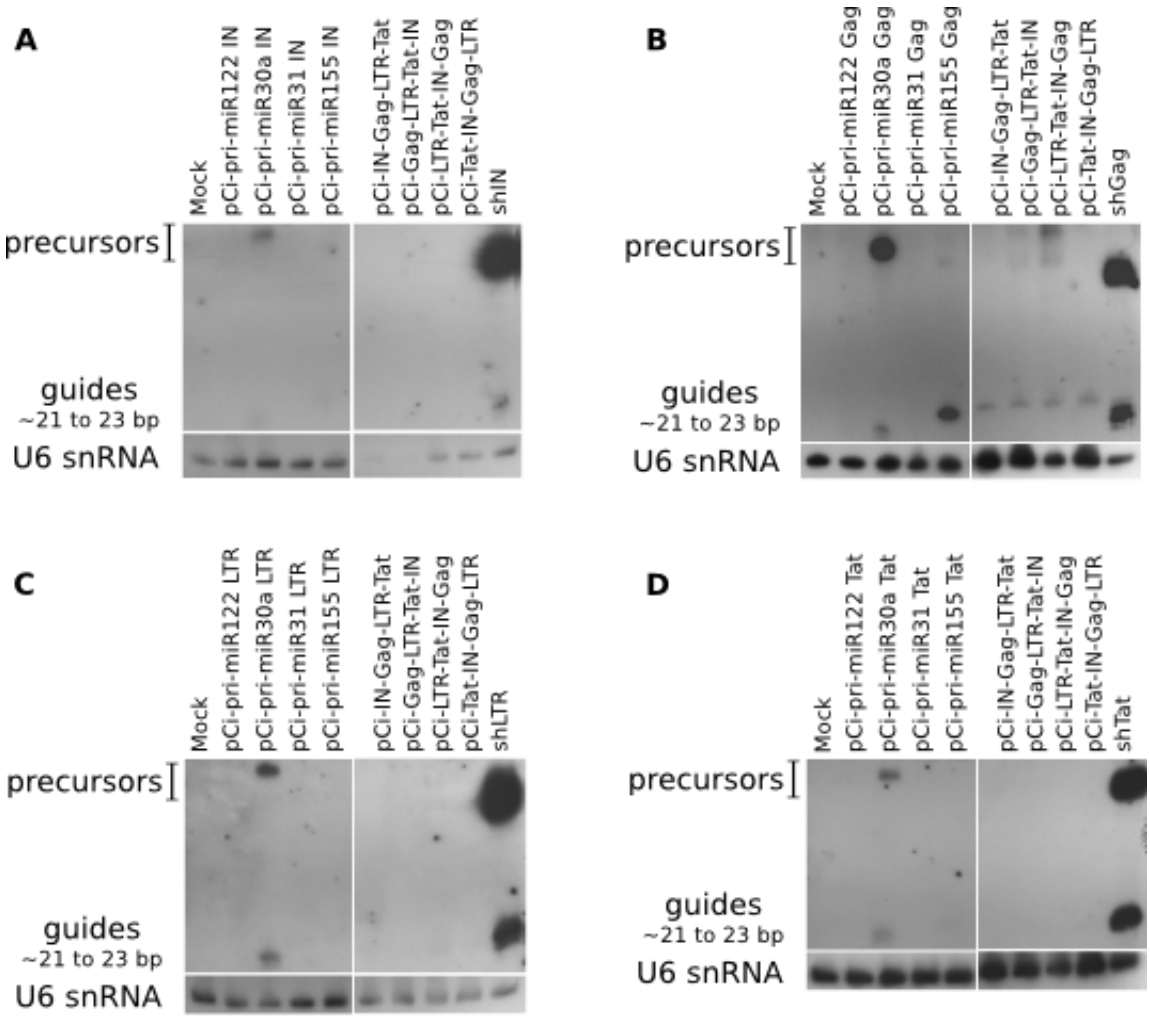


Figure 31. Northern blot analysis to detect putative guide sequences produced from expressed anti-HIV pri-miRNA mimics. Four ssDNA oligonucleotide probes targeted to HIV *int* (A), *gag* (B), LTR (C) and *tat* (D) as well as a U6 snRNA probe were hybridised to RNA extracted from HEK293 cells transfected with either single or combinatorial (4mer) pri-miRNA mimics.

3.3.6 Suppression of HIV replication using combinatorial anti-HIV pri-miRNA mimics

The main aim of testing numerous pri-miRNA scaffolds for the ability to produce highly efficacious exogenous guide sequences was to generate a combinatorial polycistronic pri-miRNA system with the capability of simultaneously and potently inhibiting four different HIV mRNAs. While all four polycistronic pri-miRNA constructs were highly effective suppressors of luciferase-based targets, an important and biologically relevant assay required the use of a fully replicative wild type HIV genome. Thus the HIV subtype C isolate FV5 provided a unique opportunity to evaluate the antiviral efficacies of all four combinatorial pri-miRNA mimics against a pool of target sequences akin to the conditions in a natural HIV infection. U87.CD4.R5

fibroblasts were transfected with either an empty vector control (mock) or one of the four polycistronic anti-HIV pri-miRNA mimics (4mers) before being challenged with FV5 using a $TCID_{50}$ of 1000/mL . For comparison, the single anti-HIV pri-miRNA mimics that had been used to generate the multimeric units were included and inhibition of HIV was assessed seven days post-infection by determining p24 antigen levels.

Cells that received pCi-pri-miR31 IN or pCi-pri-miR122 Tat showed significant differences when compared to mock-treated samples ($\alpha = 0.99$; $p < 0.05$) but the levels of suppression were poor at 29.37 % ± 1.44 % and 29.70 % ± 0.27 % respectively (Figure 32). Cells transfected with pCi-pri-miR155 Gag or pCi-pri-miR30a LTR showed highly significant inhibition ($\alpha = 0.99$; $p < 0.01$) with 98.03 % ± 0.11 % and 98.50 % ± 0.01 % knockdown respectively. As previously noted, these two mimics were the only single anti-HIV pri-miRNAs for which guide sequences had been detected via Northern blot analysis (see section 3.3.4). Three of the four 4mers mediated highly significant suppression of FV5 ($\alpha = 0.99$; $p < 0.01$) with only one 4mer, pCi-IN-Gag-LTR-Tat showing no target inhibition ($p > 0.05$). The average percentage inhibition for the efficacious 4mers was 91.91 % ± 1.39 % with pCi-Tat-LTR-IN-Gag mediating the greatest suppression at 98.37 % ± 0.01 %. Although the use of single mimics is not recommended long-term, these data provided very promising proof-of-principle results that both single and multimeric pri-miRNA mimics were able to inhibit multiple replication cycles of a subtype C HIV isolate up to seven days post-transfection.

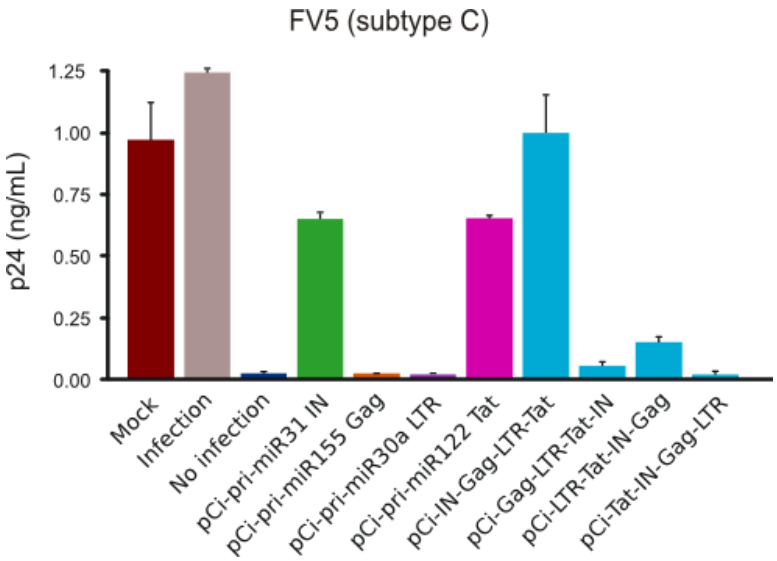


Figure 32. Challenge assay to assess combinatorial anti-HIV pri-miRNA mimics against a subtype C viral isolate of HIV. All 4 polycistronic pri-miRNA mimics (blue bars), the single anti-HIV pri-miRNA mimics used to generate the combinatorial constructs as well as an empty vector control (mock – red bar) were transfected into U87.CD4.R5 cells prior to challenge with viral isolate FV5 at a $TCID_{50}$ of 1000/mL. Viral Gag p24 antigen levels were determined 4 days post-infection by ELISA (450 nm). Values were converted to ng/mL using a standard curve supplied with the ELISA kit and data are expressed as p24 antigen levels in ng/mL (n = 3, ± SEM).

3.4 Discussion and conclusions

The generation of a combinatorial pri-miRNA mimic construct relies on the use of different backbones to ensure that the overall potency of the system is not lessened through recombination events that delete one or more of the scaffolds. However, the choice of which pri-miRNA backbones to include in such a system has been confined by the limited data available describing the abilities of pri-miRNA scaffolds to produce exogenous guides. In addition, once highly efficacious single pri-miRNA scaffolds have been chosen, the order in which to combine them to ensure that potent guides are produced from each mimic had not been thoroughly investigated. To address these issues, 16 individual pri-miRNAs encoding one of four exogenous guide sequences known to be highly efficacious against HIV were designed and their abilities to generate discrete, potent anti-HIV guides were compared. Furthermore, various combinations of the most efficacious scaffolds were examined thereby providing valuable data relating to the generation of expressed combinatorial pri-miRNA systems.

Replacement of the endogenous guide strand within pri-miR31 resulted in a set of mimics that differed significantly in both shape and thermodynamic stability as compared to the endogenous scaffold (Figure 26) and this set of constructs mediated varying levels of inhibition depending on the guide sequence present (Figure 28). In contrast, the remaining scaffolds (pri-miRNAs 122, 30a and 155) resulted in mimics that resembled the endogenous backbone both in shape and thermodynamic stability (Figure 26) and these constructs mediated similar levels of inhibition with all four anti-HIV guides (Figure 28). These data strongly suggested that the pattern of target inhibition was the same for each pri-miRNA mimic regardless of the guide sequence present as long as the folding parameters of the mimic closely resembled those of the endogenous scaffold. Of particular note, pri-miR30a-based mimics mediated the greatest levels of target suppression with all four anti-HIV guides at levels comparable to the shRNA controls (Figure 28B). The notable difference between this scaffold and the remaining three was that guide sequences were predominantly expressed from the 3' arm of pri-miR30a backbones (Figure 26 second panel from left). Drosha defines the 5' end of the mature miRNA by cleaving 11 bp from the ss/ds junction of a pri-miRNA (Zeng and Cullen, 2005; Han, Lee*, et al.*, 2006). Should the cleavage site be altered by even a few bases, 5' dominant guide sequences would be negatively affected as the cleavage point may occur *within* the seed region. However, 3' dominant guides would be less affected as Drosha cleavage defines the 3' end of these miRNAs (see Figure 10). As different anti-HIV guide strands mediated varying levels of inhibition depending on the pri-miRNA backbone from which they were processed (Figure 28), guide sequence efficacy seems to be partly dependent on the pri-miRNA scaffold. All 4 anti-HIV guides had been selected based on their potencies when expressed from shRNA backgrounds

yet comparable levels of target suppression were only mediated when the guides were expressed from pri-miRNA 30a and 155 backbones (Figure 28). This was intriguing as pri-miRNAs 122 and 31 had been selected based on previous data that showed potent exogenous guides were produced from these 2 scaffolds (Ely, Naidoo*, et al.*, 2008). Dosage may have played a role but increased concentrations of the latter two mimics did not improve their inhibitory efficacies or the ability to detect putative guide sequences by Northern blot (data no shown). This suggested that certain pri-miRNA backbones may be poor mimic scaffolds although the reasons for the discrepancies between the current body of work and that of Ely *et al* remain unclear.

The position of each individual pri-miRNA mimic within the polycistronic system did not affect the guide sequence efficacy when targeted to luciferase-based target vectors (Figure 30) suggesting that all four pri-miRNA components were effectively processed by Drosha and Dicer. The spacing between each component in the multimeric constructs had not been altered although the complete ~100 bases of ss flanking regions for each individual scaffold were included (~50 bases in the 5' flank and ~50 bases in the 3' flank). This was similar to the ~102 bases of flanking region incorporated in combinatorial HBV-specific pri-miRNA mimics that were shown to be highly efficacious against all three HBV target mRNAs (Ely, Naidoo*, et al.*, 2009). However, Drosha recognises a 33 bp stem as well as the ss/ds junction (Zeng and Cullen, 2005; Han, Lee*, et al.*, 2006) suggesting that the flanking regions could be altered. The reasons for the lack of inhibition mediated by all four polycistronic systems against the psi-Tat target vector was unclear as sequence analysis did not reveal mutations within these constructs (data not shown). Furthermore, sequence analysis and the observation that the shTat control mediated highly significant target suppression suggested that the psi-Tat construct did not contain any mutations. Despite these data, the inhibition by all four polycistronic systems against the remaining three targets provided proof-of-concept data that the multimeric pri-miRNA constructs were effective.

Importantly, the inhibitory efficacies of the individual pri-miRNA mimics could be dramatically enhanced when cloned as part of a combinatorial system suggesting that the suppressive effect was additive. This has implications for the treatment of wild type HIV as ideally each component of the system should present a formidable resistance barrier to the virus and if one mimic is producing sub-optimal concentrations of guide sequence, this may provide an escape route. Indeed, the Northern blot analysis described in section 3.3.5 suggested that pri-miRNA backbones produced more guide sequence when expressed from individual mimics as the concentration of guides produced from combinatorial mimics was undetectable. Interestingly, very faint or no guide sequence signals were detected using radio-labelled probes for either of these mimics (Ely, Naidoo*, et al.*, 2008) lending support to the data described in section 3.3.5 of this chapter. However, while the majority of multimeric pri-miRNA mimics

described here mediated highly significant inhibition of a wild type HIV isolate (see section 3.3.6), more work is needed to understand the role and contribution of each processed siRNA within the combinatorial scaffold.

The data presented in this chapter were supported by two studies published concurrently to this work that described the manipulation of innate polycistronic pri-miRNAs to express multiple guide sequences targeted to HIV (Aagaard, Zhang*, et al.*, 2008; Liu, Haasnoot*, et al.*, 2008). The first study replaced four guides within a CMV-expressed polycistonic cluster of 6 pri-miRNAs (pri-miR17-92) and targeted separate sites (*int*, *gag*, *pol* and *tat*/*rev*) within HIV (Liu, Haasnoot*, et al.*, 2008). To generate an effective anti-HIV polycistonic pri-miRNA construct, the authors initially developed an optimal pri-miRNA construct by varying the length (19 to 23 bp) and position (5' arm vs. 3' arm) of an anti-HIV guide sequence within two separate pre-miRNA backbones based on pri-miRNAs from the polycistron. The inhibition mediated by the different pre-miRNA mimics showed that guide sequences of 19 bp processed from the 5' arm of the pre-miRNA were the most efficacious and these were constructed into multimers of two, four or six anti-HIV pri-miRNAs. Similarly to the results described in section 3.3.4, the position of the single pri-miRNA mimic within the polycistronic unit did not affect guide sequence efficacy. Furthermore the inhibitory efficacies of single pri-miRNA mimics could be dramatically enhanced when cloned as part of a combinatorial system. Interestingly, when targeted to a molecular infectious clone of HIV (pLAI), the majority of the single pri-miRNA mimics and one of the 2mer constructs were poorly efficacious. The remaining 2mers, 3mer and 4mer constructs were all significantly active but in contrast to the data described in section 3.3.5 using FV5, the inhibition was measured two days post-infection and the assay was not completed on a wild type isolate of HIV. In a separate assay, stable cells that expressed single anti-HIV pri-miRNA constructs showed resistance mutants six days post-infection while cells that expressed a 4mer maintained suppression levels similar to the mock although not as great as that of a shRNA control. In the second study all three guides within an endogenous tri-cistronic pri-miRNA cluster (pri-miR106b, pri-miR-93 and pri-miR-25) were substituted with anti-HIV sequences targeted to the *tat*/*rev* overlapping ORF (Aagaard, Zhang*, et al.*, 2008). Initially, two versions of the endogenous pri-miRNA cluster were assessed – one that included the exonic sequences and one that did not, but no differences were observed in the inhibitory efficacies between these two suggesting that the exonic sequences did not enhance activity. To substitute in anti-HIV guide sequences, each entire endogenous pri-miRNA within the cluster was replaced with stem-loop structures specific to HIV. While this provided a modular system that could be readily manipulated to introduce any stem-loop sequence of choice, the hybrid construct was not a pri-miRNA multimer. Inhibition data using a luciferase target reporter showed that two of the three guide sequences were highly efficacious while the third was poorly processed and thus did not mediate significant

suppression. This could be marginally improved if the ss/ds junctions of the pri-miRNA backbones were re-introduced and the guide sequence was substituted for the natural guide within the endogenous backbone. Despite these changes, when targeted to two HIV infectious molecular clones (pHIV-IIIB and pLAI), repression was poor and significant inhibition was only achieved when a TAR RNA decoy was included in the combinatorial strategy. Thus this study represented a combinatorial RNAi approach comprising pre-miRNA mimics and a TAR decoy and not a combinatorial pri-miRNA study as suggested.

In conclusion, the pri-miRNA mimics described in this chapter were capable of expressing potent anti-HIV guide sequences although subsequent inhibition depended on the endogenous pri-miRNA scaffold used. New data has been provided regarding the construction of a modular combinatorial pri-miRNA-based system that can mediate highly significant suppression of a wild type isolate of HIV at four different target sites. Taken together, these data have provided a foundation for future investigation into combinatorial RNAi using an expressed system that can be linked to a tissue-specific promoter to provide stable, long-term gene suppression. Additional data is required to further understand how these mimics can be designed to ensure consistent and potent guide sequence production such that viral escape no longer causes treatment failure. Studies that explore different pri-miRNA backbones to those described in this chapter may provide additional examples of scaffolds that could be included in modular combinatorial systems. Investigations into the role of splicing in combinatorial pri-miRNA mimics may shed light on design criteria that could improve pri-miRNA mimic processing. For example, would the incorporation of splice donor/acceptor sites within the flanking regions of adjacent pri-miRNA mimics affect the overall processivity of the combinatorial system? In addition, data that describes the use of inducible or tissue-specific Pol II-driven pri-miRNA mimics would be valued. Finally, the delivery of combinatorial pri-miRNA mimics to target cells of interest remains to be elucidated and needs to be thoroughly tested before these constructs can be introduced into patients.

An additional mechanism that could be employed to enhance the overall post-transcriptional silencing effect of a combinatorial pri-miRNA system is the use of small RNAs such as siRNAs to induce a transcriptional silencing effect. Although highly experimental, ideally this would result in permanent silencing of the viral LTR promoter thereby limiting the chance for viral escape through spikes in transcription of proviral DNA. As the possibility of silencing the LTR of HIV has only briefly been investigated, data that explores numerous sites within the LTR for siRNA susceptibility would be of great value.

4.1 Introduction

The data in the previous two chapters described novel combinatorial RNAi based strategies designed to inhibit sequences within the HIV LTR specifically at the post-transcriptional level. Transcriptional gene silencing of promoters has also been described although the majority of work to date has been conducted in yeast (*Schizosaccharomyces pombe*), worms (*Caenorhabditis elegans*) and plants (*Arabidopsis thaliana*) [reviewed in (Hawkins and Morris 2008; Morris 2009)]. The mechanism of TGS in mammalian cells is a result of RNA directed methylation of histones and DNA at the targeted promoter although the latter observation is not a conserved feature in all cases (Morris, Chan et al. 2004; Castanotto, Tommasi et al. 2005; Kim, Villeneuve et al. 2006; Han, Kim et al. 2007). In an effort to prolong RNAi-mediated silencing of HIV, Suzuki *et al* targeted siRNAs to conserved tandem NF-κβ sites within the viral LTR (Suzuki, Shijuuku et al. 2005; Suzuki, Juelich et al. 2008). Significant long-term (30 days) inhibition of HIV that correlated with the extent of *de novo* methylation of the LTR was observed and similar TGS data was described for siRNAs targeted to the 5' LTR of SIV (Lim, Suzuki et al. 2008). While these studies demonstrated that HIV was susceptible to TGS, little data was available describing the regions within the LTR that could be targeted. One site within the HIV promoter, designated ‘247’ (located at positions 249–267 of the HXB2 molecular clone sequence, GenBank® accession number K03455) was susceptible to TGS but the silencing effect was shown to be directed by the antisense (as) strand alone (Weinberg, Villeneuve et al. 2006). Specifically, 21 nt asRNAs mediated a 2.8 fold greater suppression of target site 247 as compared to a double stranded (siRNA) counterpart. Additional data revealed that the effect was as a result of low-copy promoter-associated RNAs that are recognised by the antisense strand and function as a binding site for epigenetic silencing complexes (Han, Kim et al. 2007).

In this chapter, to further investigate the effects of 21 nt asRNAs in TGS, and to define additional asRNAs that target the HIV-1 LTR promoter, RNA Pol III (U6)-expressed asRNA constructs were generated that span ~ 50 bp up- and downstream of target site 247 (Weinberg, Villeneuve et al. 2006). A total of ten asRNAs (including as247) were assessed for their abilities to induce TGS when targeted to subtype B luciferase-based reporter constructs. Interestingly, one of the asRNAs that targeted seven bp downstream of site 247 seemed to induce a gene activation effect. The site correlated with a gene of unknown function (C10orf76) and microarray as well as chromatin immunoprecipitation (ChIP) analysis was conducted to investigate the observed activation effect.

4.2 Materials and methods

4.2.1 Cloning of expressed asRNAs targeted to site 247 and its flanking regions within the U3 region of a subtype B HIV LTR

To generate asRNAs targeted to site 247 in the HIV LTR as well as regions that span 50 bp up and downstream, PCR was carried out using a universal U6 Forward primer containing a *Spe*I restriction site (5' CTA ACT AGT GGC GCG CCA AGG TCG GGC AGG AAG AGG G 3', the *Spe*I site is underlined) and the following reverse oligonucleotide primers: LTR-247s+7 (5' AAA AAA GAG TGG AGG TTT GAC AGC CCG GTG TTT CGT CCT TTC CAC AA 3'), LTR-247s+13 (5' AAA AAA GGT TTG ACA GCC GCC TAG CGG TGT TTC GTC CTT TCC ACA A 3'), LTR-247s+19 (5' AAA AAA GAC AGC CGC CTA GCA TTT CCG GTG TTT CGT CCT TTC CAC AA 3'), LTR-247s-7 (5' AAA AAA GAG AGA AGT GTT AGA GTG GCG GTG TTT CGT CCT TTC CAC AA 3'), LTR-247s-14 (5' AAA AAA ACC CTG AGA GAG AAG TGT TCG GTG TTT CGT CCT TTC CAC AA 3'), LTR-247s-21 (5' AAA AAA ATG GAT GAC CCT GAG AGA GCG GTG TTT CGT CCT TTC CAC AA 3'),LTR-247s-1/+19 (5' AAA AAA ATG GAT GAC CCT GAG AGA GAA GTG TTA GAG TGG AGG TTT GAC AGC CGC CTA GCA TTT CCG GTG TTT CGT CCT TTC CAC AA 3'), and LTR-247s-12/+10 (5' AAA AAA CCT GAG AGA GAA GTG TTA GAG TGG AGG TTT GAC AGC CGC CCG GTG TTT CGT CCT TTC CAC AA 3'). The asRNAs were labelled according to their relative positions from site 247 with '+' denoting upstream and '-' denoting downstream. To generate a sense-strand version of LTR-247as+7 as well as mutant versions of LTR-247as+7 (mA and mB), the following reverse oligonucleotide primers were used: LTR-247s+7 Reverse (5' AAA AAA GGG CTG TCA AAC CTC CAC TCG GTG TTT CGT CCT TTC CAC AA 3'), LTR-247as+7 mA Reverse (5' AAA AAA GAG TGG AGG TTT GAC AGC GCG GTG TTT CGT CCT TTC CAC AA 3') and LTR-247as+7 mB Reverse (5' AAA AAA GAG TTG AGG TTT GAC AGC CCG GTG TTT CGT CCT TTC CAC AA 3'). PCR was carried out using Expand High Fidelity^{PLUS} PCR System reagents (Roche, Basel, Switzerland) on a Mastercycler® (BioRad, CA, USA). The PCR mixture included 0.4 µM each of forward and reverse primers, 10 ng of template, 1.5 mM $MgCl_2$, 200 µM dNTP mix, 2.5 U of Expand HiFi^{PLUS} Taq polymerase and 1 x final volume Expand HiFi^{PLUS} Reaction Buffer (proprietary information) in a final volume of 50 µL. PCR conditions comprised 3 minutes at 95 ^{o}C followed by 35 cycles of 30 seconds each at 95 ^{o}C, 58 ^{o}C and 72 ^{o}C followed by 10 minutes at 72 ^{o}C. The amplified products were excised and eluted from a 1 % agarose gel (see Appendices A1.2.3 and A1.2.4) and cloned into the TA cloning vector pTZ57R/T as described in section 2.2.1. Candidate clones were screened using *Spe*I and *EcoR*I.

The construction of pCR2.1-LTR-247as -50/+50 required the same two step PCR approach that had been used to generate anti-LTR lhRNAs (see section 2.2.1) using a universal U6 Forward primer with the following reverse oligonucleotide primers: LTR-247s-50/+50 Reverse 1 (5' AGT GGA GGT TTG ACA GCC GCC TAG CAT TTC ATC ACG TGG CCC GAG AGC TGC ATC CGG AGT ACG GTG TTT CGT CCT TTC CAC AA 3') and LTR-247s-50/+50 Reverse 2 (5' AAA AAA CAG CTT GTT ACA CCC TGT GAG CCT GCA TGG AAT GGA TGA CCC TGA GAG AGA AGT GTT AGA GTG GAG GTT TGA CAG CCG 3'). PCR fragments were cloned into the TA vector pCR2.1-TOPO (Invitrogen, CA, USA) and candidate clones were screened using sequencing as previously described for the anti-LTR lhRNAs (see section 2.2.1).

4.2.2 Cloning of HIV LTR site 247 target vectors

To generate luciferase reporter constructs containing putative LTR-247as+7 target sequences within C10orf76, a set of psiCheck2 based target vectors were generated using the same method for generating an HIV tiling array as described in section 2.2.3 with the following oligonucleotides: LTR-247as+7 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CGA GTG GAG GTT TGA CAG CCC GC 3'), LTR-247as+7 Reverse containing a *Not*I site (5' GGC CGC GGG CTG TCA AAC CTC CAC TCG ATA TC 3'), C10orf76 1049 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CCA GAG TGG GCT CCC TAA ACA GCC CCG C 3'), C10orf76 1049 Reverse containing a *Not*I site (5' GGC CGC GGG GCT GTT TAG GGA GCC CAC TCT GGA TAT C 3'), C10orf76 2230 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CGG GGA GTT GGC TTG TGG TTC CTC CCT TGG ATA GCC TCG C 3'), C10orf76 2230 Reverse containing a *Not*I site (5' GGC CGC GAG GCT ATC CAA GGG AGG AAC CAC AAG CCA ACT CCC CGA TAT C 3'), C10orf76 463 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CCA GTG GAA GGA AGC AGC TAG C 3'), C10orf76 463 Reverse containing a *Not*I site (5' GGC CGC TAG CTG CTT CCT TCC ACT GGA TAT C 3'), C10orf76 1534 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CTA GAG ATG GAG TTT TTC ACC ATG TTA CCC AGG ATG GTC TCG C 3'), C10orf76 1534 Reverse containing a *Not*I site (5' GGC CGC GAG ACC ATC CTG GGT AAC ATG GTG AAA AAC TCC ATC TCT AGA TAT C 3'), C10orf76 210 Forward containing *Xho*I and *EcoR*V sites (5' TCG AGA TAT CGG CAC AGU GGG GAG AGG CCU GC 3') and C10orf76 210 Reverse containing a *Not*I site (5' GGC CGC AGG CCT CTC CCC ACT GTG CCG ATA TC 3'). For the above 'Forward' oligonucleotide sequences, solid underlined regions correspond to *Xho*I and dotted underlined regions correspond to *EcoR*V restriction sites. For the 'Reverse' oligonucleotide sequences, solid underlined regions correspond to *Not*I restriction sites.

4.2.3 Mammalian cell culture

Both the HEK293 and HEK293-derived cell lines (HEK293-CCR5-GFP) were maintained and transfected using FCS-supplemented DMEM as described in section 2.2.4. Jurkat-derived 1G5 cells (NIH AIDS Research and Reference Reagents Program, MD, USA) were maintained in complete Roswell Park Memorial Institute 1640 medium (RPMI 1640, Lorenza Biosciences, Basel, Switzerland) supplemented with heat-inactivated 10 % FCS (Invitrogen, CA, USA).

4.2.4 Transfection and detection of asRNA activity in cultured mammalian cells

To assess the inhibitory efficacies of U6-expressed asRNAs targeted to a subtype B HIV-1 LTR promoter, 1G5 cells were electroporated and quantitative RT-PCR of luciferase mRNA was performed as follows: 1G5s are a Jurkat-derived suspension cell line that express hFluc followed by a SV40 poly A strong stop signal from an HIV-1 subtype B LTR (Aguilar-Cordova, Chinen et al. 1994; Platt, Wehrly et al. 1998). Three million cells per asRNA-expressing construct were washed once in PBS by centrifugation at 215.74 x g followed by a similar wash step using Viaspan™ (DuPont Pharmaceuticals, DE, USA) before being resuspended in 800 µL of ice-cold Viaspan™. Cells were transferred to a 0.4 cm gap 5 cuvette (BioRad, CA, USA), and a combination of 2.5 µg of pTat-dsRed (Unwalla, Li et al. 2004) and 2.5 µg of pTZ-LTR-247s+7 or pTZ-LTR-247s+13 or pTZ-LTR-247s+19 or pTZ-LTR-247s-7 or pTZ-LTR-247s-14 or pTZ-LTR-247s-21 or pTZ-LTR-247s-21/+19 or pTZ-LTR-247s-12/+10 or pTZ-LTR-247s-50/+50 or pTZ-LTR-247s+7 mA or pTZ-LTR-247s+7 mB was added. Samples were electroporated in a BioRad GenePulser (BioRad, CA, USA) using 0.3 kV, 960 µF, a time constant of 11 milliseconds and a field strength of 0.95 kV/cm, before being transferred to 4 ml of warm RPMI 1640 media. Twenty-four hours later total cellular RNA was isolated (see Appendix A1.3.2), DNAse treated (Ambion Turbo DNA-free™, Applied Biosystems, CA, USA), and 1 µg of RNA was converted to cDNA (BioRad iScript™, BioRad, CA, USA) using the same method described in section 2.2.8. Complementary DNA (cDNA) was amplified by PCR with the Quantitect® SYBR® Green PCR Kit (Qiagen, Hilden, Germany) on a Roche Lightcycler V.2. (Roche, Basel, Switzerland). To allow for absolute quantification of RNA, two primer sets were used, one for hFluc mRNA (Forward: 5´ CCT GGA ACA ATT GCT TTT AC 3´ and Reverse: 5´ GTT TCA TAG CTT CTG CCA AC 3´) and one for the GAPDH housekeeping gene (GAPDH Forward: 5´ CCA CCC ATG GCA AAT TCC 3´ and Reverse: 5´ TGG GAT TTC CAT TGA TGA CAA G 3´).

To determine the ability of LTR-247+7as to modulate episomal transcription, HEK293 cells were transiently co-transfected with 100 ng of pLTRb-luc (LAI) or pLTRc-luc-SV40pA and 800 ng of pLTR-247as+7 and 100 ng of either pCMV-Tat or pBluescript (Stratagene, CA, USA)

as described for the episomal transfection of LTR-driven luciferase vectors described in section 2.2.6.

To determine the ability of the various U6-expressed HIV-1 subtype B LTR specific asRNAs to induce off-target enhanced transcription, a CCR5 promoter-expressing GFP construct was used to develop the reporter cell line HEK293-CCR5-GFP. The vector pR5-GFPsg143 contains ~3kb of CCR5 promoter, intron, exons 1 and 2 (Moriuchi, Moriuchi et al. 1997; Mummidi, Ahuja et al. 1997; Guignard, Combadiere et al. 1998) and drives the expression of red-shifted GFP [a gift from Dr. G.N. Pavlakis (Rosati, Valentin et al. 2001; Kim, Villeneuve et al. 2006)]. A total of four million HEK293 cells were seeded into a 10 cm dish and transfected with 5 µg of pR5-GFPsg143 using 5 µL of Lipofectamine 2000 (Invitrogen, CA, USA) in 1 mL of OptiMem (Gibco, BRL, UK) as described for the stable cell line in section 2.2.5. Transfected cells were positively selected over four weeks using 800 µg/mL of Neomycin antibiotic (Merck, Darmstadt, Germany) and fluorescence microscopy. Stable HEK293-CCR5-GFP cells were maintained as described in section 4.2.3. To assess the off-target effects induced by the various U6-expressed HIV-1 subtype B LTR specific asRNAs, HEK293-CCR5-GFP stable cells were seeded and transfected as described above using 50 nM of either siRNA-C10-si1 (5′ GCU GUG AAU CAC AUA UCC CAA 3′) or siRNA-C10-si2 (5′ CCA GAG CTT TGA CAA CCT CAA 3′) or siRNA-C10-si3 (5′ CCT ACT TTG CTG GTT CCC TAA 3′) or control siRNA-AKT-19 (5′ AAC ACC ATG GAC AGG GAG AGC 3′). Twenty-four hours later, total mRNA was extracted (see Appendix A1.3.2) and quantitative RT-PCR was completed to assess mRNA levels of C10or76, *GFP*, *NFYB*, and *NSBP1* relative to *GAPDH* using the following primer sets: C10orf76 Forward (5′ ATG GCC TGG ACC AGT ATG AG 3′), C10orf76 Reverse (5′CCT TGA GCA GGA CTT CTT GG 3’), GFP (Weinberg, Villeneuve et al. 2006), NFYB Forward (5′ GGA ATT GGT GGA GCA GTC AC 3′), NFYB Reverse (5′ TGT TGT TGA CCG TCT GTG GT 3′), NSBP1 Forward (5′ AGG CAC CAG CTT CTG AAA AA 3′) and NSBP1 Reverse (5′ GCT GCC ACT GCT TCT TTC TT 3′).

To determine if the predicted LTR-247as+7 target sites within C10orf76 were susceptible to LTR-247as+7-mediated inhibition, HEK293 cells were seeded and transfected as described for the luciferase assays in section 2.2.6 with the following changes: cells were co-transfected in triplicate with 150 ng of either psi-LTR-247as+7 or psi-C10orf76-1049 or psi-C10orf76-2230 or psi-C10orf76-463 of psi-C10orf76-1534 or psi-C10orf76-210 and 100 ng of pCi-eGFP and 750 ng of pCR2.1-LTR-247as or pTZ-U6+1. Luciferase data were analysed 48 hours post-transfection as described in section 2.2.6.

Total RNA was also used to determine whether or not LTR247as+7 could directly interact with C10orf76 as follows: siRNA-247as+7 (5' CAA ACC UCC ACU CUA ACA C 3') or siRNA-C10-si1 or negative control siRNA-AKT-19 were incubated with total DNAse-treated cellular

RNA for 5 minutes at 95°C followed by cooling to room temperature and conversion to cDNA as described in section 2.2.8 although no polyA primer was added. PCR using C10orf76 Forward and Reverse primers was completed on the cDNA samples as described for IFN-related genes in section 2.2.8.

4.2.5 Chromatin immunoprecipitation (ChIP)

Previous work had shown that siRNA-mediated TGS in human cells involves histone methylation at the siRNA-targeted promoter, specifically di-methylation at histone 3 lysine 9 (H3K9me2) or tri-methylation at histone 3 lysine 27 [H3K27me3,(Kim, Villeneuve et al. 2006; Weinberg, Villeneuve et al. 2006; Han, Kim et al. 2007)]. To determine whether the observed increase in transcription was the result of the asRNA-directed histone modifications associated with increased gene expression, a ChIP assay was performed specifically for H3K4me2. Electroporation (see section 4.2.4) was used to co-transfect 1G5 cells with a combination of 2.5 µg pTat-dsRed and 2.5 µg of pTZ-LTR-247as or pTZ-LTR-247as+7 or pTZ-LTR-247as-7 or pTZ-LTR-247as+13. The following day cells were cross-linked for ten minutes at room temperature using 1 % formaldehyde (Merck, Darmstadt, Germany) and the reaction was stopped by the addition of 0.125 M glycine (Merck, Darmstadt, Germany) and incubation for ten minutes at room temperature. Cells were washed twice in PBS (Gibco, BRL, UK) supplemented with 5 µM PMSF (phenylmethanesulphonylfluoride protease inhibitor, Merck, Darmstadt, Germany), resuspended in 600 µL of ChIP lysis buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 10 % Triton X-100, 0.1 % NaDeoxycholate, 5 µM PMSF, Merck, Darmstadt, Germany) and incubated for ten minutes on ice. Samples were centrifuged for five minutes in a refrigerated centrifuge (Eppendorf 5415D/R, Eppendorf, Hamburg, Germany) at 238 x g and 4 ^{o}C followed by resuspension in 600 µL of ChIP lysis buffer and incubation for ten minutes on ice. Samples were sonicated using a Branson Sonifier® 450 (Emerson Industrial Automation, MO, USA) with six intervals of 20 second pulses and two minute rests. Sonicated samples were centrifuged for ten minutes at 16 000 x g and 4 ^{o}C, supernatants were removed and pre-cleared for 15 minutes using 30 µL of protein A/ salmon sperm DNA mix (Upstate Biotechnology Inc, NY, USA) on a Nutator^{TM} (TCS Scientific, PA, USA). Pre-cleared supernatants were centrifuged for five minutes at 16 000 x g and 4 ^{o}C. Five hundred µL aliquots of each sample were incubated overnight at 4 ^{o}C on a Nutator^{TM} with 1 µg of either anti-H3K9 antibody (Upstate Biotechnology Inc, NY, USA) or anti-H3K27 antibody (Upstate Biotechnology Inc, NY, USA) or no antibody (no-antibody control). The following day samples were incubated for 15 minutes with 10 µL of protein A/salmon sperm DNA at room temperature on a Nutator^{TM} followed by centrifugation for one minute at 16 000 x g and 4^{o}C. The ‘no-antibody control’ supernatants were stored at -20 ^{o}C and

used as ‘total chromatin input’ control samples. The remaining pellets were washed as follows: two washes with 1 mL of ChIP lysis buffer, two washes with 1 mL of ChIP lysis buffer with high salt (50 mM HEPES pH 7.5, 500 mM NaCl, 1 % Triton X-100, 0.1 % NaDeoxycholate, 0.5 µM PMSF) and two washes with 1 mL of ChIP wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5 % NP-40, 0.5 % NaDeoxycholate, 1 mM EDTA, Merck, Darmstadt, Germany). Per wash, samples were incubated for three minutes at room temperature on a Nutator^{TM} followed by centrifugation for three minutes at 16 000 x g. After the final wash, DNA/protein complexes were resuspended in 100 µL of elution buffer (50 mM Tris-HCl pH 8.0, 1 % SDS, 10 mM EDTA, Merck, Darmstadt, Germany), incubated for 10 minutes at 65 ^{o}C followed by room temperature centrifugation for three minutes at 16 000 x g. The eluted complexes as well as 200 µL of the ‘no antibody control’ were incubated for six hours at 65 ^{o}C with 1 µL of RNAse A (10 mg/mL, Invitrogen, CA, USA) and 20 µL of 5 M NaCl (Merck, Darmstadt, Germany) to reverse cross-link the complexes. Reverse cross-linked samples were incubated for one hour at 45 ^{o}C with 10 µL of 0.5 M EDTA, 20 µL of 1 M Tris-HCl pH 6.5 and 2 µL of 10 mg/mL Proteinase K (Invitrogen, CA, USA) and the DNA was recovered by phenol/chloroform extraction (see Appendix A1.2.2). Samples were diluted 300 fold in nuclease-free water (Qiagen, Hilden, Germany) and 2 µL per sample was amplified by PCR using Quantitect® SYBR® Green (Qiagen, Hilden, Germany) on a Roche Lightcycler V.2. (Roche, Basel, Switzerland) as described in section 2.2.8 with the following HIV-1 LTR specific oligonucleotide primers: LTR Forward (5´CAC ACA AGG CTA CTT CCC TGA 3´) and LTR Reverse (5´ GGC CAT GTG ATG AAA TGC TA 3´).

4.2.6 Microarray analysis

Microarray analysis was used to specifically determine the genes involved in the pTZ-LTR-247as+7 mediated increase in gene expression. 1G5 cells were co-transfected in triplicate with either 2.5 µg of pTZ-LTR-247as and 2.5 µg of pTat-dsRed or 2.5 µg of pTZ-GFPas (Weinberg, Villeneuve et al. 2006) and 2.5 µg pTat-dsRed. Twenty-four hours later total RNA was isolated (see Appendix A1.3.1) and 1 µg of RNA per sample was labelled and hybridised to the Affymetrix Human HG-U133 Plus 2.0 array (Affymetrix Inc, CA, USA) at the Scripps Research Institute Core Microarray facility. The HG-U133 Plus 2.0 array includes probes for ~52,000 human probe-IDs. The Affymetrix chips were scanned using an Affymetrix GeneChip Scanner 3000 (Affymetrix Inc, CA, USA). Chips have a background of less than 50 intensity units and a GAPDH 3´/5´ ratio of less than three. Robust Multichip Average (RMA) was used to convert the intensity values to expression values (Bolstad, Irizarry et al. 2003; Irizarry, Hobbs et al. 2003). RMA consists of a three-step approach which uses a background correction on the Perfect Match (PM) probes, a quantile normalisation and summarisation of the probe set information

using Tukey's median polish algorithm. Present and absent calls were calculated in the R software package as implemented in the Affymetrix Microarray Suite version 5. This algorithm uses a Wilcoxon signed rank-based calculation to assign presence or absence of probe sets. Probe sets were filtered out if they were assigned absent for all the samples, which left ~30,000 probes for further analysis. ANOVA was performed using BRB Array Tools, developed by Dr. Richard Simon and Amy Peng Lam. BRB utilises multivariate permutation tests to ensure that the number or proportion of false discoveries is controlled and is effective when the number of samples is at least three per treatment. The ANOVA analysis identified 185 human genes on the chip that were significantly altered at a P-value of 0.05, which were considered as false positives. Heat-maps were generated with Cluster and TreeView programs written by Michael Eisen (Eisen and Lucchesi 1998). Based on multiple positive matches using miRNA target-site online search algorithms 'RNA22' (Miranda, Huynh et al. 2006) and 'RNAhybrid 2.2' (Rehmsmeier, Steffen et al. 2004), three genes were chosen for microarray analysis validation by RT-PCR using the same RNA samples and primer sets specific to C10or76, *NFYB*, and *NSBP1* as described for off-target determination in section 4.2.4.

4.2.7 Statistical analysis

Data are expressed as the mean ± the standard error of the mean (SEM) and where appropriate are normalised to a control sample within the experiment. Statistical difference was considered to be significant when $p < 0.05$ and highly significant when $p < 0.01$. In all cases $\alpha = 0.99$ and data was analysed using a one-way ANOVA followed by a Dunnett's Multiple Comparison post-test or a Student's t-test calculated using GraphPad Prism (GraphPad Software Inc, CA, USA).

4.3 Results

4.3.1 Selection of HIV-1 subtype B LTR-targeted asRNAs

Data had been published showing that only the antisense strand of a promoter-directed siRNA is required to initiate transcriptional silencing in human cells (Weinberg, Villeneuve et al. 2006). Specifically, a U6-expressed 21 nt asRNA targeted to site 247 (denoted as LTR-247as), which lies 205 bp upstream of the HIV-1 transcription start site, was capable of suppressing Tat-mediated transcription of HIV-1 LTR-luciferase expression in 1G5 cells (Weinberg, Villeneuve et al. 2006). 1G5 cells possess an integrated LTR-driven luciferase reporter gene construct, which is responsive to the HIV transactivating protein Tat (Aguilar-Cordova, Chinen et al. 1994). In order to comprehensively characterise the suppressive efficacy of LTR-247as to the 247 target

site, a series of U6-expressed asRNAs were generated to target the region surrounding site 247 including sequences that spanned 50 bp up- and downstream of this site (Figure 33). It was hypothesised that site 247 is intrinsically unique with regards to siRNA and/or asRNA targeted TGS.

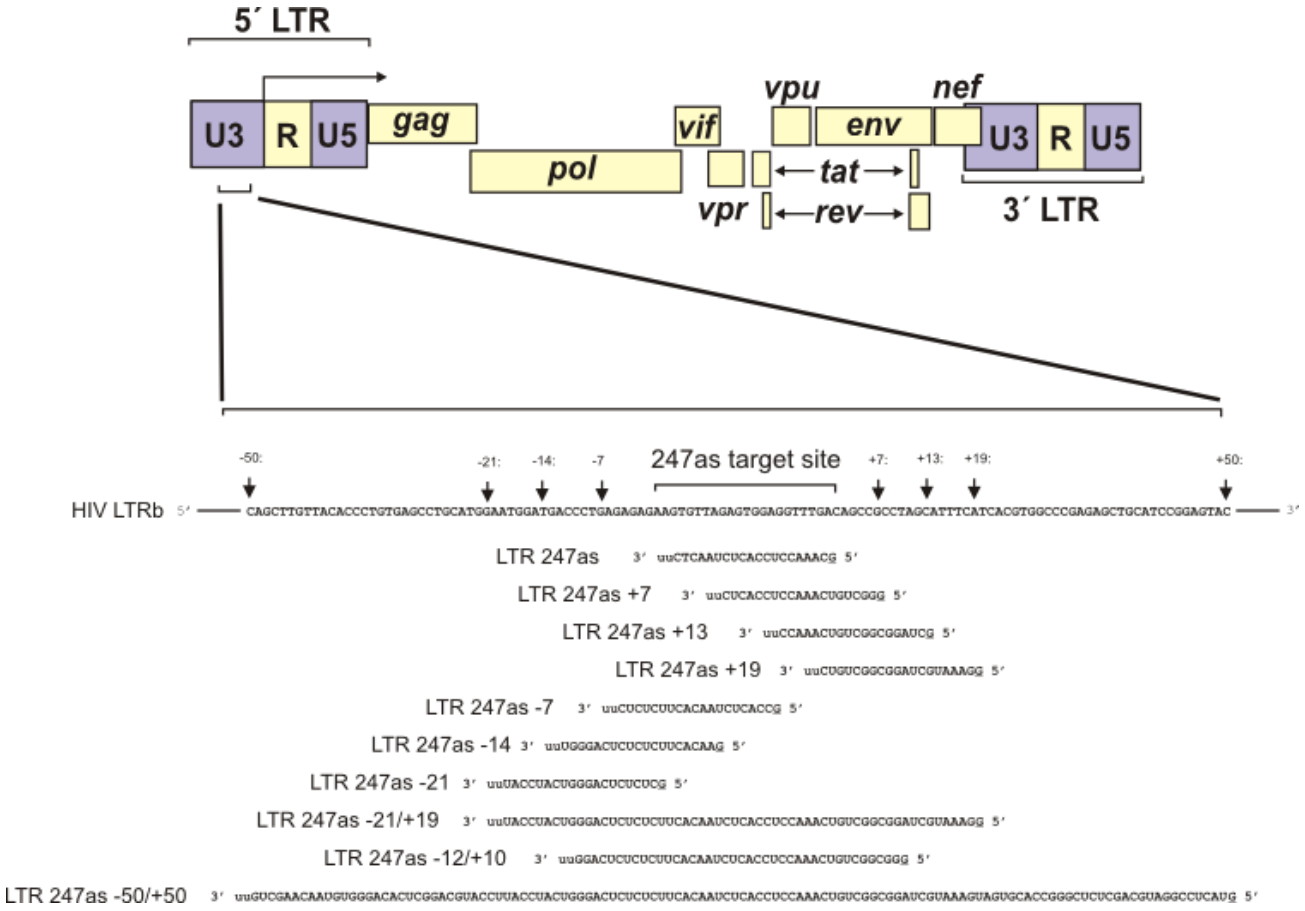


Figure 33. Region within the HIV LTR targeted by U6-expressed asRNAs. A set of nine asRNAs were selected that targeted either site 247, or regions which span upstream (+) and/or downstream (-) of 247. The respective asRNAs are shown aligned with the HIV-1 subtype B LTR and site 247 (targeted by LTR-247as).

4.3.2 Screening of HIV-1 LTR targeted U6-expressed asRNAs

To determine asRNA-mediated promoter suppression, the U6-expressed asRNA plasmids were transiently co-transfected into 1G5 cells with a Tat expression plasmid, pTat-dsRed (Platt, Wehrly et al. 1998). Total RNA was extracted 24 hours later and hFluc mRNA relative to *GAPDH* mRNA was assessed (Figure 34). Interestingly, one of the U6-expressed asRNA constructs, LTR-247as+7, which is shifted 7 bases downstream relative to LTR-247as (Figure 33), produced

a significant increase in luciferase mRNA expression when compared to an asRNA targeted to GFP (α = 0.99, p < 0.01). LTR-247as (Figure 34 dark blue bar) and LTRc-247as [targeted to the same sequence on HIV subtype C, isolate Du151 (GenBank® accession number DQ411851), Figure 34 turquoise bar], suppressed Tat-mediated activation of the LTR similarly to the subtype B-specific LTR-247as.

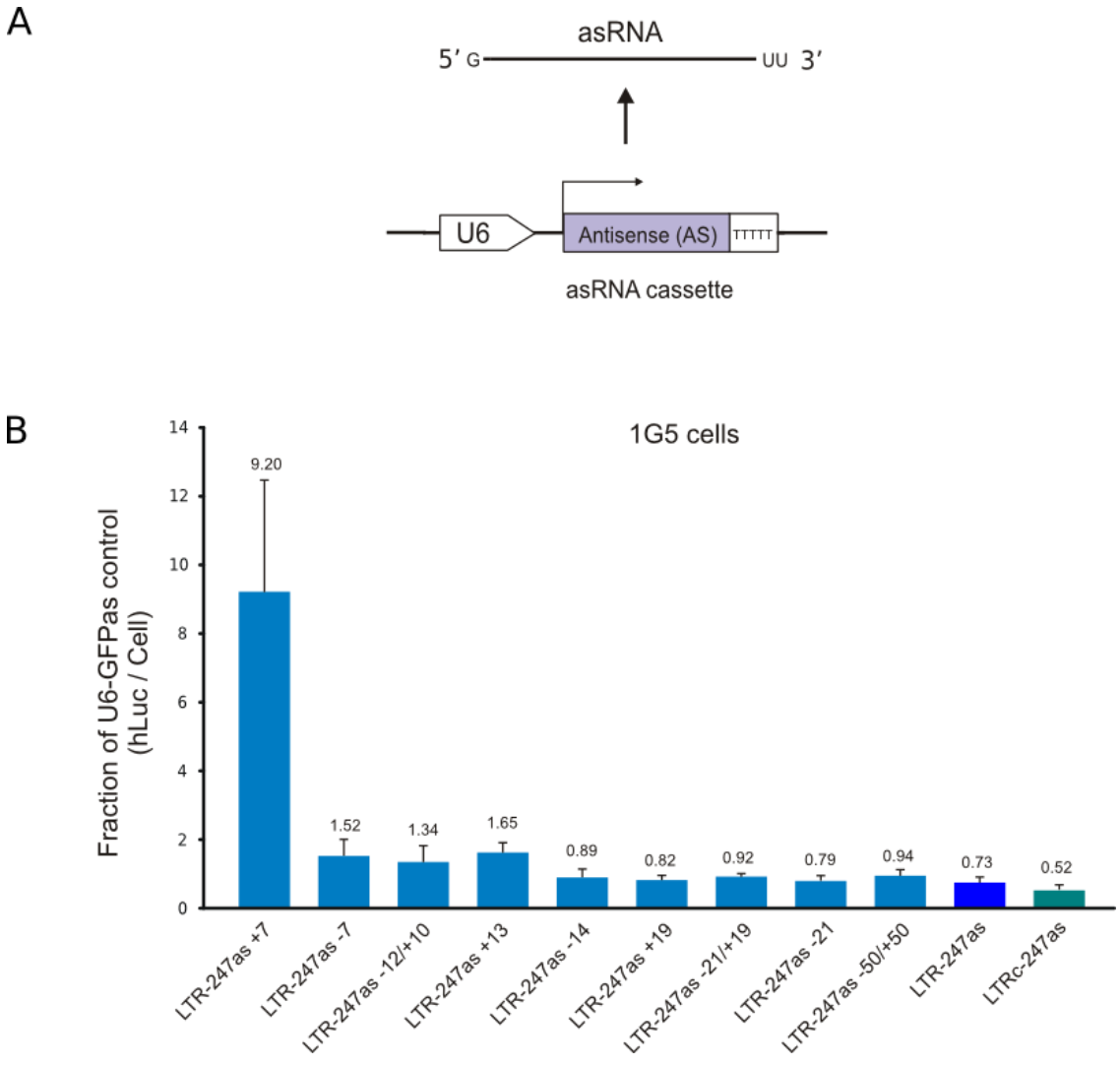


Figure 34. Effects of various U6-expressed antisense RNAs on luciferase expression. A) Design of expressed antisense RNA cassettes. Pol III (U6) cassettes were designed to express a single asRNA with characteristic 2 bp 3' overhangs. B) 1G5 cells containing an integrated HIV-1 subtype B LTR driving hFluc were co-electroporated with pTat-dsRed and each of the LTR-247as vectors (blue bars). Luciferase expression was assessed 24 hours post-transfection relative to *GAPDH* expression by real-time RT-PCR. Data are normalised to control-treated cultures (pU6-GFPas, n = 3, ± SEM).

To determine if the increased gene activation was independent of the activity of HIV-1 Tat, 1G5 cells were co-transfected with either LTR-247as+7 and the HIV-1 Tat expression plasmid, pTat-dsRed, or LTR-247as+7 and an irrelevant plasmid, pBluescript (pBSK). There was a ~1.5 fold enhancement of LTR-mediated luciferase expression in the absence of HIV-1 Tat (Figure 35), suggesting that the previously observed increase in LTR activity (Figure 34) was specifically mediated by LTR-247as+7.

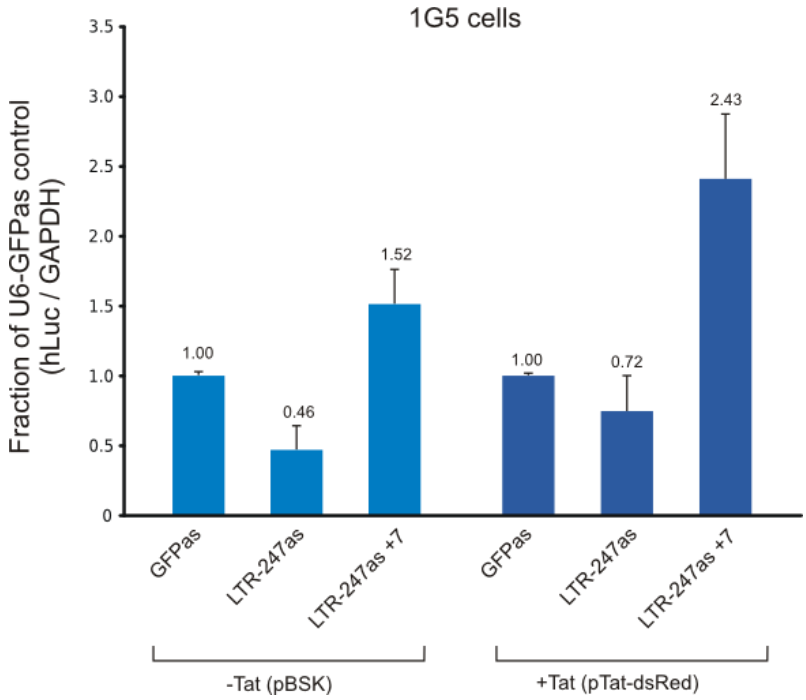


Figure 35. LTR-247as activation is independent of HIV-1 Tat. 1G5 cells were co-transfected with either pBSK (light blue bars) or pTat-dsRed (dark blue bars) and either pU6-LTR-247as, pU6-LTR-247as+7 or pU6-GFPas control. Twenty-four hours later quantitative RT-PCR was performed. Data are expressed as hFluc/*GAPDH* mRNA levels normalised to U6-GFPas control (n = 3, ± SEM).

The activation mediated by LTR-247as shown in Figures 34 and 35 was determined using an integrated LTR target. To assess if these observations were related to the integrated nature of the reporter, transient co-transfections of LTR-247as with episomal subtype B and C LTR-reporter vectors were completed in HEK293 cells. Luciferase expression was assessed 24 hours post-transfection by real-time RT-PCR. Treatment with U6-expressed LTR-247as+7 plasmids resulted in a significant increase ($\alpha = 0.99$, $p < 0.01$) in episomal LTR-driven expression in a Tat-independent manner (Figure 36). While significant, this increase was far less than that mediated against the integrated target vector (compare Figures 35 and 36).

Taken together, the data in Figures 34 to 36 suggested that the U3 region of the LTR spanning 50 bp up-and downstream of site 247as was susceptible to antisense mediated inhibition. However, the stark contrast mediated by 247as+7 showed that Tat-independent target transactivation could occur. This activation was further shown to be independent of the integrated chromatic state of the target. To assess if the 247as+7-mediated activation induced chromatin changes at the target site, ChIP analysis was completed to determine the histone modifications present.

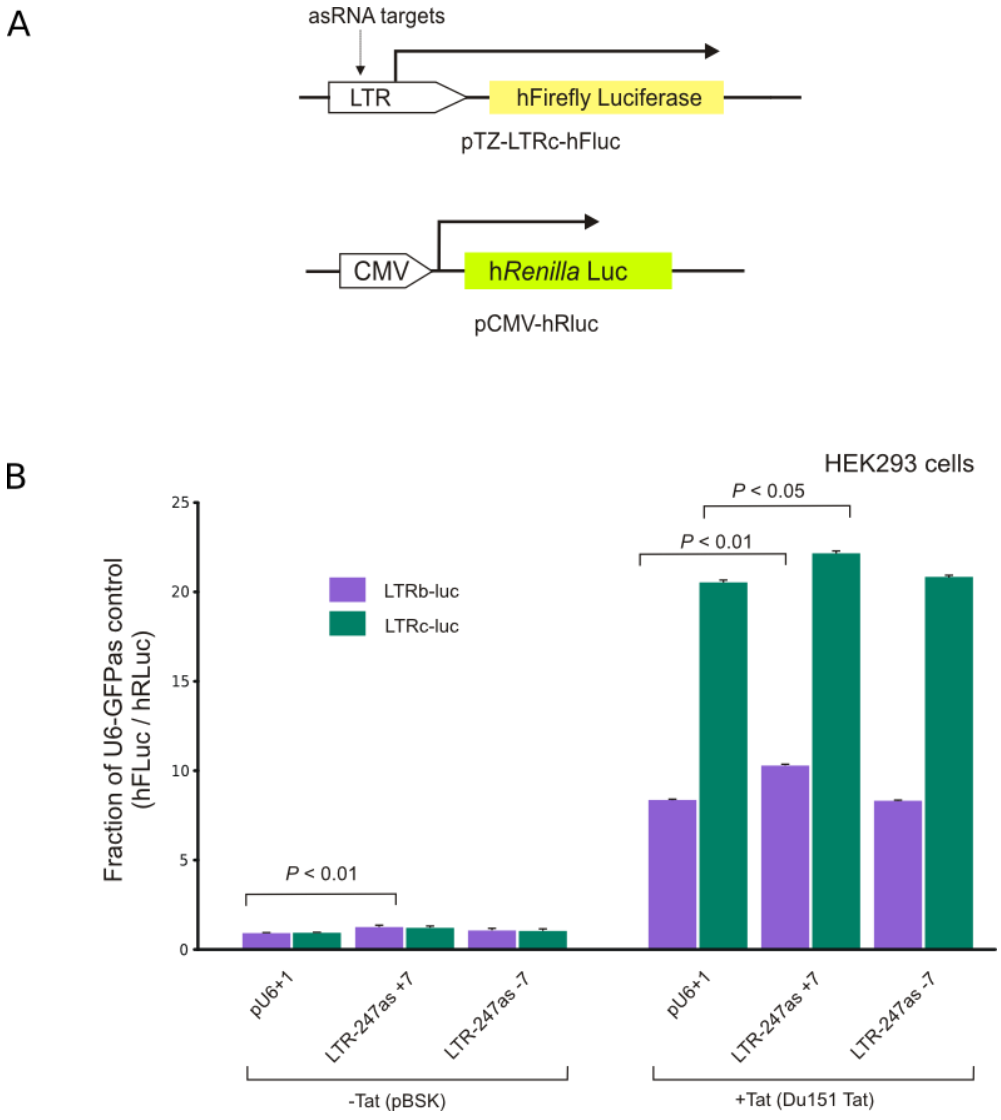


Figure 36. LTR-247as+7 modulates episomal gene expression. A) Either a subtype B or C LTR-driven hFluc expression plasmid was co-transfected in triplicate with the LTR targeted pU6-LTR-247as+7, pU6-LTR-247as-7, or control pU6+1 plasmids with or without HIV-1 Tat and a hRluc-expressing plasmid in HEK293 cells. B) Twenty-four hours post-transfection, luciferase expression was determined. Data are expressed as a ratio of hFluc to hRluc luciferase activity normalised to a pU6-GFPas control (n = 3, ± SEM). A Student´s t-test was used to determine statistical significance.

4.3.3 The role of histone modifications in LTR-247+7as-mediated increase in LTR-driven luciferase expression

Methylation of histone 3 at lysines 9 and 27 (H3K9me2 or H3K27me3) is observed in genes silenced by siRNAs targeted to their respective promoter regions in human cells (Morris, Chan et al. 2004; Janowski, Huffman et al. 2005; Suzuki, Shijuuku et al. 2005; Janowski, Huffman et al. 2006; Kim, Villeneuve et al. 2006), *Schizosaccharomyces pombe* and plants [reviewed in (Matzke and Birchler 2005; Morris 2005)]. While H3K9me2 and H3K27me3 correlates with transcriptionally silent genes (Jenuwein and Allis 2001), H3K4me2 associates with transcriptionally active genes (Krogan, Kim et al. 2003). Furthermore, siRNAs have been shown to participate in the activation of genes via epigenetic modifications that include the acetylation and/or methylation of histone N-terminal residues (Li, Okino et al. 2006; Janowski, Younger et al. 2007). In light of this observation, it was hypothesized that the increase in LTR-luciferase reporter gene expression observed with asRNA LTR-247as+7 may be mediated by asRNA-directed H3K4me2.

Levels of enrichment of H3K4me2 at the targeted LTR-247 promoter site with asRNAs LTR-247as, LTR-247as+7, LTR-247as-7 and LTR-247as+13 was assessed by ChIP. Relative to the suppressive effects of asRNA LTR-247as which did not enrich for H3K4me2, LTR-247as+7, LTR-247as-7 and LTR-247as+13 were enriched for H3K4me2 in 1G5 cells (Figure 37).

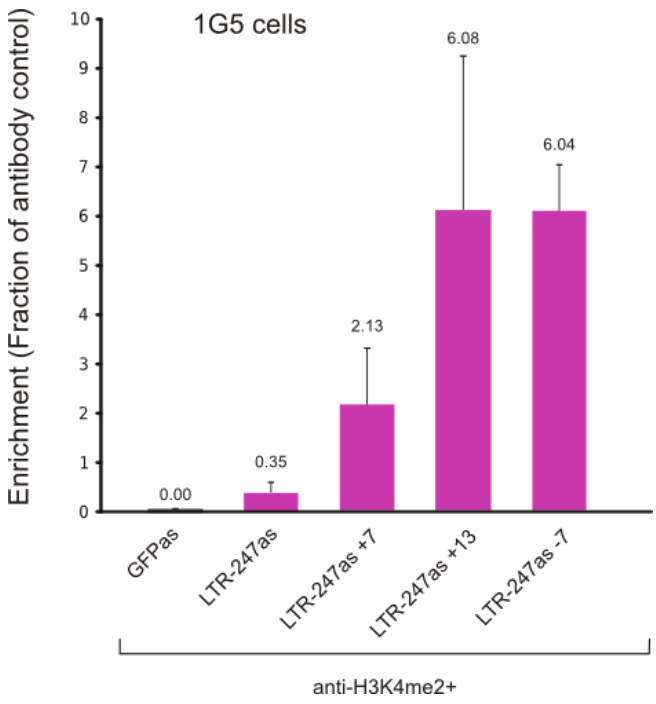


Figure 37. H3K4me2+ in various asRNA-treated cells. 1G5 cells were co-transfected with pTat-dsRed and pTZ-LTR-247as or pTZ-LTR-247as+7 or pTZ-LTR-247as-7 or pTZ-LTR-247as+13 or pTZ-GFPas. The following day, ChIP analysis was performed using anti-H3K4me2+ antibodies and LTR-specific oligonucleotides. Data represent relative enrichment normalised to a 'no antibody' control (n = 3, ± SEM).

Although LTR-247as-7 and LTR-247as+13 showed a six fold increased enrichment of H3K4me2 (relative to a two fold enrichment by LTR-247as+7), these asRNAs were only capable of marginal gene activation when compared to LTR-247as+7 (Figure 34B). The presence of activating H3K4me2 marks may have been a contributing factor in LTR-247as+7-mediated gene activation. However, this alone was not sufficient to explain the marked activation observed relative to LTR-247as-7 and LTR-247as+13.

4.3.4 Single changes in the LTR-247as+7 sequence can abrogate LTR-targeted activation

To determine whether LTR-247as+7 activation was sequence-specific, two U6-expressed LTR-247as+7 RNAs containing single mutations aimed at disrupting sequence-specific interactions between the asRNA and the target were generated (LTR-247as+7mA and mB). In addition, a sense sequence construct (LTR-247s+7) was also included. Construct LTR-247as+7mA included a nucleotide change within the first 7 bases that corresponded to the seed region typically required for microRNA target specificity (Watanabe, Yachie et al. 2006). Furthermore, the nucleotide change discriminated the mA construct from its suppressive LTR-247as

counterpart. Construct LTR-247as+7mB included a nucleotide change within the 3' region of the asRNA that should not have affected seed region binding (Figure 38A).

Co-transfection of 1G5 cells with a plasmid containing the LTR-257as+7 target site, the mutated constructs and HIV-1 Tat showed that the single base change in LTR-247as+7mA (G to C at position 2) was sufficient to partially reverse the activation potential of LTR-247as+7 whereas LTR-247as+7mB did not mediate such a reversal (Figure 38B). This suggested that the activation mediated by LTR-247as+7 was sequence specific and required seed sequence pairing to induce the effect.

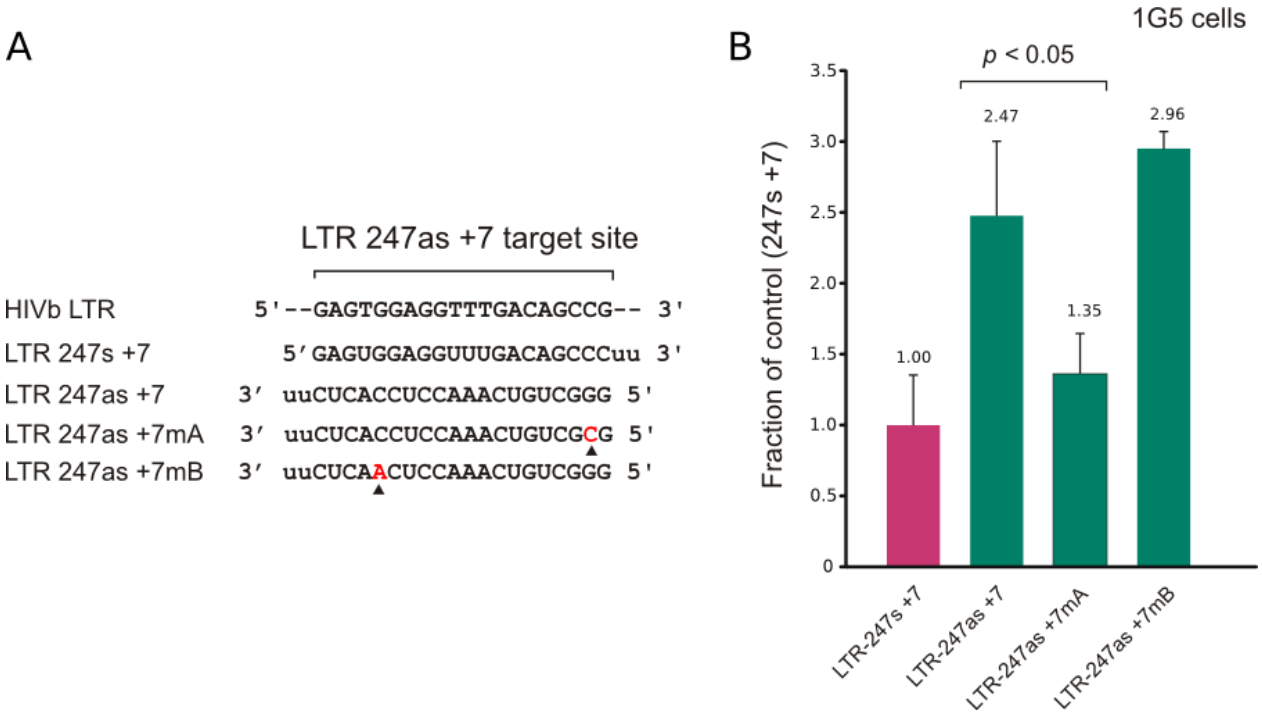


Figure 38. A single point mutation in LTR-247as+7 at nucleotide 7 abrogates off-target gene activation in 1G5 cells. A) Two U6-expressed LTR-247as+7 RNAs containing single mutations aimed at disrupting sequence-specific interactions between the asRNA and the LTR-247as+7 target site were generated (LTR-247as+7mA and mB). A sense sequence construct, LTR-247s+7 was also included. All sequences are aligned relative to the HIV LTRb region corresponding to the LTR-247as+7 target site. Single point mutations are highlighted in red with arrows. B) 1G5 cells were co-transfected in triplicate with a combination of psi-LTR-257as+7 and pCMV-Tat and either one of the mutated constructs or LTR-247as+7 or LTR-247s+7. The following day quantitative RT-PCR was performed. Data are presented as a ratio of hRluc to *GAPDH* mRNA expression normalised to LTR-247as+7 (n = 3, ± SEM). A Student´s t-test was used to determine statistical significance.

4.3.5 Microarray analysis of LTR-247as+7-treated cells

The data from sections 4.3.2 to 4.3.4 showed that LTR247as+7 mediated Tat-independent gene activation from episomal and integrated targets that correlated with a H3K9me2+ marker of

transcriptional activation. To determine if the observed increase in gene expression by LTR-247as+7 was unique to the LTR-luciferase target system present in 1G5 cells, co-transfections with LTR-247as+7 were performed in a cell line that contains an integrated CCR5 promoter constitutively driving GFP expression (Kim, Villeneuve et al. 2006).

Similar to the observations in 1G5 cells, the addition of LTR-247as+7 resulted in an increase in CCR5-expressed GFP mRNA (Figure 39).

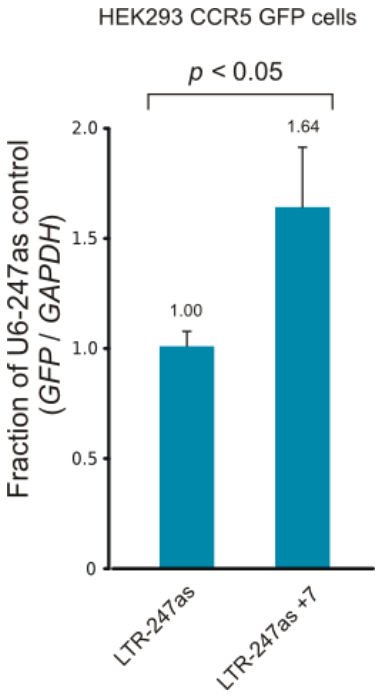


Figure 39. CCR5-expressed GFP shows increased expression in the presence of LTR-247as+7 relative to the control LTR-247as. HEK293 CCR5 GFP cells were co-transfected in triplicate with LTR-247as+7 or LTR-247as as a control and quantitative RT-PCR was performed 24 hours later. Data are expressed as a ratio of *GFP* to *GAPDH* mRNA expression normalised to LTR-247as (n = 3, ± SEM). A Student´s t-test was used to determine statistical significance.

Importantly, no corresponding homologous target site for LTR-247as+7 existed within the CCR5 promoter, indicating a possible off-target effect induced by LTR-247as+7. To determine which genes may have been involved in LTR-247as+7-mediated activation, microarray analysis was performed on 1G5 cells treated with LTR-247as+7 or GFPas control.

A total of 185 genes showed significant alterations in their expression profile when compared to control treated cultures, with RNA-associated proteins such as La, DEAD box proteins (polypeptide 50 and 26B), and a cold inducible RNA binding protein, being of interest (see Table A3.4.1 in Appendix A3.4). Of the 185 deregulated genes, 23 genes could have been involved in producing the observed off-target effect based on their annotation (Figure 40A). The

microarray analysis was further validated by RT-PCR for 3 of the 23 genes which confirmed the corresponding trend of increased expression in the presence of LTR-247as+7 (Figure 40B).

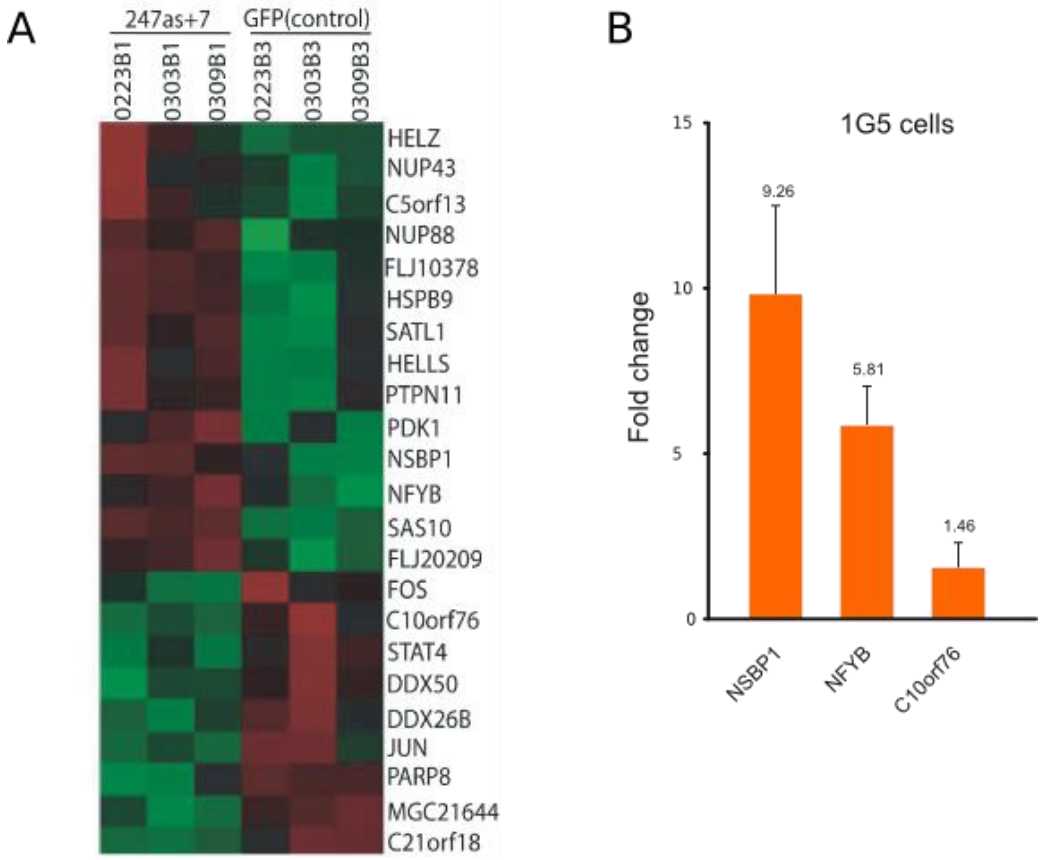


Figure 40. Microarray analysis of 23 deregulated genes. A) A heat-map was produced depicting the variation in gene expression for 1G5 cells treated with LTR-247as+7 or GFPas. Data represent triplicate samples (labelled at the top of the heat-map as 0223B, 0303B and 0309B) with the names of all 23 deregulated genes on the right hand side of the heat-map. B) Microarray analysis was further validated for 3 of the 23 genes that had shown an increase in fold expression relative to GFPas. Real-time quantitative RT-PCR was performed using PCR primers specific to *NSBP1*, *NFYB* and *C10orf76* and *GAPDH*. Data are expressed as a fold change in mRNA expression relative to *GAPDH* (n = 3, ± SEM).

To determine which of the 23 deregulated genes were interacting with LTR-247+7as, the potential miRNA binding characteristics of each gene were assessed using RNA22 (Miranda, Huynh et al. 2006) and RNAhybrid 2.2 (Rehmsmeier, Steffen et al. 2004) software analysis. Both programs find the energetically most favourable hybridisations of small RNAs within large RNAs, thus predicting potential binding sites of miRNAs within large target mRNAs. A suspected non-coding gene expressing a ~2730 nt transcript from chromosome 10 of unknown function (C10orf76, GenBank® accession number NM_024541) was shown to have multiple pairings to LTR-247as+7 (Figure 41).

Figure 41. Screening the miRNA binding potential of LTR-247as+7 and C10orf76. Complementary miRNA for LTR-247as+7 and mRNA target sites for C10orf76 were predicted using RNAhybrid 2.2 and RNA22 software analysis. The predicted miRNA-mRNA duplexes required a total free energy value of less than -20.0 kcal/mol. In addition the first 8 nucleotides of the miRNA in the predicted duplexes required a total free energy value less than -10.0 kcal/mol.

4.3.6 The effect of C10orf76 on LTR-247as+7-mediated gene activation

To assess if C10orf76 functioned to modulate the suppression of genes that were deregulated in the microarray study, C10orf76 was targeted with three different siRNAs and the expression of *NSBP1*, *NFYB* and CCR5-expressed GFP was simultaneously observed (Figure 42). Interestingly, when C10orf76 was suppressed, increased expression of *NSBP1*, *NFYB* and CCR5-expressed GFP was observed that was inversely proportional to the suppression of C10orf76 by the siRNAs. These results matched the increase seen in the microarray data for *NSBP1*, *NFYB* and CCR5-expressed GFP following treatment with LTR-247as+7 (compare Figures 40B and 42B).

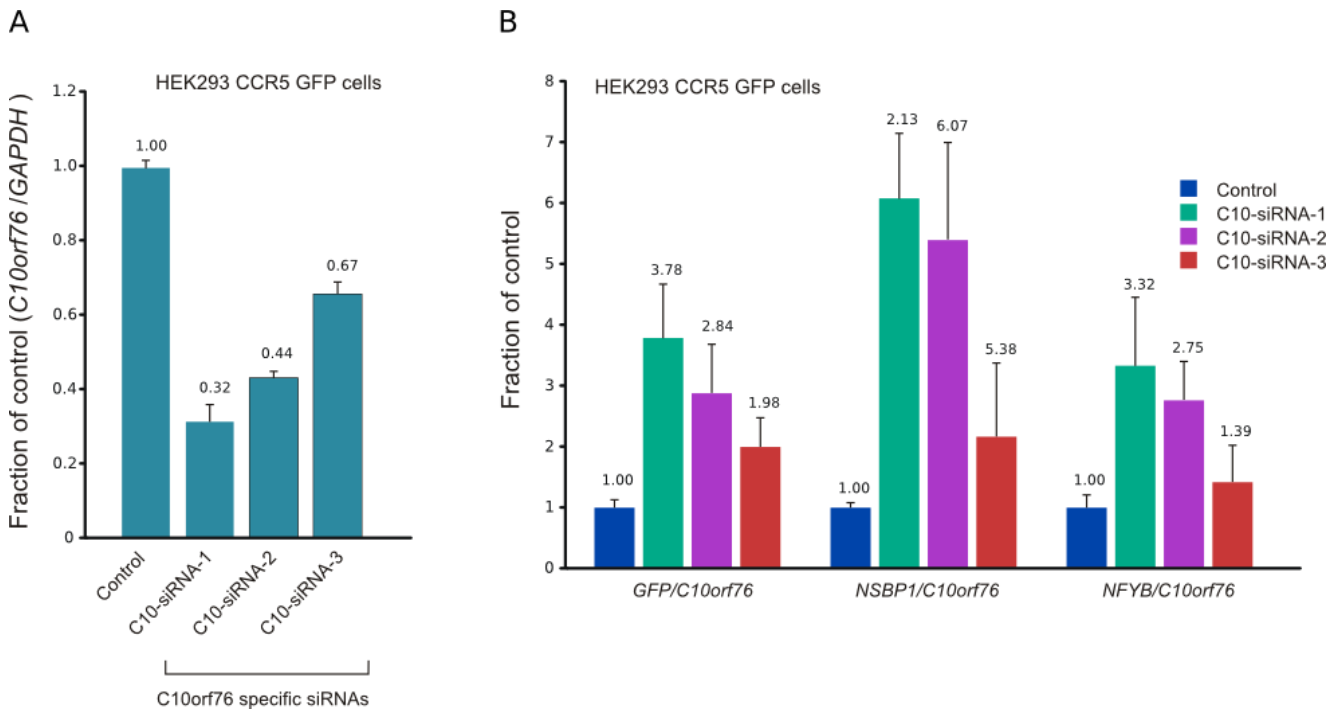


Figure 42. The effect of C10orf76 on 247as+7-mediated gene activation. A) HEK293 CCR5 GFP cells were transfected with siRNAs specific to C10orf76 and quantitative RT-PCR was performed 24 hours later. Data are presented as a ratio of *C10orf76* to *GAPDH* mRNA expression normalised to control siRNA-AKT19 (n = 3, ± SEM). B) Suppression of C10orf76 is inversely proportional to the activation of *GFP*, *NSBP1* and *NFYB* as measured by quantitative RT-PCR 24 hours post-transfection with C10orf76-specific siRNAs in HEK293 CCR5 GFP cells. Data are presented as a ratio of each gene relative to *C10orf76* mRNA expression normalised to control siRNA-AKT19 (n = 3, ± SEM).

The target site search algorithms described in section 4.3.5 showed putative matches between LTR-247as+7 and C10orf76 (Figure 41). To determine if the putative LTR-247as+7 target sites within C10orf76 were susceptible to LTR-247as+7-mediated inhibition, HEK293 cells were co-transfected with LTR-247as+7 and various psiCheck2-based dual-luciferase reporter constructs, each containing individual putative C10orf76 target sequences for LTR-247as+7 (Figure 43A). Relative to a complete LTR-247as+7 target as a control, two of the five putative C10orf76 target sites tested were susceptible to LTR-247as+7 (C10orf76-1049 and C10orf76-463), suggesting that C10orf76 was being directly suppressed via LTR-247as+7 (Figure 43B). Moreover, the C10orf76 transcript could be selectively reverse-transcribed using the LTR-247as+7-siRNA as a primer (Figure 43C). Taken together these data suggested that off-target inhibition of C10orf76 expression was predominantly responsible for the observed gene activation mediated by LTR-247as+7.

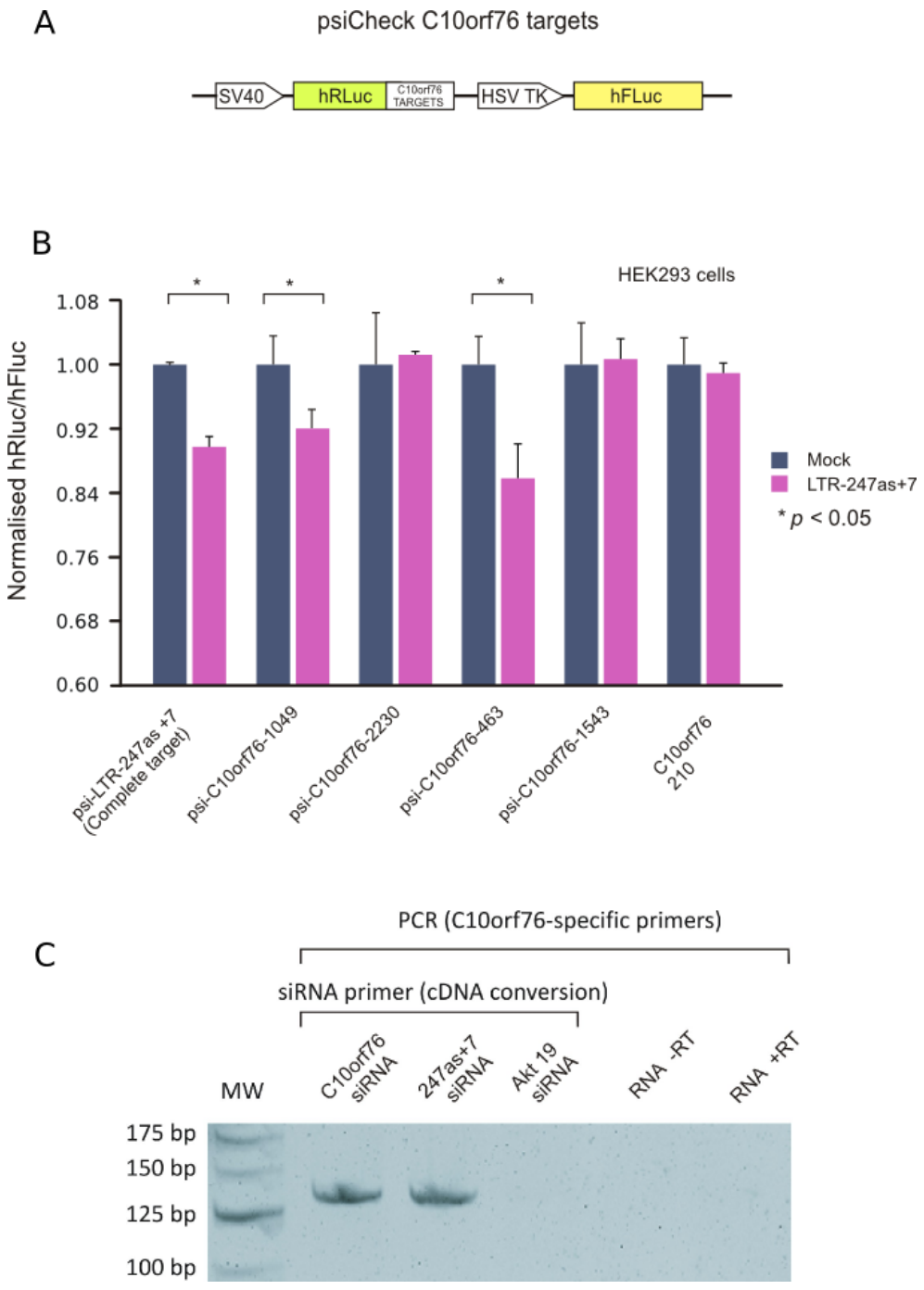


Figure 43. Select putative LTR-247as+7 target sites within C10orf76 are susceptible to LTR247as+7-mediated suppression. A) The complete LTR-247as+7 target site and putative LTR-247as+7 target sites within C10orf76 were inserted downstream of hRluc within the psiCheck2 vector. B) Knockdown efficiency of LTR-247as+7 was determined against each of the C10orf76 targets following transient co-transfections with LTR-247as+7 or a mock (pTZU6+1). Data are expressed as a ratio of hRluc to hFluc normalised to the mock (n = 3, ± SEM). A single-tailed Student's t-test was used to determine statistical significance. C) C10orf76 RNA can be reverse-transcribed using LTR-247as+7-siRNA as a primer. PCR was completed using C10orf76-specific primers on cDNA that had been converted using siRNAs specific to C10orf76 (C10-siRNA-1), LTR-247as+7 or a control (siRNA-AKT19). Additional PCR controls included RNA that had not been reverse transcribed (RNA -RT) and RNA in the presence of RT (RNA +RT) but without primers. The expected PCR product was ~136 bp as compared to a molecular weight ladder (MW).

4.4 Discussion and conclusions

Previous work had identified a TGS-susceptible region, site 247, which when targeted by an asRNA (LTR-247as), was capable of suppressing a HIV-1 subtype B LTR promoter in the presence of Tat (Weinberg, Villeneuve et al. 2006). Indeed, long-term TGS of HIV was induced using HIV-2 delivered asRNAs targeted to site 247 (Turner, De La Cruz et al. 2009). The observed silencing correlated with an increase in silent-state epigenetic marks (including histone and DNA methylation), a loss of NF-κβ recruitment to the viral promoter, and required AGO1, HDAC1 and DNMT3a localisation to the targeted region. Thus it was unexpected that a 22 nt asRNA sequence targeted only seven bases downstream of site 247 (LTR-247as+7), and still largely overlapping with the suppressor LTR-247as, showed the opposite effect of increased gene activation. Intriguingly, transcriptional activation of genes (TGA) in response to synthetic siRNAs or short dsRNAs targeted to the promoters has been described for E-cadherin, $p21^{WAF1/CIP1}$ (p21), VEGF, and progesterone (Li, Okino et al. 2006; Janowski, Younger et al. 2007).

TGA (also referred to as RNA activation or RNAa) appears to be a robust sequence-specific phenomenon supported by the observation that asRNAs are also involved in human genetic diseases (Tufarelli, Stanley et al. 2003) and point to a biological role for short RNAs in the epigenetic control of gene expression in human cells (Han, Kim et al. 2007; Kapranov, Cheng et al. 2007). However, since a mechanistic explanation for RNA activation remains to be elucidated, it was unclear whether the phenomenon observed with asRNA LTR247as+7 represented a unique pathway, or formed part of an existing framework of post-transcriptional inhibitory effects caused by antisense RNAs or RNAi effector sequences. A number of mechanisms exist whereby RNAs can modulate gene activity. For example, gene activation has been observed for 20 bp dsRNAs containing neuron restrictive silencer element (*NRSE/RE1*) sequences. *NRSE* dsRNA activates expression of *NRSE*/*RE1*-containing genes by recognising the NRSF/REST transcriptional regulator to stimulate neural stem cells differentiation (Kuwabara, Hsieh et al. 2004).

The gene activation observed for LTR-247as+7 was probably as a result of a sequence-specific off-target effect. This has been ubiquitously observed when using RNAi and antisense effector sequences to induce post-transcriptional inhibition (Jackson, Bartz et al. 2003). The data presented here suggested that a complex cascade exists which possibly includes several of the genes which have been shown to be deregulated by LTR-247as+7. However, in this chapter attention was focused on a gene of unknown function, C10orf76, which was down-regulated by LTR-247as+7. The link between C10orf76 and LTR247as+7 was determined via an *in silico*

analysis of putative microRNA-like binding sites, two of which were established to be susceptible to post-transcriptional inhibition by LTR-247as+7 (Figures 41 and 42).

Some of the experimental criteria used by Li (Li, Okino et al. 2006) and Janowski (Janowski, Younger et al. 2007) in establishing the validity of RNA-mediated gene activation may also be true for off-target effects. These include the sequence specificity and length characteristics of the targeting asRNA or dsRNA. Firstly, no notable increase in gene expression was observed with LTR-247as-50/+50, which includes the LTR-247as+7 sequence, indicating that the LTR-247as+7 sequence motif alone was not a determinant for off-target gene activation. Moreover, a mutant version of LTR-247as+7, with seed region mismatches, could abrogate activation, similarly to what was observed by others (Li, Okino et al. 2006). Since sequence specificity may not be a determinant for ruling out off-target or RNA-protein interactions, more stringent criteria may be necessary for determining sequence-dependent effects induced by RNA when targeted to promoter sequences. As was shown for antisense RNAs that induce TGS (Weinberg, Villeneuve et al. 2006), direct association of effector RNAs at the targeted promoter would be of value, although this was attempted without success by Janowski *et al* for AGO1 and AGO2 (Janowski, Younger et al. 2007).

Secondly, RNA activation has been associated with the removal of silencing histone methylation marks and/or the addition of activating acetylation and methylation marks (Li, Okino et al. 2006; Janowski, Younger et al. 2007). While some increase in H3K4me2+ elicited by LTR-247as+7 was shown (Figure 37), even greater enrichment was elicited by upstream and downstream overlapping asRNAs, which did not show potent gene activation. It is possible that the off-target inhibition of a repressor may function to induce active chromatin. The data in this chapter suggested that showing active or suppressive chromatin marks was largely correlative, and did not provide a mechanistic basis for the underlying result.

Lastly, both Li and Janowski used siRNAs and not asRNAs to induce gene activation. Li *et al* (Li, Okino et al. 2006) showed that the inhibition of AGO2 negatively affected the ability of siRNAs to inhibit RNA activation. Taken together with the inability to localise AGO2 to the activating promoter (Janowski, Younger et al. 2007), the data are consistent with the possibility that siRNAs were inadvertently inhibiting a transcriptional regulator. In addition, most studies utilising siRNAs to induce TGS made use of very high concentrations of siRNAs in order to direct their nuclear localisation (> 30 nM). However to date, much lower concentrations have been used to elicit RNA-mediated gene activation [1 nM for (Li, Okino et al. 2006) and 12 nM for (Janowski, Younger et al. 2007)].

The mechanism of TGS has been fairly well described [reviewed in (Hawkins and Morris 2008; Morris 2009)] and involves AGO1 (Janowski, Huffman et al. 2006; Kim, Villeneuve et al. 2006) as well as RNA Pol II-expressed promoter-associated RNAs that act as target binding

sites for siRNAs (Han, Kim et al. 2007). Less data is available regarding TGA although promoter-associated RNAs as well as AGO proteins were required to induce TGA at the human progesterone receptor and E-cadherin promoters (Place, Li et al. 2008; Schwartz, Younger et al. 2008). Interestingly, a mechanism for TGA has been proposed involving deregulation of endogenous bidirectional transcription (Morris, Santoso et al. 2008; Morris 2009). The binding of asRNAs to bidirectionally transcribed human gene promoters directs epigenetic silencing protein complexes to the region resulting in TGS. However, if the asRNA is suppressed or prevented from binding to the target promoter, deregulation occurs resulting in a TGA effect. This was suggested as a reason for tumour suppressor gene p21 activation (Morris, Santoso et al. 2008) and may also be true of previously observed TGA effects (Place, Li et al. 2008; Schwartz, Younger et al. 2008).

Overall, the findings of this chapter indicated that targeting promoter regions with short RNAs may have unintended effects, especially if partial homology to transcripts within the treated cells exists. Here it was shown that indiscriminate cell-wide gene activation could be elicited by off-target down-regulation of a suppressor gene in a manner similar to that proposed by Morris *et al* (Morris, Santoso et al. 2008). Just as off-target effects mask the efficacy of siRNAs for post-transcriptional gene knockdown, caution should be taken when RNAs (either siRNAs or asRNAs) are generated to target promoter regions with the intention of inducing specific RNA-mediated transcriptional modulation.

Just over ten years have passed since RNAi was discovered and already RNAi-based gene expression systems are being tested in a clinical setting. Although much has been achieved, the manipulation of the mammalian RNAi pathway for therapeutic purposes requires a more complete understanding of the intricate networks involved to ensure robust, reliable and safe clinical outcomes. The emerging field of combinatorial RNAi has been given a lot of recent attention because it represents a powerful tool for use against highly mutable viruses such as HIV. The main aim of such a combinatorial approach is to prevent viral escape through the simultaneous and constitutive application of numerous guide sequences. The majority of combinatorial RNAi-based systems tried to date have relied on the co-expression of more than one shRNA and a major drawback has been the production of saturating levels of these mimics in mammalian cells (Grimm, Streetz et al. 2006). In addition, HIV rapidly generates resistance in the presence of low levels or too few unique siRNAs rendering the guide sequences ineffective (ter Brake, Konstantinova et al. 2006). The studies presented in this thesis add significantly to the newly emerging body of research on combinatorial RNAi strategies by focusing on two novel technological approaches for expressed combinatorial RNAi mimic systems. Both long-hairpin RNAs and combinatorial polycistronic pri-miRNA systems were generated and shown to produce multiple siRNAs capable of simultaneously targeting several HIV sites. Importantly, as a prelude toward further studies that prove that these combinatorial RNAi modalities provide robust long-term inhibition of HIV, the focus of this thesis was to explore the functional attributes of different mimics to effectively express and produce multiple anti-HIV guides. Solving the problem of viral escape remains a crucial future objective, and while not the immediate focus of this thesis, further discussion is warranted. Furthermore, as HIV is highly mutable it has developed counter-strategies against RNAi that may require the use of additional RNAi target sites and/or RNA-based inhibitory strategies such as TGS to ensure maximum efficacy. Related to this, the data described in Chapter 4 showed the unintended development of significant OTEs following TGS-inducing RNAs and this topic is discussed further in this chapter. Finally, improvements in delivery strategies used to introduce RNAi mimics to HIV-infected cells have enabled rigorous clinical testing but the wide-spread application of some of these strategies remains contentious. Although it is currently unclear how AIDS may be cured by harnessing RNAi, the potential of this gene therapy for countering HIV infection is exciting and the role of combinatorial RNAi is central to this.

5.1 Combinatorial RNAi and viral escape

The combined expression of numerous anti-HIV siRNAs is effective at reducing HIV replication over short periods of time. However, as HIV causes a chronic infection, constitutive long-term inhibition of HIV is required. Importantly, as HIV resistance develops over time, monitoring of RNAi-resistant strains is crucial to the success of these strategies. Early studies by Westerhout *et al* using single anti-HIV siRNAs revealed that there is an inverse correlation between the level of HIV resistance and the stability of the siRNA/target RNA interaction (Westerhout, Ooms et al. 2005). Nine different HIV escape mutants were detected between 27 and 80 days following the introduction of anti-*nef* siRNAs. The majority of the resistance mutations were present in the centre portion of the siRNA guide sequence and included base substitutions or deletions. Two escape mutants that showed particularly high levels of siRNA-*nef* resistance had developed mutations that induced alternative RNA folding of the target mRNA thus negatively affecting siRNA binding. This result showed that the efficacy of RNAi-mediated inhibition depends on siRNA/target RNA binding. While the study by Westerhout *et al* revealed vital information regarding HIV resistance mutations, only single siRNAs targeted to an accessory gene were used. In a subsequent study by ter Brake *et al*, the simultaneous application of four expressed anti-HIV shRNAs prevented viral escape for 40 days (ter Brake, Hooft et al. 2008). However, viral escape mutants were detected following prolonged culturing and sequence analysis revealed point mutations within the central portion of the target sites. Notably, the study by ter Brake *et al* used siRNAs targeted to *gag*, *pol* and the *tat*/*rev* overlap and no deletion mutations were detected. This observation alluded to the notion that the types of HIV resistance mutations are restricted when conserved mRNAs are targeted. This was clearly demonstrated in a thorough investigation by von Eije *et al*. When highly conserved HIV sequences were targeted, escape mutants only showed point mutations and no deletions were detected (von Eije, ter Brake et al. 2008). Similarly to the other studies, the point mutations were all located within the central portions of the target sites and intriguingly, G:U mismatches were frequently selected. These observations were also in line with previous data showing that mutations in the central region of target mRNAs have the greatest effect on RNAi efficacy *in vitro* (Amarzguioui, Holen et al. 2003; Pusch, Boden et al. 2003). Notably, positions 15 and 16 can display discriminatory effects such that a single mismatch in either of these two positions is sufficient to cause mutation-specific target selectivity of mutant alleles only (Schwarz, Ding et al. 2006; Scholefield, Greenberg et al. 2009). Taken together, these observations have important ramifications for the combinatorial RNAi systems described in this thesis as a key element to their overall success lies in target choice. The combinatorial expressed pri-miRNA mimics described in Chapter 3 encoded siRNAs against *int*, *gag*, LTR and *tat* mRNAs. The *int*, *gag* and *tat* target sites were

chosen based on previously published data that was generated using molecular clones of HIV. This latter point is important because HIV molecular clones comprise isogenic sequences that do not reflect the diversity of target sequences present in wild type HIV infection. A better option would have been to select target sites based on multiple HIV sequence alignments as this may have provided more robust siRNAs capable of inhibiting a broader range of cognate mRNAs. This is particularly important if long-term suppression in the absence of viral escape is desired. The abilities of the pri-miRNA mimics to suppress non-isogenic target sequences was assessed using the primary HIV isolate FV5 (Figure 32). However, while potent suppression was noted following transient transfection, data was only collected for seven days following infection and thus potential escape mutants were not detected. The LTR target sites encoded in the lhRNAs described in Chapter 2 were chosen as together they comprised the initial ~180 bp of HIV transcripts. LhRNA LTR 1 overlapped completely with TAR loop and despite mediating significant (albeit low) target inhibition, the strong secondary structure of this site renders it refractory to potent RNAi. As the stability of the siRNA/target RNA interaction has an inverse correlation to the level of HIV resistance (Westerhout, Ooms et al. 2005), long-term suppression mediated by lhRNA LTR 1 would be unlikely. However, understanding the rules of lhRNA processing can lead to improvements in the design and subsequent siRNA/target interactions of these mimics (see section 2.4).

Transcriptional gene silencing provides an alternative mechanism to induce long-term proviral silencing. TGS can be mediated by asRNAs targeted to promoter elements and various sequences, including site 247, within the HIV LTR are susceptible to this (Weinberg, Villeneuve et al. 2006; Suzuki, Juelich et al. 2008; Turner, De La Cruz et al. 2009; Yamagishi, Ishida et al. 2009). As the LTR is highly conserved, selective pressure mediated by silencing RNAs may result in escape mutants with decreased fitness. However, as TGS of the LTR induces changes consistent with latent HIV infection (Williams, Chen et al. 2006) (i.e. basal levels of LTR-mediated transcription), the detection of potential escape mutants and analysis of the types of resistance mutations present remains to be completed. Data on the ability to control the evolution of wild type HIV using an RNA-based strategy would be highly valuable. Furthermore, targeting of the LTR using a combination of TGS and traditional post-transcriptional RNAi, may present a resistance barrier that is too high for HIV to escape and is discussed further in section 5.4.1.

RNAi-mediated inhibition of host cell targets required for HIV infection poses an additional set of sequences that can be incorporated into combinatorial RNAi systems (see section 1.11.3). The number of novel host proteins available for targeting has been boosted by the recent discovery of 200 to 300 new host dependency factors required for HIV replication (Brass, Dykxhoorn et al. 2008; Konig, Zhou et al. 2008; Zhou, Xu et al. 2008; Yeung, Houzet et

al. 2009). Surprisingly little concordance of specific genes was found between the different studies and no single gene was common to all four data sets. The Zhou and Brass papers for example had an overlap of only seventeen genes, while the Brass and König papers had five overlapping genes (Brass, Dykxhoorn et al. 2008; Zhou, Xu et al. 2008). All three of these data sets were generated with different HIV strains, various siRNA design algorithms and unequal efficacy criteria. Should the same stringent criteria as used in the König study be used on the genes identified in the Brass paper, the level of overlapping genes rises to 60 (Konig, Zhou et al. 2008). The Yeung study utilised shRNAs delivered by lentiviral vectors to T cell clones which was a completely different method as compared to the other three studies. Taken together, these discrepancies highlight the importance of thorough experimental design, multi-faceted data validation and the fact that variations in these can account for significant differences between similar experiments.

The impact of these discrepancies is on target choice which ties back to viral escape once again. Assessment of the effect(s) each new HDF has on wild type HIV replication will require large-scale screening and independent validations that may not reach agreement. Thus it was suggested that attention should be paid to the modestly concordant genes only (Konig, Zhou et al. 2008). These include genes coding for nuclear pore complex scaffolds, transcription factors, Golgi-associated proteins and trans-membrane transporters. As these genes are involved in generic cytoplasmic and nuclear transport as well as global gene expression, RNAi-mediated targeting may be detrimental to the host cell as well. However, HIV is less likely to evolve an entirely new set of cellular associations to bypass the requirement for an important host factor. Clearly, the incorporation of HDFs into combinatorial systems would place significant selective pressure on HIV and escape mutations would most likely occur. This remains to be elucidated and should receive ample attention.

5.2 HIV-mediated modulation of the RNAi pathway

HIV-1 is able to escape the RNAi pathway via a number of routes (see section 1.11.2) but the virus also demonstrates the capacity to modulate RNAi through specific protein suppressors [reviewed in (Corbeau 2008)]. Understanding these interactions may shed light on how combinatorial RNAi effectors can be used to reduce or even prevent HIV's modulation of the pathway. Viral Tat can directly inhibit Dicer processing of siRNAs *in vitro* via the interaction of Tat and TRBP (Bennasser, Le et al. 2005; Bennasser, Yeung et al. 2006). Functional Dicer requires TRBP to provide a stable link between itself and AGO2 from RISC (Chendrimada, Gregory et al. 2005) but TRBP also specifically binds the HIV-1 TAR loop in order to increase Tat-mediated viral expression [(Gatignol, Lainé et al. 2005) and see section 1.4.4]. Recent data

indicated that sequestration of TRBP by interactions with the HIV-1 TAR loop, inhibits the action of Dicer thereby reducing the overall effectiveness of RNAi (Bennasser, Yeung et al. 2006). Moreover, TRBP shares a similar domain (dsRBD3) with the dsRNA binding protein PACT that forms part of the Dicer/TRBP/AGO2 complex (Lee, Hur et al. 2006). This means the HIV-1 TAR loop could bind both PACT and TRBP, either of which action would negatively impact Dicer (Bennasser, Yeung et al. 2006; Lee, Hur et al. 2006). Inhibition of the HIV TAR loop and/or Tat would dramatically reduce these interactions thereby ensuring Dicer functioning. As the TAR loop was previously been shown to be refractive to siRNA-mediated inhibition, it was surprising to see significant (albeit low) levels of suppression mediated by lhRNA LTR 1 (Figure 12). However, low level inhibition of TAR may be sufficient to hinder TRBP sequestration, especially if combined with additional potent siRNAs. Alternatively, HIV transcription could be inhibited *before* threshold levels of Tat have been produced. The ability of lhRNAs LTR 2 and 3 to inhibit non-processive LTR-driven transcription provides an exciting observation that RNAi-based tools can reduce early-stage HIV replication and this area should be fully explored.

Another RNAi suppressor mechanism exhibited by HIV-1 has been linked to viral latency and miRNA regulatory loops. Both Drosha and Dicer inhibit early stage HIV-1 replication in PBMCs and latently infected cells taken from HIV-1 positive donors [(Triboulet, Mari et al. 2007) and see section 1.5]. The mechanism is incomplete but a miRNA cluster that affects a Tat cofactor, P300/CBP-associated factor (PCAF) is involved. Two members of the cluster, miR17-5p and miR20a, down-regulate PCAF mRNAs causing a negative impact on HIV replication although the virus can suppress miR17-5p/miR20a expression to complete the feedback loop. HIV *nef* encodes a miRNA (miR-N367) that down-regulates transcription of viral mRNAs through the negative response element (NRE) of the LTR U3 region (Omoto, Ito et al. 2004; Omoto and Fujii 2005). Two human-encoded miRNAs (hsa-miR29a and hsa-miR29b) detected in PBMCs were shown to inhibit Nef expression and HIV replication *in vitro* (Ahluwalia, Khan et al. 2008) thereby completing another loop. Although the exact mechanism is unclear, HIV further manipulates RNAi to induce proviral latency by generating a TAR-derived miRNA that degrades viral transcripts in HIV infected cells and primary T cell precursors (Klase, Kale et al. 2007). Conventional drugs such as histone deacetylase (HDAC) inhibitors have been used to prevent HIV latency but the generic action of these compounds causes global HDAC inhibition which is toxic to the host cells. Recent data revealed that Class II HDAC inhibitors can selectively repress histones located in the LTR of integrated HIV but further development is required before these drugs become part of a treatment strategy (Archin, Keedy et al. 2009). The use of gene therapy-based alternatives such as the siRNAs targeted to site 247 in the HIV LTR (see Chapter 4) provide a ready approach to limit the viral/host interactions described above as well as prevent HIV replication.

The ability to modulate RNAi is not exclusive to HIV and many examples exist of pathogenic viruses that encode modulators of the pathway (Manjunath, Kumar et al. 2006; Haasnoot, de Vries et al. 2007). An interesting question is why these strategies evolved. Plants use the RNAi pathway as an antiviral mechanism and the presence of countless RNAi suppressor proteins encoded in plant viruses is evidence of this (Haasnoot, Westerhout et al. 2007). Due to the increased complexity of vertebrates, antiviral mechanisms here are the responsibility of immune systems. Why then are there consistent examples of rapidly evolving vertebrate viruses that conserve specific anti-RNAi mechanisms if the pathway is not primarily antiviral in nature? An interesting suggestion, put forward in the context of HCV infection, is that viruses manipulate the RNAi pathway as a means to ensure viral persistence in the face of elimination (Mahajan, Drake et al. 2008). HCV uses a low fidelity polymerase for replication thereby generating heterogenous virions during infection, yet two host miRNA target sequences within its genome are maintained across all six genotypes. Upon infection HCV induces a potent IFN response that leads to sterilising immunity. To avoid immediate annihilation, HCV seems to have evolved a mechanism whereby host anti-HCV miRNAs suppress viral expression (at the two conserved sites) to levels that do not induce IFN but allow for persistent low level infection (Mahajan, Drake et al. 2008). This strategy may also be true of HIV and adds another hurdle to anti-HIV RNAi.

5.3 Unintended off-target effects related to RNAi

Although RNAi mimics, both synthetic and expressed, are highly effective inducers of RNAi, toxicities related to their sequences, structures or processing as well as possible OTEs remain a problem. The Pol III-expressed lhRNAs presented in this thesis did not induce cellular immunostimulation and there are several explanations for this. Firstly, the characteristic 2 nt 3' overhangs present in the anti-LTR lhRNAs likely prevented these mimics from being detected by the mammalian pathogen recognition receptors retinoic acid helicase type I (RIG-I) and protein kinase R (PKR). RIG-I plays a central role in the discrimination between self and non-self cytoplasmic RNAs by eliciting a downstream cascade of cytokines including type I interferons that in turn up-regulate expression of IFN-inducible genes and affect protein production, cell cycle growth and apoptosis [reviewed in (Kawai and Akira 2008)]. Blunt-ended dsRNAs and those with 5' overhangs complex stably with RIG-I in the presence of ATP, causing downstream binding of IFN signalling molecules. In contrast 3' overhanging dsRNAs do not bind RIG-I tightly and are released (Takahasi, Yoneyama et al. 2008; Schlee, Roth et al. 2009). PKR also binds dsRNAs but requires the presence of a 5' triphosphate (Sledz, Holko et al. 2003; Kim, Longo et al. 2004) to cause phosphorylation of eukaryotic translation initiation factor 2 (eIF2) and

subsequent cellular apoptosis [(McAllister and Samuel 2009) and reviewed in (Borden, Sen et al. 2007)]. As the anti-LTR lhRNAs described in this thesis were expressed from a Pol III (U6) promoter, the 5' ends do not have triphosphates and the sequence design ensured that the overall hairpin folded into a structure with a 2 nt 3' overhang.

The mammalian IFN cascade is not only elicited via the recognition of dsRNA ends. Indeed, regardless of the presence of 3' overhangs, longer siRNAs (> 33 bp) have been shown to elicit an IFN response (Puthenveetil, Whitby et al. 2006; Nallagatla and Bevilacqua 2008). Investigations revealed that PKR has multiple dsRNA binding domains and that numerous PKR complexes can assemble on the same siRNA duplex *in vitro* (Puthenveetil, Whitby et al. 2006). Importantly, modification of the sense strand of the siRNA duplex inhibited PKR binding but did not decrease the potency of the siRNAs (Nallagatla and Bevilacqua 2008). This suggests that the G:U mismatches included within the sense strands of the anti-LTR lhRNAs described in Chapter 2 may play a role in decreasing the immunostimulatory potential of these constructs. Several groups have reported potent target-specific inhibition using long dsRNAs (38 to 448 bp) without G:U mismatches but these constructs did not contain hairpin loops, and were either chemically synthesised (Akimov, Kabilova et al. 2009; Chang, Kang et al. 2009) or transcribed from T7 promoters. This is in direct contrast earlier reports (Elbashir, Harborth et al. 2001; Akashi, Miyagishi et al. 2005; Akimov, Kabilova et al. 2009) as well as the current body of work but the reasons for the discrepancies remain unclear.

It has been suggested that certain sequence motifs can elicit an IFN response, although this usually occurs when viral nucleic acids interact with toll-like receptors (TLRs) located on endosomes [(Forsbach, Nemorin et al. 2008) and reviewed in (Montero Vega and de Andrés Martín 2008)]. Importantly, because expressed constructs (such as the anti-LTR lhRNAs and pri-miRNA mimics described in this work) are introduced cytoplasmically, they do not encounter TLRs. Of greater concern, is the capacity for expressed mimics to overwhelm components of the endogenous RNAi pathway. Acute liver toxicity and mortality in mice following expressed shRNA administration (Grimm, Streetz et al. 2006) as well as toxic effects in the brains of mice treated with expressed shRNAs targeted to Huntington's Disease-specific mRNAs (McBride, Boudreau et al. 2008) has been observed. Accumulation of shRNA precursors in the livers of mice suggested that the observed toxicity was due to saturation of the export machinery (Exp5) (Grimm, Streetz et al. 2006). This was supported by the observation that a decrease in shRNA expression reduced liver damage although more recent data suggested that TRBP and AGO2 could also be the bottleneck (Castanotto, Sakurai et al. 2007). Lowering shRNA expression levels did not abrogate toxicity in the brains of mice and poor detection of unprocessed mimics in those samples suggested that export machinery saturation was not responsible for the adverse

effects observed. Alternatively OTEs may have been involved as all three shRNAs tested, regardless of sequence, elicited toxicities (McBride, Boudreau et al. 2008).

Since siRNA-mediated activity requires a match of only six nucleotides (seed region), at least one hit can be expected by random chance every 16 384 nucleotides within a stretch of DNA or RNA (Jackson, Burchard et al. 2006; Mahajan, Drake et al. 2008) making OTEs statistically inevitable. The data in Chapter 4 describes an unintended yet significant OTE that occurred as a result of an asRNA targeted to a single site within the promoter region of HIV. While TGS was the intended outcome, the opposite result occurred and involved a complex cascade of interactions including a gene of unknown function. This result emphasises the importance of using potent RNAi effector sequences that are effective at low concentrations. One option to address this is to express RNAi mimics of choice from less active promoters such as H1 or tRNA. Transcription from these Pol III promoters generates RNAi mimics with characteristic 5' and 3' ends but at low concentrations that do not induce cellular immunostimulation (Scherer, Frank et al. 2007). Furthermore, as $tRNA^{LYS3}$ and $tRNA^{VAL}$ Pol III promoter transcripts are processed in the nucleus by tRNA processing enzymes prior to export, the transport of these shRNAs may occur via an alternative exportin pathway (Kawasaki and Taira 2003; Rodriguez, Dargemont et al. 2004). Another option that would also allow for greater flexibility in terms of expression is the use of Pol II promoters. These latter promoters provide inducible and tissue-specific expression but as they do not generate transcripts with defined 5' and 3' ends, pri-miRNAs would be the mimic of choice. Interestingly, expressed pri-miRNA scaffold versions of the same shRNAs used in the McBride *et al* mouse brain study did not induce adverse effects and showed similar silencing profiles despite being expressed at lower levels (McBride, Boudreau et al. 2008). Similar data was described for expressed pri-miRNA mimics targeted to HBV in the livers of mice (Ely, Naidoo et al. 2009). Furthermore, the exogenous constructs did not affect the expression of an endogenous liver-specific pri-miRNA suggesting that the endogenous miRNA pathway was not disrupted following the introduction of these constructs into liver cells.

Regardless of promoter choice, it remains crucial that each guide sequence within a combinatorial RNAi system is expressed at high enough doses to induce potent target suppression. The observation that the multimeric Pol II-expressed anti-HIV pri-miRNA mimics were highly efficacious against wild type HIV yet almost no guide sequences were detected (Figures 31 and 32) suggested that these constructs were effective at very low concentrations. In contrast, the anti-LTR lhRNAs were less efficacious against wild type HIV despite readily detectable guide sequences. This latter observation was most likely related to the poor Dicer-mediated processing of these constructs. The tiling arrays described in Chapter 2 demonstrated that functional guides were only being processed from the first position relative to the 3'

overhang of the lhRNAs (Figures 16 and 17). Thus, these mimics were not functioning as a combinatorial system. The design and generation of expressed lhRNAs capable of producing functional siRNAs from two to three positions along the duplex has been described (see section 2.4). These latter mimics are transcribed off a single Pol III promoter (H1 or U6) with two similarly expressed lhRNA constructs introduced into cells as a combinatorial system (Saayman, Barichievy et al. 2008; Sano, Li et al. 2008; Liu, von Eije et al. 2009). While this approach has been shown to induce target specific inhibition for up to 49 days, the use of multiple Pol III promoters once again raises the concern of Exp5 saturation (although TRBP or AGO2 may also play a role in saturation). One option would be to vary the type of Pol III promoter used (to H1 or tRNA for example) or to utilise a completely alternative method of synthetic RNAi mimics such as that described by Kumar *et al* (Kumar, Ban et al. 2008). A CD7-specific single chain antibody conjugated to an oligo-9-arginine peptide was used to deliver peptide-bound anti-HIV siRNAs directly into T cells of NOD/SCID/IL2rγ^{null} mice (see section 1.12) prior to challenge with HIV. The delivery method was highly specific and Dicer was able to efficiently cleave the siRNAs in the absence of an IFN response. Subsequent viral replication was controlled and the disease-associated CD4 T cell loss was dramatically reduced despite the use of a single siRNA (anti-CCR5). Notably, this strategy provided a pre-infection or prophylactic approach that would be as valuable as a post-infection treatment. However, as with the latter option, combinations of discrete siRNAs are needed and peptide-bound lhRNAs could be explored in this context.

In summary, cellular immunostimulation and unintended OTEs can be avoided through the use of expressed RNAi mimics that are transcribed from less active promoters or synthetic mimics introduced directly to the target site(s) of choice. Pri-miRNA mimics may be better suited to this approach but additional data is still required to assess whether or not these mimics can induce an IFN response. Expressed lhRNAs will become more applicable if their relative expression levels can be controlled and their design improved such that Dicer efficiently processes each siRNA within the mimic.

5.4 Into the future: delivery of RNAi mimics to patients

5.4.1 Vectors for the delivery of RNAi mimics

For RNAi to be effective against chronic viral infections such as HIV, constitutive expression of combinatorial mimics in the target cells of interest is required. An elegant means to achieve this involves delivery of the RNAi constructs to host cells using various vector-based methods. Therapeutic delivery of RNAi mimics presents a significant challenge in the context of HIV as many cell types are infected by the virus and latent reservoirs of viral pools exist in discrete,

inaccessible locations such as the brain. Historically retroviral vectors have been used to deliver genetic payloads with reports dating back to the early 1980s. Within the retrovirus family, lentiviruses have emerged as the most promising candidate largely due to their ability to transduce non-dividing cells and integrate within introns of active transcriptional sites [reviewed in (Quinonez and Sutton 2002)]. Since 1996 when an HIV-1 based vector was compared to a Murine Leukaemia Virus (MLV) counterpart and only the former was shown to stably transduce terminally differentiated rat neurons (Naldini, Blomer et al. 1996; Naldini, Blomer et al. 1996), various other lentiviral strains have been tested as gene delivery vehicles. Feline Immunodeficiency Virus (FIV), HIV-2, Bovine Immunodeficiency Virus (BIV) as well as Equine Infectious Anaemia Virus (EIAV) have all been shown to transduce various cell types and express their transgene payloads (Quinonez and Sutton 2002). The introduction of pseudotyping allowed these vectors to target specific cell types, with the use of the Vesicular Stomatitis Virus Glycoprotein (VSV-G) extending the host range of pseudotyped viruses extensively (Sutton 2000). However, due to the integrating nature of lentiviral vectors, they are potentially oncogenic and recombination with wild type virus already present in cells presents another safety concern. Various strategies have been employed to improve the overall safety of lentiviral vectors such as the removal of the 5' LTR and accessory genes (Morris and Rossi 2006). This has also provided space within the genome to insert genes of interest or RNAi mimic expression cassettes but an upper size limit exists related to the packaging of lentiviruses.

One particular lentiviral vector system, termed a conditionally replicating HIV lentiviral vector (crHIV) has provided something of a ‘Trojan horse’ approach to dealing with HIV infection because crHIV vectors can carry an anti-HIV payload as well as compete with natural viral RNAs in already infected host cells (Sachdeva, D’Costa et al. 2007). crHIV vectors carrying a triple combination ribozyme targeted to the U5 portion of the LTR, significantly reduced wild type HIV replication and the infection of neighbouring cells (Dropulic, Hermankova et al. 1996). The effect occurred for 22 days which may be a sufficiently long time period as infected CD4+ T cells display a short half life *in vivo*. Interestingly, cultures that contained both wild type HIV and crHIV vectors carrying only single or double ribozymes showed viral escape, emphasising once again the importance of targeting multiple sites. The ability of crHIV vectors to spread to uninfected CD4 cells (i.e. vector mobilisation) has also been addressed. PBMCs infected with HIV-1 were readily transduced with crHIV-2 vectors carrying up to four anti-HIV ribozyme sequences and while the antiviral effects of the vector could be passed on to neighbouring cells, the subsequent titres were lower as compared to the initial crHIV-2 titre (Morris and Looney 2005). These data suggested that crHIV-based vectors with maximum potency are needed due to the decrease in vector mobilisation following the initial infection. Notably, the asRNA (LTR-247as) described in Chapter 4 was successfully delivered to an HIV reporter cell line using a HIV-2-based

conditionally replicating vector and TGS was observed for a minimum of one month (Turner, De La Cruz et al. 2009). Addition of the HIV-2 vector alone (devoid of LTR-247 asRNAs) also mediated some HIV inhibition suggesting that the vector was competing with wild type HIV for proteins involved in replication. This is not a new observation (Dropulic, Hermankova et al. 1996; Corbeau and Wong-Staal 1998; An, Morizono et al. 1999) but the effect(s) of 'parasitic' delivery lentivectors on the evolution of escape mutants remains to be determined.

There are a number of concerns related to lentiviral vectors including the production of clinically relevant titres and the naturally high recombination rate of HIV that may lead to the deletion of multiple RNAi expression cassettes [reviewed in (ter Brake and Berkhout 2007)]. Related to the first point, lentiviral vector titres can be significantly lowered due to self-targeting when they carry RNAi mimic payloads complementary to lentiviral genes (Banerjea, Li et al. 2003; Chang, Liu et al. 2005). However, significant advances have been made such as using human codon-optimised versions of the target gene or selecting target sequences not present in the lentiviral genome (Westerhout, Ooms et al. 2005; Westerhout and Berkhout 2007) and numerous examples have been published describing the successful lentiviral-mediated delivery of shRNAs (Chang, Liu et al. 2005; Lee 2005; Manjunath, Kumar et al. 2006; von Eije, ter Brake et al. 2008). The second concern relates to the recombination prone nature of HIV. During transduction, repeat sequences such as multiple copies of the same promoter each driving expression of a single shRNA, may be recombined resulting in the loss of the RNAi cassettes (An and Telesnitsky 2001; ter Brake, Konstantinova et al. 2006). This can be overcome by using different promoter sequences or a different RNAi mimic such as those described in Chapter 3 or elsewhere (Liu, Haasnoot et al. 2007; Weinberg, Ely et al. 2007; Ely, Naidoo et al. 2008; ter Brake, Hooft et al. 2008). Indeed, a central reason for exploring the ability of four different pri-miRNA scaffolds to produce highly efficacious anti-HIV guides was to incorporate them into a single expressed system that would be averse to recombination events (Chapter 3). Importantly though, the use of pri-miRNA mimics introduces an additional problem as Drosha cleavage sites within the construct can reduce vector titres due to destruction of genomic DNA when the pri-miRNA mimic is processed (Berkhout and ter Brake 2009). To date, only one publication has described lentiviral packaging of a combinatorial pri-miRNA construct targeted to HIV (Liu, Haasnoot et al. 2008). While there was no difference in cell viability between transduced and un-transduced SupT1 cells, no data was included on the effect(s) of pri-miRNA packaging on vector titres. Uncovering ways to enhance vector titre production will further boost the use of combinatorial pri-miRNA mimics in a clinical setting.

5.4.2 Clinical studies in human patients

The details of some anti-HIV RNA-based clinical trials have already been discussed (see section 1.13) but the wide-spread application of successful gene therapies in human patients remains a contentious topic of discussion. Two major approaches are theoretically feasible: T cell and haematopoietic stem cell gene therapies. In the former, an entire T cell population or sub-sets thereof are manipulated and grown *in vivo* allowing for HIV-resistant cells to be expanded (Levine, Mosca et al. 1996; Levine, Bernstein et al. 1997). These cells are reintroduced into patients and while the safety and persistence data are promising, either no or modest effects on viral load have been reported (Rossi, June et al. 2007). CD4, CD8 and T cell receptors have all been the foundation of various T cell therapies and are an attractive treatment as the impact on the patients is minimal. The best way to measure this is to discontinue HAART and monitor the ability of engineered T cells to control viral load and CD4 decline. However, as treatment interruption schemes are detrimental to patients this is not advised (Abbas and Mellors 2002). The second gene therapy option uses haematopoietic stem cells and is particularly attractive as these cells produce all the subsequent cells infected by HIV (CD4+ T cells, macrophages, microglia and dendritic cells). Genetic manipulation of these cells to produce an HIV resistant precursor could theoretically provide protection to the entire spectrum of HIV susceptible cells. Haematopoietic precursors are found in bone marrow and umbilical chord blood and can be isolated by their CD34+ cell-surface marker (Sutherland, Keeney et al. 2003). Contrary to T cells, CD34+ cells cannot be cultured *in vivo* but rather require an *ex vivo* approach and subsequently, lower numbers of genetically modified haematopoietic cells are generated. In addition, the reintroduction of modified CD34+ cells usually requires complete bone marrow ablation of the patient but this is highly toxic and requires intensive patient care. Several clinical trials (including the triple combination therapy against HIV described in section 1.13) are underway to explore the feasibility of this therapeutic approach but early data suggested that low numbers of peripheral blood cells were carrying the transgene indicating poor engraftment of the haematopoietic precursors (Amado, Mitsuyasu et al. 2004; Podsakoff, Engel et al. 2005).

Clearly this approach is expensive, requires intensive care and the treatment regimen is highly invasive and toxic to the patient. Furthermore, while effective on a small scale, this therapy is not feasible in Africa where the number of patients would overwhelm the health care system. In addition, HIV reservoirs that allow for replication competent forms of the virus to accumulate and persist remain inaccessible to treatment (Blankson, Persaud et al. 2002). Reservoirs include quiescent CD4+ T cells, lymphoid organs, the genitourinary tract and the central nervous system and can persist for the lifetime of the patient despite the use of HAART (Finzi, Blankson et al. 1999). Thus for an RNAi-based gene therapy approach to be successful,

not only do the RNAi mimics need to be highly efficacious yet non-toxic, they need to be delivered to all target sites with a vector capable of transducing actively infected cells as well as reservoir sites using a methodology that is feasible and not too demanding on the patient. This will require a concerted multi-disciplinary approach but as more data are published, the techniques will be refined and improved. Combinatorial RNAi-based approaches such as those described in this thesis could form part of either T cell or HSC strategies. Combinatorial pri-miRNA mimics need to be delivered to the nucleus of cells to ensure Drosha-mediated cleavage and lentiviral vectors could be used. Expressed lhRNAs could be delivered by lentiviral vectors but as they only require Dicer-mediated cleavage prior to RISC loading, they could be also delivered by the method of Kumar *et al*. Given the incredible success of HAART, perhaps the best option would be to combine RNAi therapy with current antiretrovirals using an approach that targets infected *and* susceptible cells.

5.5 Concluding remarks

Evidence from protein conservation across plants and animals suggests that the last common ancestor of eukaryotes had siRNA-based mechanisms of RNA regulation and that the modern RNAi pathway is an amalgamation of ancestral archaeal, bacterial and bacteriophage proteins required for DNA repair and RNA processing (Shabalina and Koonin 2008). The information related to RNAi in mammalian biology is continually increasing as more details about the various molecular interactions are uncovered and intersecting pathways are described. As our understanding of the RNAi pathway is extended, the ability to manipulate the mechanisms involved expands. By exploiting the exquisite sensitivity and potency of the pathway, exogenous RNAi mimics are often far more efficient than other nucleic acid based strategies and as such, the field of antiviral RNAi has received extensive attention [reviewed in (Devincenzo 2009; von Eije and Berkhout 2009)]. Various stages of clinical trial are underway in humans to assess the safety and efficacies of numerous RNAi mimics targeted against a range of viruses. One of the most prolific viruses worldwide is HIV and it has been on the receiving end of many RNAi-based therapeutic strategies with the main reason for failure lying in the ability of HIV to rapidly mutate and thus escape suppression. HAART has been incredibly successful at decreasing HIV-related morbidity and mortality with the chief reason for success lying in the combinatorial approach whereby a minimum of three anti-HIV drugs are simultaneously administered. Unfortunately, treatment interruption leads to resistance and many factors including poor patient compliance due to drug toxicity, intermittent supply of drugs in resource-strained areas or lack of patient education can lead to sub-optimal therapies. Thus there is a need for alternatives and RNAi provides an emerging and enthralling new gene therapy approach. Akin to HAART, RNAi needs

to be applied in combination for the maximum efficacy to be achieved but to date few successful combinatorial strategies have been explored.

In the current body of work presented here, two new discrete methods of combinatorial RNAi, namely long hairpin RNAs and multimeric pri-miRNA mimics, were investigated. Long hairpins can be expressed off a single promoter but can be concurrently processed into numerous active siRNAs thereby providing a combinatorial RNAi option that won't lead to unintended promoter occlusion and related toxicity. The transcribed region of the HIV 5' LTR promoter was shown to be effectively suppressed using lhRNAs that were non-toxic in mammalian cells. The portions of the lhRNA duplex that produced the most active siRNAs were dissected out and inhibition of an HIV infectious molecular clone as well as a wild type HIV isolate was demonstrated. Importantly, this work laid the foundation for subsequent studies that went on to improve the design of lhRNAs such that highly active siRNAs could be produced from all positions along the stem of the duplex (Saayman, Barichievy et al. 2008; Liu, von Eije et al. 2009). The second major combinatorial RNAi strategy explored was the use of multimeric pri-miRNA mimics. The ability of various endogenous pri-miRNA backbones to generate active guide sequences complementary to various highly susceptible HIV target sites was analysed and it was shown that the backbone affects guide sequence efficacy. A strategy to combine four of the most effective pri-miRNA scaffolds into one expression cassette was developed and significant inhibition of an HIV infectious molecular clone as well as a wild type HIV isolate was demonstrated. Furthermore, it was shown that the suppression mediated by each of the guide strands within the multimeric vector was target specific and that the constructs were non-toxic in mammalian cells. Finally, indiscriminate cell-wide gene activation was shown to occur as an unintended off-target effect following the introduction of short RNAs with partial homology to non-target RNAs. These observations demonstrated that caution should be exercised when targeting untranscribed regions of promoters such as the HIV LTR. Overall, new and exciting data describing the antiviral activities of two novel combinatorial RNAi approaches targeted to HIV as well as valuable data related to the application of RNAi effectors was generated.

These data have laid vital groundwork for what will surely become an important research area as antiviral RNAi moves into the clinical arena. The ability to exploit and tailor-make RNAi mimics allows them to be widely applied to a vast number of viral pathogens and future experiments will need to build on previous work if laboratory applications are to reach patients timeously. Given that the development of a successful preventative vaccine for HIV may continue to frustrate the scientific community, therapeutics will become even more important. Research in the coming years will need to focus not only on combining HAART with new gene based therapies such as RNAi, but also on devising methods to introduce antiviral genetic modalities in a manner that is suitable to infected patients in the developed *and* developing

world. RNAi was discovered a mere decade ago and yet patients are already enjoying the benefits of efforts to exploit the pathway. By furthering the field through the development of combinatorial RNAi, we have contributed to the most exciting new therapy in the 21st century.

Aagaard, L., J. Zhang, et al. (2008). "Engineering and optimization of the miR-106b cluster for ectopic expression of multiplexed anti-HIV RNAs." Gene Therapy: 1-14.

Abbas, U. L. and J. W. Mellors (2002). "Interruption of antiretroviral therapy to augment immune control of chronic HIV-1 infection: risk without reward." Proceedings of the National Academy of Sciences. USA **99**(21): 13377-13378.

Agrawal, L., C. R. Maxwell, et al. (2009). "Complexity in human immunodeficiency virus type 1 (HIV-1) co-receptor usage: roles of CCR3 and CCR5 in HIV-1 infection of monocyte-derived macrophages and brain microglia." Journal of General Virology **90**: 710-722.

Aguilar-Cordova, E., J. Chinen, et al. (1994). "A sensitive reporter cell line for HIV-1 tat activity, HIV-1 inhibitors, and T cell activation effects." AIDS Research and Human Retroviruses **10**(3): 295-301.

Ahluwalia, J. K., S. Z. Khan, et al. (2008). "Human cellular microRNA hsa-miR-29a interferes with viral nef protein expression and HIV-1 replication." Retrovirology **5**(117): 1-25.

Akashi, H., M. Miyagishi, et al. (2005). "Escape from the interferon response associated with RNA interference using vectors that encode long modified hairpin-RNA." Molecular BioSystems **1**: 382-390.

Akimov, I. A., T. O. Kabilova, et al. (2009). "Inhibition of human cancer-cell proliferation by long double-stranded RNAs." Oligonucleotides **19**(1): 31-40.

Akkina, R. K., J. D. Rosenblatt, et al. (1994). "Modeling human lymphoid precursor cell gene therapy in the SCID-hu mouse." Blood **84**(5): 1393-1398.

Amado, R. G., R. T. Mitsuyasu, et al. (2004). "Anti-human immunodeficiency virus hematopoietic progenitor cell-delivered ribozyme in a phase I study: myeloid and lymphoid reconstitution in human immunodeficiency virus type-1–infected patients." Human Gene Therapy **15**: 251-262.

Amarzguioui, M., T. Holen, et al. (2003). "Tolerance for mutations and chemical modifications in a siRNA." Nucleic Acids Research **31**(2): 589-595.

Ameres, S. L., J. Martinez, et al. (2007). "Molecular basis for target RNA recognition and cleavage by human RISC." Cell Cycle **130**: 101-112.

An, D. S., R. E. Donahue, et al. (2007). "Stable reduction of CCR5 by RNAi through hematopoietic stem cell transplant in non-human primates." Proceedings of the National Academy of Sciences. USA **104**(32): 13110-13115.

An, D. S., K. Morizono, et al. (1999). "An inducible human immunodeficiency virus type 1 (HIV-1) vector which effectively suppresses HIV-1 replication." Journal of Virology **73**(9): 7671-7677.

An, W. and A. Telesnitsky (2001). "Frequency of direct repeat deletion in a human immunodeficiency virus type 1 vector during reverse transcription in human cells." Virology **286**(2): 475-482.

Andersen, J. L., E. Le Rouzic, et al. (2008). "HIV-1 Vpr: mechanisms of G2 arrest and apoptosis." Experimental and Molecular Pathology **85**(1): 2-10.

Anderson, J., A. Banerjea, et al. (2003). "Bispecific short hairpin siRNA constructs targeted to CD4, CXCR4 and CCR5 confer HIV-1 resistance." Oligonucleotides **13**: 303-312.

Anderson, J., A. Banerjea, et al. (2003). "Potent suppression of HIV type 1 infection by a short hairpin anti-CXCR4 siRNA." AIDS Research and Human Retroviruses **19**(8): 699-706.

Anderson, J., M. Li, et al. (2007). "Safety and efficacy of a lentiviral vector containing three anti-HIV genes—CCR5 ribozyme, Tat-Rev siRNA, and TAR decoy—in SCID-hu mouse–derived T cells." Molecular Therapy **15**(6): 1182-1188.

Archin, N. M., K. S. Keedy, et al. (2009). "Expression of latent human immunodeficiency type 1 is induced by novel and selective histone deacetylase inhibitors." AIDS **23**(14): 1799-1806.

Asparuhova, M. B., I. Barde, et al. (2008). "Development and characterization of a triple combination gene therapy vector inhibiting HIV-1 multiplication." The Journal of Gene Medicine **10**(10): 1059-1070.

Azuma-Mukai, A., H. Oguri, et al. (2008). "Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing." Proceedings of the National Academy of Sciences. USA **105**(23): 7964-7969.

Bai, J., S. Gorantla, et al. (2000). "Characterisation of anti-CCR5 ribozyme-transduced CD34+ hematopoeticprogenitor cells *in vitro* and in SCID-hu mouse model *in vivo*." Molecular Therapy **1**(3): 244-254.

Bai, J., J. J. Rossi, et al. (2001). "Multivalent anti-CCR5 ribozymes for stem cell-based HIV type 1 gene therapy." AIDS Research and Human Retroviruses **17**(5): 385-399.

Banerjea, A., M. Li, et al. (2003). "Inhibition of HIV-1 by lentiviral vector-transduced siRNAs in T lymphocytes differentiated in SCID-hu mice and CD34+ progenitor cell-derived macrophages." Molecular Therapy **8**(1): 62-71.

Banerjea, A., M. Li, et al. (2004). "Lentiviral transduction of Tar decoy and CCR5 ribozyme into CD34+ progenitor cells and derivation of HIV-1 resistant T cells and macrophages." AIDS Research and Therapy **1**(2): 1-11.

Barboric, M. and B. M. Peterlin (2005). "A new paradigm in eukaryotic biology: HIV Tat and the control of transcriptional elongation." PLoS Biology **3**(2): 0200-0203.

Barnor, J. S., N. Miyano-Kurosaki, et al. (2005). "Lentiviral-mediated delivery of combined HIV-1 decoy TAR and *vif* siRNA as a single RNA molecule that cleaves to inhibit HIV-1 in transduced cells." Nucleosides, Nucleotides, and Nucleic Acids **24**(5-7): 431-434.

Barnor, J. S., N. Miyano-Kurosaki, et al. (2004). "Intracellular expression of antisense RNA transcripts complementary to the human immunodeficiency virus type-1 vif gene inhibits viral replication in infected T-lymphoblastoid cells." Biochemical and Biophysical Research Communications **320**: 544-550.

Bartel, D. P. and C. Chen (2004). "Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs." Science **5**: 396-400.

Bartlett, D. W. and M. E. Davis (2006). "Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging." Nucleic Acids Research **34**: 322-333.

Behlke, M. A. (2008). "Chemical modification of siRNAs for *in vivo* use." Oligonucleotides **18**: 305-319.

Benkirane, M., R. F. Chun, et al. (1998). "Activation of integrated provirus requires histone acetyltransferase." The Journal of Biological Chemistry **273**(38): 24898-24905.

Bennasser, Y., S. Y. Le, et al. (2005). "Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing." Immunity **22**: 607-619.

Bennasser, Y., M. L. Yeung, et al. (2006). "HIV-1 TAR RNA subverts RNA interference in transfected cells through sequestration of TAR RNA-binding protein, TRBP." The Journal of Biological Chemistry **281**(38): 27674-27678.

Bera, S., K. K. Pandey, et al. (2009). "Molecular Interactions between HIV-1 integrase and the two viral DNA ends within the synaptic complex that mediates concerted integration." Journal of Molecular Biology **389**(1): 183-198.

Berezikov, E., W. Chung, et al. (2007). "Mammalian mirtron genes." Molecular Cell **28**: 328-336.

Berkhout, B. and J. Haasnoot (2006). "The interplay between virus infection and the cellular RNA interference machinery." FEBS Letters **580**(12): 2896-2902.

Berkhout, B. and O. ter Brake (2009). "Towards a durable RNAi gene therapy for HIV-AIDS." Expert Opinion **9**(2): 161-170.

Bernstein, E., A. A. Caudy, et al. (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." Nature **409**: 363-366.

Bitko, V. and S. Barik (2008). "Nasal delivery of siRNA." Methods in Molecular Biology **442**: 75-82.

Bitko, V., A. Musiyenko, et al. (2005). "Inhibition of respiratory viruses by nasally administered siRna." Nature Medicine **11**(1): 50-55.

Blankson, J. N., D. Persaud, et al. (2002). "The challenge of viral reservoirs in HIV-1 infection." Annual Review of Medicine **53**: 557-593.

Boden, D., O. Pusch, et al. (2003). "Human immunodeficiency virus type 1 escape from RNA interference." Journal of Virology **77**(21): 11531-11535.

Boden, D., O. Pusch, et al. (2003). "Promoter choice affects the potency of HIV-1 specific RNA interference." Nucleic Acids Research **31**(17): 5033-5038.

Boden, D., O. Pusch, et al. (2004). "Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins." Nucleic Acids Research **32**(3): 1154-1158.

Bohnsack, M. T., K. Czaplinski, et al. (2004). "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs." RNA **10**(2): 185-191.

Bolstad, B. M., R. A. Irizarry, et al. (2003). "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias." Bioinformatics **19**(2): 185-93.

Borchert, G. M., W. Lanier, et al. (2006). "RNA polymerase III transcribes human microRNAs." Nature Structural and Molecular Biology **13**(12): 1097-1101.

Borden, E. C., G. C. Sen, et al. (2007). "Interferons at age 50: past, current and future impact on biomedicine." Nature Reviews Drug Discovery **6**: 975-990.

Boudreau, R. L., A. Mas Monteys, et al. (2008). "Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs." RNA **14**: 1834–1844.

Brass, A. L., D. M. Dykxhoorn, et al. (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." Science **319**(5865): 921-926.

Brelot, A. and M. Alizon (2001). "HIV-1 entry and how to block it." AIDS **15**(Supplement 5): S3-S11.

Brennecke, J., A. Stark, et al. (2005). "Principles of microRNA-target recognition." PLoS Biology **3**(3): e85.

Brodersen, P. and O. Voinnet (2009). "Revisiting the principles of microRNA target recognition and mode of action." Nature Reviews Molecular Cell Biology **10**: 141-148.

Browning, C. M., L. Cagnon, et al. (1999). "Potent inhibition of human immunodeficiency virus type 1 (HIV-1) gene expression and virus production by an HIV-2 Tat activation-response RNA decoy." Journal of Virology **73**(6): 5191-5195.

Cagnon, L. and J. Rossi (2000). "Downregulation of the CCR5 beta-chemokine receptor and inhibition of HIV-1 infection by stable VA1-ribozyme chimeric transcripts." Antisense Nucleic Acid Drug Development **10**: 251-261.

Campbell, E. M. and T. J. Hope (2008). "Live cell imaging of the HIV-1 life cycle." TRENDS in Microbiology **16**(12): 580-587.

Caplen, N. J., S. Parrish, et al. (2001). "Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems." Proceedings of the National Academy of Sciences. USA **98**(17): 9742-9747.

Capodici, J., K. Kariko, et al. (2002). "Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference." The Journal of Immunology **169**: 5196-5201.

Caputo, A., M. P. Grossi, et al. (1996). "Inhibition of HIV-1 replication and reactivation from latency by Tat transdominant negative mutants in the cysteine rich region " Gene Therapy **3**: 235-245.

Carmona, S., A. Ely, et al. (2006). "Effective inhibition of HBV replication *in vivo* by anti-HBx short hairpin RNAs." Molecular Therapy **13**(2): 411-421.

Castanotto, D., J. R. Li, et al. (2002). "Intracellular ribozyme applications." Biochemical Society Transactions **30**(6): 1140-1145.

Castanotto, D., K. Sakurai, et al. (2007). "Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC." Nucleic Acids Research **35**(15): 5154-5164.

Castanotto, D., S. Tommasi, et al. (2005). "Short hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells." Molecular Therapy **12**(1): 179-183.

Cave, E., M. Weinberg, et al. (2006). "Silencing of HIV-1 subtype C primary isolates by expressed small hairpin RNAs targeted to *gag*." AIDS Research and Human Retroviruses **22**(5): 401-410.

Chan, J. H., S. Lim, et al. (2006). "Antisense oligonucleotides: from design to therapeutic application." Clinical and Experimental Pharmacology and Physiology **33**(5/6): 533-540.

Chang, C. I., H. S. Kang, et al. (2009). "Dual-target gene silencing by using long, synthetic siRNA duplexes without triggering antiviral responses." Molecules and Cells **27**(6): 689-695.

Chang, L. J., X. Liu, et al. (2005). "Lentiviral siRNAs targeting multiple highly conserved RNA sequences of human immunodeficiency virus type 1." Gene Therapy **12**(14): 1133-1144.

Chaubey, B., S. Tripathi, et al. (2007). "Mechanism of RNA cleavage catalyzed by sequence specific polyamide nucleic acid-neamine conjugate." Oligonucleotides **17**(3): 302-313.

Chekulaeva, M. and W. Filipowicz (2009). "Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells." Current Opinion in Cell Biology **21**: 452-460.

Chendrimada, T. P., R. I. Gregory, et al. (2005). "TRBP recruits Dicer complex to Ago2 for microRNA processing and gene silencing." Nature **436**: 740-744.

Cheng, T. L. and W. T. Chang (2007). "Construction of simple and efficient DNA vector-based short hairpin RNA expression systems for specific gene silencing in mammalian cells." Methods in Molecular Biology **408**: 223-241.

Chiu, Y. L., H. Cao, et al. (2004). " Inhibition of human immunodeficiency virus type 1 replication by RNA interference directed against human transcription elongation factor P-TEFb (CDK9/CyclinT1)." Journal of Virology **78**(5): 2517-2529.

Chiu, Y. L., V. B. Soros, et al. (2005). "Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells." Nature **435**(7038): 108-114.

Choe, H., M. Farzan, et al. (1996). "The b-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates." Cell **85**: 1135-1148.

Christensen, H. S., A. Daher, et al. (2007). "Small interfering RNAs against the TAR RNA binding protein, TRBP, a Dicer cofactor, inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production." Journal of Virology **81**(10): 5121-5131.

Chun, T. W., D. Finzi, et al. (1995). "*In vivo* fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency." Nature Medicine **1**(12): 1284-1290.

Chung, K., C. C. Hart, et al. (2006). "Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155." Nucleic Acids Research **34**(7): 1-14.

Clapham, P. R. and A. McKnight (2002). "Cell surface receptors, virus entry and tropism of primate lentiviruses." Journal of General Virology **83**: 1809-1829.

Coburn, G. A. and B. R. Cullen (2002). "Potent and specific Iinhibition of human immunodeficiency virus type 1 replication by RNA interference." Journal of Virology **76**(18): 9225–9231.

Cohen, M. S., J. A. Anderson, et al. (2005). "Acute HIV infection: implications for HIV spread, disease progression and vaccine development." HIV Molecular Immunology: 21-31.

Connor, R. I., B. Chen, et al. (1995). "Vpr is required for efficient replication of human immunodeficiency virus type 1 in mononuclear phagocytes." Virology **206**: 935-944.

Corbeau, P. (2008). "Interfering RNA and HIV: reciprocal interferences." PLoS Pathogens **4**(9): e1 to e7.

Corbeau, P. and F. Wong-Staal (1998). "Anti-HIV effects of HIV vectors." Virology **243**(2): 268-274.

Cramer, P., K. J. Armache, et al. (2008). " Structure of eukaryotic RNA polymerases." Annual Review of Biophysics **37**: 337-352.

D'Orso, I. and A. D. Frankel (2009). "Tat acetylation modulates assembly of a viral-host RNA–protein transcription complex." Proceedings of the National Academy of Sciences. USA **106**(9): 3101-3106.

Daniels, S. M., C. E. Melendez-Peña, et al. (2009). "Characterization of the TRBP domain required for Dicer interaction and function in RNA interference." BMC Molecular Biology **10**(38): e1-e13.

Das, A. T., T. R. Brummelkamp, et al. (2004). "Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition." Journal of Virology **78**(5): 2601-2605.

Das, S. R. and S. Jameel (2005). "Biology of the HIV Nef protein." Indian Journal of Medical Research **121**(4): 315-332.

Davis, A. J., J. M. Carr, et al. (2008). "Human Immunodeficiency Virus type-1 reverse transcriptase exists as post-translationally modified forms in virions and cells." Retrovirology **5**(115): 1-12.

Denli, A. M., B. B. Tops, et al. (2004). "Processing of primary microRNAs by the Microprocessor complex." Nature **432**(7014): 231-235.

Devincenzo, J. P. (2009). "Harnessing RNA interference to develop neonatal therapies: From Nobel Prize winning discovery to proof of concept clinical trials." Early human development **Advanced online publication**.

Diallo, M., C. Arenz, et al. (2003). "Long endogenous dsRNAs can induce complete gene silencing in mammalian cells and primary cultures." Oligonucleotides **13**: 381-392.

Dickins, R. A., K. McJunkin, et al. (2007). "Tissue-specific and reversible RNA interference in transgenic mice." Nature Genetics **39**(7): 914-921.

Dingwall, C., I. Ernberg, et al. (1990). "HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure." The EMBO Journal **9**(12): 4145-4153.

Doench, J. G. and P. A. Sharp (2004). "Specificity of microRNA target selection in translational repression." Genes & Development **18**(5): 504-511.

Doranz, B. J., J. Rucker, et al. (1996). "A dual-tropic primary HIV-1 isolate that uses fusin and the b-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors." Cell **85**: 1149-1158.

Dropulic, B., M. Hermankova, et al. (1996). "A conditionally replicating HIV-1 vector interferes with wild-type HIV-1 replication and spread." Proceedings of the National Academy of Sciences. USA **93**: 11103-11108.

Duverger, A., J. Jones, et al. (2009). "Determinants of the establishment of human immunodeficiency virus type 1 latency." Journal of Virology **83**(7): 3078-3093.

Dvir, A., R. C. Conaway, et al. (1996). "Promoter escape by RNA polymerase II. A role for an ATP cofactor in suppression of arrest by polymerase at promoter-proximal sites." Journal of Biological Chemistry **271**(38): 23352-23356.

Dvorin, J. D. and M. H. Malim (2003). "Intracellular trafficking of HIV-1 cores: journey to the center of the cell." Current Topics in Microbiology and Immunology **281**: 179-208.

Ehrlich, L. S., T. Liu, et al. (2001). "HIV-1 capsid protein forms spherical (immature-like) and tubular (mature-like) particles in vitro: structure switching by pH-induced conformational changes." Biophysical Journal **81**: 586-594.

Eisen, M. B., P. T. Spellman, et al. (1998). "Cluster analysis and display of genome-wide expression patterns." Proceedings of the National Academy of Sciences. USA **95**(25): 14863-14868.

Elbashir, S. M., J. Harborth, et al. (2001). "Duplexes of 21±nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**: 494-498.

Elbashir, S. M., W. Lendeckel, et al. (2001). "RNA interference is mediated by 21-and 22-nucleotide RNAs." Genes and Development **15**: 188-208.

Ely, A., T. Naidoo, et al. (2009). "Efficient silencing of gene expression with modular trimeric Pol II expression cassettes comprising microRNA shuttles." Nucleic Acids Research **37**(13): e91.

Ely, A., T. Naidoo, et al. (2008). "Expressed anti-HBV primary microRNA shuttles inhibit viral replication efficiently *in vitro* and *in vivo*." Molecular Therapy **16**(6): 1105-1112.

Endres, M. J., P. R. Clapham, et al. (1996). "CD4-independent infection by HIV-2 is mediated by Fusin/CXCR4." Cell **87**: 745-756.

Engelman, A. (2003). "The roles of cellular factors in retroviral integration." Current Topics in Microbiology and Immunology **281**: 209-238.

Engelman, A., I. Oztop, et al. (2009). "Quantitative analysis of HIV-1 preintegration complexes." Methods **Advanced online publication**.

Eulalio, A., E. Huntzinger, et al. (2008). "Getting to the root of miRNA-mediated gene silencing." Cell **132**: 9-14.

Felice, B., C. Cattoglio, et al. (2009). "Transcription factor binding sites are genetic determinants of retroviral integration in the human genome." PLoS ONE **4**(2): e4571.

Filipowicz, W., S. N. Bhattacharyya, et al. (2009). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nature Reviews Genetics **9**: 102-114.

Finzi, D., J. Blankson, et al. (1999). "Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy." Nature Medicine **5**(5): 512-517.

Fire, A., S. Xu, et al. (1998). "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans* " Nature **391**: 806-811.

Flexner, C. (1998). "HIV-protease inhibitors." New England Journal of Medicine **338**: 1281-1293.

Forsbach, A., J. G. Nemorin, et al. (2008). "Identification of RNA sequence motifs stimulating sequence-specific TLR8-dependent immune responses." Journal of Immunology **180**(6): 3729-3738.

Foster, J. L. and J. V. Garcia (2008). "HIV-1 Nef: at the crossroads." Retrovirology **5**(84): 1-13.

Freed, E. O. (2001). "HIV-1 Replication." Somatic Cell and Molecular Genetics **26**(1/6): 13-33.

Fridell, R. A., H. P. Bogerd, et al. (1996). "Nuclear export of late HIV-1 mRNAs occurs via a cellular protein export pathway." Proceedings of the National Academy of Sciences. USA **93**(9): 4421-4424.

Ganser-Pornillos, B. K., M. Yeager, et al. (2008). "The structural biology of HIV assembly " Current Opinions in Structural Biology **18**(2): 203-217.

Gao, F., E. Bailes, et al. (1999). "Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*." Nature **397**: 436-441.

Gao, F., L. Yue, et al. (1992). "Human infection by genetically diverse SIVsm-related HIV-2 in West Africa." Nature **358**: 495-499.

Gao, Y., M. A. Lobritz, et al. (2008). "Targets of small interfering RNA restriction during human immunodeficiency virus type 1 replication." Journal of Virology **82**(6): 2938-2951.

Garcia, J. A., D. Harrich, et al. (1989). " Human immunodeficiency virus type 1 LTR TATA and TAR region sequences required for transcriptional regulation." The EMBO Journal **8**(3): 765-778.

Gatfield, D., G. Le Martelot, et al. (2009). "Integration of microRNA miR-122 in hepatic circadian gene expression." Genes & Development **23**: 1313-1326.

Gatignol, A., S. Lainé, et al. (2005). "Dual role of TRBP in HIV replication and RNA interference: viral diversion of a cellular pathway or evasion from antiviral immunity?" Retrovirology **2**(65): 1-16.

Geary, R. S., S. P. Henry, et al. (2002). "Fomivirsen: clinical pharmacology and potential drug interactions." Clinical Pharmacokinetics **41**(4): 255-260.

Gervaix, A., X. li, et al. (1997). "Multigene antiviral vectors inhibit diverse human immunodeficiency virus type 1 clades." Journal of Virology **71**(4): 3048-3053.

Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nature Reviews Genetics **10**: 94-108.

Giering, J. C., D. Grimm, et al. (2008). "Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic." Molecular Therapy **16**(9): 1630-1636.

Gifford, R. J., A. Katzourakis, et al. (2008). "A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution." Proceedings of the National Academy of Sciences. USA **105**(51): 20362–20367.

Gitlin, L., S. Karelsky, et al. (2002). "Short interfering RNA confers intracellular antiviral immunity in human cells." Nature **418**: 430-434.

Goff, S. P. (2001). "Intracellular trafficking of retroviral genomes during the early phase of infection: viral exploitation of cellular pathways." Journal of Gene Medicine **3**: 517-528.

Goila-Gaur, R. and K. Strebel (2008). "HIV-1 Vif, APOBEC, and Intrinsic Immunity." Retrovirology **5**(51): 1-16.

Goila, R. and A. Banerjea (1998). "Sequence specific cleavage of the HIV-1 coreceptor CCR5 gene by a hammer-head ribozyme and a DNA-enzyme: inhibition of the coreceptor function by DNA-enzyme." FEBS Letters **436**: 233-238.

Gottlinger, H. G. (2001). "The HIV-1 assembly machine." AIDS **15**(Supplement 5): S13-S20.

Göttlinger, H. G. (2008). "Virus kept on a leash." Nature **451**: 406-408.

Gregory, R. I., K. Yan, et al. (2004). "The Microprocessor complex mediates the genesis of microRNAs." Nature **432**: 235-240.

Grimm, D., K. L. Streetz, et al. (2006). "Fatality in mice due to oversaturation of celular micro/short hairpin RNA pathways." Nature **441**: 537-541.

Grishok, A., A. E. Pasquinelli, et al. (2001). "Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing." Cell **106**: 23-34.

Guignard, F., C. Combadiere, et al. (1998). "Gene organization and promoter function for CC chemokine receptor 5 (CCR5)." Journal of Immunology **160**(2): 985-992.

Haasnoot, J., W. de Vries, et al. (2007). "The Ebola Virus VP35 protein is a suppressor of RNA silencing." PLoS Pathogens **3**(6): 0794-0803.

Haasnoot, J., E. M. Westerhout, et al. (2007). "RNA interference against viruses: strike and counterstrike." Nature Biotechnology **25**(12): 1435-1443.

Hahn, B. H., G. M. Shaw, et al. (2000). "AIDS as a zoonosis: scientific and public health implications." Science **287**: 607-614.

Hammond, S. M., E. Bernstein, et al. (2000). "An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells." Nature **404**: 293-296.

Han, J., D. Kim, et al. (2007). "Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells." Proceedings of the National Academy of Sciences. USA **104**(30): 12422–12427.

Han, J., Y. Lee, et al. (2004). "The Drosha-DGCR8 complex in primary microRNA processing." Genes and Development **18**: 3016-3027.

Han, J., Y. Lee, et al. (2006). "Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex." Cell **125**: 887-901.

Han, J., J. S. Pedersen, et al. (2009). "Posttranscriptional crossregulation between Drosha and DGCR8." Cell **136**: 75-84.

Han, W., M. Wind-Rotolo, et al. (2004). "Inhibition of human immunodeficiency virus type 1 replication by siRNA targeted to the highly conserved primer binding site." Virology **330**(1): 221-232.

Han, Y. and R. F. Siliciano (2007). "Keeping quiet: microRNAs in HIV-1 latency." Nature Medicine **13**(10): 1138-1140.

Hang, J. Q., Y. Li, et al. (2007). "Substrate-dependent inhibition or stimulation of HIV RNase H activity by non-nucleoside reverse transcriptase inhibitors (NNRTIs)." Biochemistry and Biophysics Research Communications **352**(2): 341-350.

Harrich, D. and B. Hooker (2002). "Mechanistic aspects of HIV-1 reverse transcription initiation." Reviews in Medical Virology **12**(1): 31-45.

Haussecker, D. (2008). "The business of RNAi therapeutics." Human Gene Therapy **19**(5): 451-462.

Hawkins, P. G. and K. V. Morris (2008). "RNA and transcriptional modulation of gene expression." Cell Cycle **7**(5): 602-607.

Hawkins, P. G., S. Santosom, et al. (2009). "Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells." Nucleic Acids Research **Advanced online publication**.

He, J., S. Choe, et al. (1995). "Human immunodeficiency virus type 1 Viral Protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity." Journal of Virology **69**(11): 6705-6711.

He, Y., B. Vogelstein, et al. (2008). "The antisense transcriptomes of human cells." Science **322**(5909): 1804-1805.

Hefner, E., K. Clark, et al. (2008). "Increased potency and longevity of gene silencing using validated Dicer substrates." Journal of Biomolecular Techniques **19**: 231-237.

Heinrich, J., S. Mathur, et al. (2009). "Oligonucleotide-mediated retroviral RNase H activation leads to reduced HIV-1 titer in patient-derived plasma." AIDS **23**(2): 213-221.

Heinzinger, N. K., M. I. Bukrinsky, et al. (1994). "The Vpr protein of human immunodeficiency virus type 1 influences nuclear localization of viral nucleic acids in nondividing host cells." Proceedings of the National Academy of Sciences. USA **91**: 7311-7315.

Hogan, P. G., L. Chen, et al. (2003). "Transcriptional regulation by calcium, calcineurin, and NFAT." Genes & Development **17**: 2205-2232.

Huang, J., F. Wang, et al. (2007). "Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes." Nature Medicine **13**(10): 1241-1247.

Hunt, J. S. and F. Romanelli (2009). "Maraviroc, a CCR5 coreceptor antagonist that blocks entry of human immunodeficiency virus type 1." Pharmacotherapy **29**(3): 295-304.

Irizarry, R. A., B. Hobbs, et al. (2003). "Exploration, normalization, and summaries of high density oligonucleotide array probe level data." Biostatistics **4**(2): 249-264.

Ishikawa, F., M. Yasukawa, et al. (2005). "Development of functional human blood and immune systems in NOD/SCID/IL2 receptor gamma chain null mice " Blood **106**(5): 1565-1573.

Jackson, A. L., S. R. Bartz, et al. (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nature Biotechnology **21**(6): 635-637.

Jackson, A. L., J. Burchard, et al. (2006). "Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing." RNA **12**(7): 1197-1205.

Jacque, J., K. Triques, et al. (2002). "Modulation of HIV-1 replication by RNA interference." Nature **418**: 435-438.

Janowski, B. A., K. E. Huffman, et al. (2005). "Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs." Nature Chemical Biology **1**(4): 216-222.

Janowski, B. A., K. E. Huffman, et al. (2006). "Involvement of AGO1 and AGO2 in mammalian transcriptional silencing." Nature Structural and Molecular Biology **13**(9): 787-92.

Janowski, B. A., S. T. Younger, et al. (2007). "Activating gene expression in mammalian cells with promoter-targeted duplex RNAs." Nature Chemical Biology **3**(3): 166-173.

Jaskiewicz, L. and W. Filipowicz (2008). "Role of Dicer in posttranscriptional RNA silencing." Current Topics in Microbiology and Immunology **320**: 77-97.

Jeeninga, R. E., M. Hoogenkamp, et al. (2000). "Functional differences between the long terminal repeat transcriptional promoters of human immunodeficiency virus type 1 subtypes A through G." Journal of Virology **74**(8): 3740-3751.

Jenuwein, T. and C. D. Allis (2001). "Translating the Histone Code." Science **293**: 1074-1080.

Jinek, M. and J. A. Doudna (2009). "A three-dimensional view of the molecular machinery of RNA interference." Nature **457**: 405-412.

Johnson, V. A., F. Brun-Vézinet, et al. (2008). "Update of the drug resistance mutations in HIV-1: December 2008." Topics in HIV Medicine **16**(5): 138-145.

Jopling, C. L., M. Yi, et al. (2005). "Modulation of hepatitis C virus RNA abundance by a liver-specific microRNA." Science **309**(5740): 1577-1581.

Jordan, A., P. Defechereux, et al. (2001). "The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation." The EMBO Journal **20**(7): 1726-1738.

Kapranov, P., J. Cheng, et al. (2007). "RNA maps reveal new RNA classes and a possible function for pervasive transcription." Science **316**(5830): 1484-1488.

Kawai, T. and S. Akira (2008). "Toll-like receptor and RIG-1-like receptor signaling." Annuals of the New York Academy of Sciences **1143**: 1-20.

Kawamata, T., H. Seitz, et al. (2009). "Structural determinants of miRNAs for RISC loading and slicer-independent unwinding." Nature Structural and Molecular Biology **16**(9): 953-960.

Kawasaki, H. and k. Taira (2003). "Short hairpin type of dsRNAs that are controlled by tRNAVal promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells." Nucleic Acids Research **31**(2): 700-707.

Keele, B. F., F. van Heuverswyn, et al. (2006). "Chimpanzee reservoirs of pandemic and nonpandemic HIV-1." Science **313**(5786): 523-526.

Keiser, O., C. Orrell, et al. (2008). "Public-health and individual approaches to antiretroviral therapy: township South Africa and Switzerland compared." PLoS Medicine **5**(7): e148.

Kertesz, M., N. Iovino, et al. (2007). "The role of site accessibility in microRNA target recognition." Nature Genetics **39**: 1278-1284.

Khvorova, A., A. Reynolds, et al. (2003). "Functional siRNAs and miRNAs exhibit strand bias." Cell **115**: 209-216.

Kim, D. H., M. Longo, et al. (2004). "Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase." Nature Biotechnology **22**(3): 280-282.

Kim, D. H., P. Sætrom, et al. (2008). "MicroRNA-directed transcriptional gene silencing in mammalian cells." Proceedings of the National Academy of Sciences. USA **105**(42): 16230–16235.

Kim, D. H., L. M. Villeneuve, et al. (2006). "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells." Nature Structural and Molecular Biology **13**(9): 793-797.

Kim, T. K., R. H. Ebright, et al. (2000). "Mechanism of ATP-dependent promoter melting by transcription factor IIH." Science **288**(5470): 1418-1422.

Kim, V. N., J. Han, et al. (2009). "Biogenesis of small RNAs in animals." Nature Reviews Molecular Cell Biology **10**(2): 126-139.

Kim, V. N. and J. W. Nam (2006). "Genomics of microRNA." Trends in Genetics **22**(3): 165-173.

Kim, Y. K., C. F. Bourgeois, et al. (2002). "Phosphorylation of the RNA Polymerase II carboxyl-terminal domain by CDK9 is directly responsible for Human Immunodeficiency Virus Type 1 Tat-activated transcriptional elongation." Molecular and Cellular Biology **22**(13): 4622-4637.

Kim, Y. K. and V. N. Kim (2007). "Processing of intronic microRNAs." The EMBO Journal **26**(3): 775-783.

Klase, Z., P. Kale, et al. (2007). "HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR." BMC Molecular Biology **8**(63): 1-19.

Klug, W. S. and M. R. Cummings (2000). Concepts of Genetics New Jersey, Prentice Hall, Inc.

Kohn, D. B., G. Bauer, et al. (1999). "A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children." Blood **94**(1): 368-371.

Konig, R., Y. Zhou, et al. (2008). "Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication." Cell **135**: 49-60.

Konstantinova, P., W. de Vries, et al. (2006). "Inhibition of human immunodeficiency virus type 1 by RNA interference using long-hairpin RNA." Gene Therapy **13**(19): 1403-1413.

Korber, B., M. Muldoon, et al. (2000). "Timing the ancestor of the HIV-1 pandemic strains." Science **288**(5472): 1789-1796.

Krogan, N. J., M. Kim, et al. (2003). "Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II." Molecular and Cellular Biology **23**(12): 4207-4218.

Kumar, P., H. Ban, et al. (2008). "T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice." Cell **134**: 1-10.

Kuwabara, T., J. Hsieh, et al. (2004). "A small modulatory dsRNA specifies the fate of adult neural stem cells." Cell **116**(6): 779-93.

Lagos-Quintana, M., R. Rauhut, et al. (2001). "Identification of novel genes coding for small expressed RNAs." Science **294**: 853-858.

Lagos-Quintana, M., R. Rauhut, et al. (2003). "New microRNAs from mouse and human." RNA **9**(2): 175-179.

Lanman, J., J. Sexton, et al. (2002). "Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly in vitro." Journal of Virology **76**(14): 6900-6908.

Lassen, K., Y. Han, et al. (2004). "The multifactorial nature of HIV-1 latency." TRENDS in Molecular Medicine **10**(11): 525-531.

Lassen, K. G., J. R. Bailey, et al. (2004). "Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting CD4+ T cells *in vivo*." Journal of Virology **78**(17): 9105-9114.

Le Rouzic, E. and S. Benichou (2005). "The Vpr protein from HIV-1: distinct roles along the viral life cycle." Retrovirology **2**(11): 1-14.

Lee, H. Y., A. S. Perelson, et al. (2008). "Dynamic correlation between intrahost HIV-1 quasispecies evolution and disease progression." PLoS Computational Biology **4**(12): e1-e14.

Lee, J. H., G. Culver, et al. (2008). "Analysis of the EIAV Rev-responsive element (RRE) reveals a conserved RNA motif required for high affinity Rev binding in both HIV-1 and EIAV." PLoS ONE **3**(6): e2272.

Lee, J. T. (2009). "Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome." Genes & Development **23**: 1831-1842.

Lee, N. S., T. Dohjima, et al. (2002). "Expression of small interfering RNAs targeted against HIV-1 *rev* transcripts in human cells." Nature Biotechnology **19**: 500-505.

Lee, R. C. and V. Ambros (2001). "An extensive class of small RNAs in *Caenorhabditis elegans*." Science **294**(5543): 862-864.

Lee, S. K. (2005). "Lentiviral delivery of short hairpin RNAs protects CD4 T cells from multiple clades and primary isolates of HIV." Blood **106**(3): 818-826.

Lee, Y., C. Ahn, et al. (2003). "The nuclear RNase III Drosha initiates microRNA processing " Nature **425**: 415-419.

Lee, Y., I. Hur, et al. (2006). "The role of PACT in the RNA silencing pathway." The EMBO Journal **25**: 522-532.

Lee, Y., K. Jeon, et al. (2002). "MicroRNA maturation: stepwise processing and subcellular localisation." The EMBO Journal **21**(17): 4663-4670.

Lee, Y., M. Kim, et al. (2004). "MicroRNA genes are transcribed by RNA polymerase II." The EMBO Journal **23**: 4051-4060.

Leonard, J. N. and D. Schaffer (2005). "Computational design of antiviral RNA interference strategies that resist human immunodeficiency virus escape." Journal of Virology **79**(3): 1645-1654.

Leonard, J. N., P. S. Shah, et al. (2008). "HIV evades RNA interference directed at TAR by an indirect compensatory mechanism." Cell Host Microbe **4**(5): 484-494.

Levine, B. L., W. Bernstein, et al. (1997). "Ex vivo replicative potential of adult human peripheral blood CD4+ T cells." Transplant Proc **29**(4): 2028.

Levine, B. L., L. M. Humeau, et al. (2006). "Gene transfer in humans using a conditionally replicating lentiviral vector." Proceedings of the National Academy of Sciences. USA **103**(46): 17372–17377.

Levine, B. L., J. D. Mosca, et al. (1996). "Antiviral effect and *ex vivo* CD4+ T cell proliferation in HIV-positive patients as a result of CD28 costimulation." Science **272**(5270): 1939-43.

Levy, J. A. (2009). "HIV pathogenesis: 25 years of progress and persistent challenges." AIDS **23**(2): 147-160.

Lewis, B. P., I. H. Shih, et al. (2003). "Prediction of mammalian microRNA targets." Cell **115**(7): 787-798.

Li, L., X. Lin, et al. (2007). "Defining the optimal parameters for hairpin-based knockdown constructs." RNA **13**(10): 1765-1774.

Li, L. C., S. T. Okino, et al. (2006). "Small dsRNAs induce transcriptional activation in human cells." Proceedings of the National Academy of Sciences. USA **103**(46): 17337-17342.

Li, M. and R. Craigie (2009). "Nucleoprotein complex intermediates in HIV-1 integration." Methods **47**(4): 237-242.

Li, M., H. Li, et al. (2006). "RNAi in combination with a ribozyme and TAR decoy for the treatment of HIV infection in haematopoietic cell gene therapy." Annals of the New York Academy of Science **1082**: 172-179.

Li, M. and J. J. Rossi (2005). "Lentiviral vector delivery of siRNA and shRNA encoding genes into cultured and primary hematopoietic cells." Methods in Molecular Biology **309**: 261-272.

Li, M. J., J. Kim, et al. (2005). "Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy." Molecular Therapy **12**(5): 900-909.

Li, P., A. J. Stephenson, et al. (1993). "Double-stranded strong-stop DNA and the second template switch in human immunodeficiency virus (HIV) DNA synthesis." Virology **194**(1): 82-88.

Li, Y., H. Hui, et al. (1992). "Complete nucleotide sequence, genome organization, and biological properties of Human Immunodeficiency Virus Type 1 *in vivo*: evidence for limited defectiveness and complementation." Journal of Virology **66**(11): 6587-6600.

Lim, H. G. W., K. Suzuki, et al. (2008). "Promoter-targeted siRNAs induce gene silencing of simian immunodeficiency virus (SIV) infection *in vitro*." Molecular Therapy **16**(3): 565-570.

Lingel, A., B. Simon, et al. (2003). "Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain." Nature **426**(6965): 465-469.

Liu, J., A. Bartesaghi, et al. (2008). "Molecular architecture of native HIV-1 gp120 trimers." Nature **455**(7209): 109-113.

Liu, J., M. A. Carmell, et al. (2004). "Argonaute2 is the catalytic engine of mammalian RNAi." Science **305**(5689): 1437-1441.

Liu, Y. P., J. Haasnoot, et al. (2007). "Design of extended short hairpin RNAs for HIV-1 inhibition." Nucleic Acids Research **35**(17): 5683-5693.

Liu, Y. P., J. Haasnoot, et al. (2008). "Inhibition of HIV-1 by multiple siRNAs expressed from a single microRNA polycistron." Nucleic Acids Research **36**(9): 2811-2824.

Liu, Y. P., K. J. von Eije, et al. (2009). "Combinatorial RNAi against HIV-1 using extended short hairpin RNAs." Molecular Therapy **17**(10): 1712-1723.

López-Fraga, M., N. Wright, et al. (2008). "RNA interference-based therapeutics: new strategies to fight infectious disease." Infectious Disorders Drug Targets **8**(4): 262-273.

Loseke, S., E. Grage-Griebenow, et al. (2006). "In vitro-generated viral double-stranded RNA in contrast to polyinosinic : polycytidylic acid induces interferon-alpha in human plasmacytoid dendritic cells " Scandinavian Journal of Immunology **63**: 264-274.

Lund, E. and J. E. Dahlberg (2006). "Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs." Cold Spring Harbor Symposia on Quantitative Biology **71**: 59-66.

Lusic, M., A. Marcello, et al. (2003). "Regulation of HIV-1 gene expression by histone acetylation and factor recruitment at the LTR promoter." The EMBO Journal **22**(24): 6550-6561.

Ma, E., I. J. Macrae, et al. (2008). "Autoinhibition of human Dicer by its internal helicase domain." Journal of Molecular Biology **380**: 237-243.

Macpherson, J. L., M. P. Boyd, et al. (2005). "Long-term survival and concomitant gene expression of ribozyme-transduced CD4+ T-lymphocytes in HIV-infected patients." The Journal of Gene Medicine **7**(5): 552-564.

Macrae, I. J., F. Li, et al. (2006). "Structure of Dicer and mechanistic implications for RNAi." Cold Spring Harbor Symposia on Quantitative Biology **71**: 73-80.

Mahajan, V. S., A. Drake, et al. (2008). "Virus-specific host miRNAs: antiviral defenses or promoters of persistent infection?" Trends in Immunology **30**(1): 1-7.

Malim, M. H. and M. Emerman (2008). "HIV-1 accessory proteins--ensuring viral survival in a hostile environment." Cell Host & Microbe **3**(6): 388-398.

Mallon, P. W. (2007). "Pathogenesis of lipodystrophy and lipid abnormalities in patients taking antiretroviral therapy." AIDS Reviews **9**(1): 3-15.

Manjunath, N., p. Kumar, et al. (2006). "Interfering antiviral immunity: application, subversion, hope?" TRENDS in Immunology **27**(7): 328-335.

Manoharan, M. (2004). "RNA interference and chemically modified small interfering RNAs." Current Opinion in Chemical Biology **8**(6): 570-579.

Marcello, A. (2006). "Latency: the hidden HIV-1 challenge." Retrovirology **3**(7): 1-9.

Martin, D. E., R. Blum, et al. (2007). "Multiple-dose pharmacokinetics and safety of bevirimat, a novel inhibitor of HIV maturation, in healthy volunteers." Clinical Pharmacokinetics **46**(7): 589-598.

Martin, D. E., R. Blum, et al. (2007). "Safety and pharmacokinetics of Bevirimat (PA-457), a novel inhibitor of human immunodeficiency virus maturation, in healthy volunteers." Antimicrobial Agents and Chemotherapy **51**(9): 3063-3066.

Martinez, J., A. Patkaniowska, et al. (2002). "Single-stranded antisense siRNAs guide target RNA cleavage in RNAi." Cell **110**: 563-574.

Martinez, J. and T. Tuschl (2004). "RISC is a 5' phosphomonoester-producing RNA endonuclease." Genes & Development **18**(9): 975-980.

Martinez, M. A., M. I. Gutierrez, et al. (2002). "Suppression of chemokine receptor expression by RNA interference allows for inhibition of HIV-1 replication." AIDS **16**: 2385-2390.

Mathews, D. H., J. Sabina, et al. (1999). "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure." Journal of Molecular Biolgy **288**(5): 911-940.

Matranga, C., Y. Tomari, et al. (2005). "Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes." Cell **123**: 1-14.

Matzke, M. A. and J. A. Birchler (2005). "RNAi-mediated pathways in the nucleus." Nature Reviews Genetics **6**: 24-35.

McAllister, C. S. and C. E. Samuel (2009). "The RNA-activated protein kinase enhances the induction of interferon-beta and apoptosis mediated by cytoplasmic RNA sensors." The Journal of Biological Chemistry **284**(3): 1644-1651.

McBride, J. L., R. L. Boudreau, et al. (2008). "Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi." Proceedings of the National Academy of Sciences. USA **105**(15): 5868-5873.

McCaffrey, A. P., H. Nakai, et al. (2003). "Inhibition of hepatitis B virus in mice by RNA interference." Nature Biotechnology **21**(6): 639-644.

McIntyre, G. J. and G. C. Fanning (2006). "Design and cloning strategies for constructing shRNA expression vectors." BMC Biotechnology **6**(1): 1-8.

McIntyre, G. J., J. L. Groneman, et al. (2009). "96 shRNAs designed for maximal coverage of HIV-1 variants." Retrovirology **6**(55): e1-e15.

McManus, M. T., C. P. Petersen, et al. (2002). "Gene silencing using micro-RNA designed hairpins." RNA **8**: 842-850.

Meister, G., M. Landthaler, et al. (2004). "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs." Molecular Cell **15**(2): 185-197.

Michienzi, A., D. Castanotto, et al. (2003). "RNA-mediated inhibition of HIV in a gene therapy setting." Annals of the new York Academy of Sciences **1002**: 63-71.

Michienzi, A., F. G. de Angelis, et al. (2006). "A nucleolar localizing Rev binding element inhibits HIV replication." AIDS Research and Therapy **3**(13): 1-9.

Miller, M. D., C. M. Farnet, et al. (1997). "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition." Journal of Virology **71**(7): 5382-5390.

Miranda, K. C., T. Huynh, et al. (2006). "A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes." Cell **126**(6): 1203-17.

Mitsuyasu, R. T., T. C. Merigan, et al. (2009). "Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells." Nature Medicine **15**(3): 285-292.

Miyoshi, K., H. Tsukumo, et al. (2005). "Slicer function of *Drosophila* Argonautes and its involvement in RISC formation." Genes & Development **19**(23): 2837-2848.

Montero Vega, M. T. and A. de Andrés Martín (2008). "Toll-like receptors: a family of innate sensors of danger that alert and drive immunity." Allergol et Immunopathol **36**(6): 347-357.

Mook, O. R., F. Baas, et al. (2007). "Evaluation of locked nucleic acid–modified small interfering RNA in vitro and *in vivo*." Molecular Cancer Therapeutics **6**(3): 833-843.

Moore, M. D., W. Fu, et al. (2008). "Suboptimal inhibition of protease activity in human immunodeficiency virus type 1: effects on virion morphogenesis and RNA maturation." Virology **379**(1): 152-160.

Morgan, R. A., R. Walker, et al. (2005). "Preferential survival of CD4+ T lymphocytes engineered with anti-human immunodeficiency virus (HIV) genes in HIV-infected individuals." Human Gene Therapy **16**(9): 1065-1074.

Moriuchi, H., M. Moriuchi, et al. (1997). "Cloning and analysis of the promoter region of CCR5, a coreceptor for HIV-1 entry." Journal of Immunology **159**(11): 5441-9.

Morlando, M., M. Ballarino, et al. (2008). "Primary microRNA transcripts are processed co-transcriptionally." **15**(9): 902-909.

Morris, K., S. Chan, et al. (2004). "Small interfering RNA-induced transcriptional gene silencing in human cells." Science **305**: 1289-1292.

Morris, K. V. (2005). "siRNA-mediated transcriptional gene silencing: the potential mechanism and a possible role in the histone code." Cellular and Molecular Life Sciences **62**: 3057–3066.

Morris, K. V. (2009). "Long antisense non-coding RNAs function to direct epigenetic complexes that regulate transcription in human cells." Epigenetics 4:5, 296-301; 1 July 2009 **4**(5): 296-301.

Morris, K. V., C. H. Chung, et al. (2004). "Inhibition of HIV-1 replication by siRNA targeting conserved regions of gag/pol." RNA Biology **1**: 114-117.

Morris, K. V. and D. J. Looney (2005). "Characterization of human immunodeficiency virus (HIV)-2 vector mobilization by HIV-1." Human Gene Therapy **16**: 1463-1472.

Morris, K. V. and J. J. Rossi (2006). "Lentivirus-mediated RNA interference therapy for human immunodeficiency virus type 1 infection." Human Gene Therapy **17**: 479-486.

Morris, K. V., S. Santoso, et al. (2008). "Bidirectional transcription directs both transcriptional gene activation and suppression in human cells." PloS Genetics **4**(11): e1000258.

Muhonen, P., T. Tennilä, et al. (2007). "RNA interference tolerates 2'-fluoro modifications at the Argonaute2 cleavage site." Chemistry and Biodiversity **4**(5): 858-873.

Mummidi, S., S. S. Ahuja, et al. (1997). "The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts with 5'-end heterogeneity, dual promoter usage, and evidence for polymorphisms within the regulatory regions and noncoding exons." Journal Biological Chemistry **272**(49): 30662-71.

Naito, Y., K. Nohtomi, et al. (2007). "Optimal design and validation of antiviral siRNA for targeting HIV-1." Retrovirology **4**(80): 1-5.

Naldini, L., U. Blomer, et al. (1996). "Efficient transfer, integration, and sustained long-term expression of the trans-gene in adult rat brains injected with a lentiviral vector." Proceedings of the National Academy of Sciences. USA **93**: 11382-11388.

Naldini, L., U. Blomer, et al. (1996). "*In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector." Science **272**(5259): 263-267.

Nallagatla, S. R. and P. C. Bevilacqua (2008). "Nucleoside modifications modulate activation of the protein kinase PKR in an RNA structure-specific manner." RNA **14**(6): 1201-1213.

Napoli, C., C. Lemieux, et al. (1990). "Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in *trans*." Plant cell **2**: 279-289.

Nassal, M. (1992). "The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly." Journal of Virology **66**(7): 4107-4116.

Ngok, F. K., R. T. Mitsuyasu, et al. (2004). "Clinical gene therapy research utilizing ribozymes: application to the treatment of HIV/AIDS." Methods in Molecular Biology **252**: 581-598.

Nilsen, T. W. (2007). "Mechanisms of microRNA-mediated gene regulation in animal cells." Trends in Genetics **23**(5): 243-249.

Nishina, K., T. Unno, et al. (2008). "Efficient *in vivo* delivery of siRNA to the liver by conjugation of alpha-tocopherol." Molecular Therapy **16**: 734-740.

Nishitsuji, H., M. Kohara, et al. (2006). "Effective suppression of human immunodeficiency virus type 1 through a combination of short or long-hairpin RNAs targeting essential sequences for retroviral integration." Journal of Virology **80**(15): 7658-7666.

Nisole, S. and A. Saïb (2004). "Early steps of retrovirus replicative cycle." Retrovirology **1**(9): 1-20.

Novina, C. D., M. F. Murray, et al. (2002). "siRNA-directed inhibition of HIV-1 infection." Nature Medicine **8**(7): 681-686.

O'Connell, R. M., A. A. Chaudhuri, et al. (2009). "Inositol phosphatase SHIP1 is a primary target of miR-155." Proceedings of the National Academy of Sciences. USA **106**(17): 7113-7118.

Obernosterer, G., H. Tafer, et al. (2008). "Target site effects in the RNA interference and microRNA pathways." Biochemical Society Transactions **36**(6): 1216-1219.

Ojwang, J. O., A. Hampel, et al. (1992). "Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme." Proceedings of the National Academy of Sciences. USA **89**: 10802-10806.

Okamura, K., J. W. Hagen, et al. (2007). "The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*." Cell **130**(1): 89-100.

Omoto, S. and Y. R. Fujii (2005). "Regulation of human immunodeficiency virus 1 transcription by nef microRNA." Journal of General Virology **86**: 751-755.

Omoto, S., M. Ito, et al. (2004). "HIV-1 nef suppression by virally encoded microRNA." Retrovirology **1**(44): e1-e12.

Paiardini, M., i. Frank, et al. (2008). "Mucosal immune dysfunction in AIDS pathogenesis." AIDS Reviews **10**: 36-46.

Pal, M. and D. S. Luse (2002). "Strong natural pausing by RNA polymerase II within 10 bases of transcription start may result in repeated slippage and re-extension of the nascent RNA." Molecular and Cell Biology **22**(1): 30-40.

Palella, F. J., K. M. Delaney, et al. (1998). "Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection." New England Journal of Medicine **338**(13): 853-860.

Passman, M., M. Weinberg, et al. (2000). "In situ demonstration of inhibitory effects of hammerhead ribozymes that are targeted to the hepatitis Bx sequence in cultured cells." Biochemical and Biophysical Research Communications **268**: 728-733.

Paul, C. P., P. D. Good, et al. (2002). "Effective expression of small interfering RNA in human cells." Nature Biotechnology **20**(5): 505-508.

Peden, K., M. Emerman, et al. (1991). "Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI." Virology **185**: 661-672.

Pelchen-Matthews, A., B. Kramer, et al. (2003). "Infectious HIV-1 assembles in late endosomes in primary macrophages." Journal of Cell Biology **162**(3): 443-55.

Perdomo, M. F., M. Levi, et al. (2008). "Neutralization of HIV-1 by redirection of natural antibodies." Proceedings of the National Academy of Sciences. USA **105**(34): 12515-12520.

Pereira, L. A., K. Bentley, et al. (2000). "A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter." Nucleic Acids Research **28**(3): 663-668.

Perelson, A. S., A. U. Neumann, et al. (1996). "HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life span and viral generation time." Science **271**(5255): 1582-1586.

Pierson, T. C., Y. Zhou, et al. (2002). "Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection." Journal of Virology **76**(17): 8518-8531.

Pillay, V., C. Pillay, et al. (2008). "HIV Type 1 subtype C drug resistance among pediatric and adult South African patients failing antiretroviral therapy." AIDS Research and Human Retroviruses **24**(11): 1449-1454.

Place, R. F., L. C. Li, et al. (2008). "MicroRNA-373 induces expression of genes with complementary promoter sequences." Proceedings of the National Academy of Sciences. USA **105**(5): 1608-1613.

Plantier, J. C., Leoz, M, et al (2009). “A new human immunodeficiency virus derived from gorillas.” Nature Medicine **15** (8): 871-872.

Platt, E. J., K. Wehrly, et al. (1998). "Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1." Journal of Virology **72**(4): 2855-2864.

Podsakoff, G. M., B. C. Engel, et al. (2005). " Selective survival of peripheral blood lymphocytes in children with HIV-1 following delivery of an anti-HIV gene to bone marrow CD34+ cells." Molecular Therapy **12**(1): 77-86.

Pollard, V. W. (1998). "The HIV-1 Rev protein." Annual Review of Microbiology **52**: 491-532.

Pond, S. J. K., W. K. Ridgeway, et al. (2008). "HIV-1 Rev protein assembles on viral RNA one molecule at a time." Proceedings of the National Academy of Sciences. USA **106**(5): 1404–1408.

Popov, S., M. Rexach, et al. (1998). "Viral protein R regulates nuclear import of the HIV-1 pre-integration complex." The EMBO Journal **17**(4): 909-917.

Prakash, T. P. and B. Bhat (2007). "2'-modified oligonucleotides for antisense therapeutics." Current Topics in Medicinal Chemistry **7**(7): 641-649.

Prater, C. E., A. D. Saleh, et al. (2007). "Allosteric inhibition of the HIV-1 Rev/RRE interaction by a 3'-methylphosphonate modified antisense oligo-2'-O-methylribonucleotide." Oligonucleotides **17**(3): 275-290.

Preall, J. B. and E. J. Sontheimer (2005). "RNAi: RISC gets loaded." Cell **123**: 543-553.
Price, D. H. (2000). "P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II." Molecular and Cell Biology **20**(8): 2629-2634.

Provost, P., D. Dishart, et al. (2002). "Ribonuclease activity and RNA binding of recombinant human Dicer." The EMBO Journal **21**(21): 5864-5874.

Pusch, O., D. Boden, et al. (2003). "Nucleotide sequence homology requirements of HIV-1-specific short hairpin RNA." Nucleic Acids Research **32**(22): 6444-6449.

Puthenveetil, S., L. Whitby, et al. (2006). "Controlling activation of the RNA-dependent protein kinase by siRNAs using site-specific chemical modification." Nucleic Acids Research **34**(17): 4900-4911.

Qin, X., D. S. An, et al. (2003). "Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5." Proceedings of the National Academy of Sciences. USA **100**(1): 183-188.

Quinonez, R. and R. E. Sutton (2002). "Lentiviral vectors for gene delivery into cells." DNA and Cell Biology **21**(12): 937-951.

Ramanathan, Y., S. M. Rajpara, et al. (2001). "Three RNA Polymerase II carboxyl-terminal domain kinases display distinct substrate preferences." The Journal of Biological Chemistry **276**(14): 10913-10920.

Rand, T. A., S. Petersen, et al. (2005). "Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation " Cell **123**: 1-9.

Rayburn, E. R. and R. Zhang (2008). "Antisense, RNAi, and gene silencing strategies for therapy: mission possible or impossible?" Drug Discovery Today **13**(11-12): 513-521.

Rehmsmeier, M., P. Steffen, et al. (2004). "Fast and effective prediction of microRNA/target duplexes." RNA **10**(10): 1507-1517.

Ren, G. L., Y. Fang, et al. (2007). "The short hairpin RNA driven by polymerase II suppresses both wild-type and lamivudine-resistant hepatitis B virus strains." Antiviral Therapy **12**(6): 865-876.

Reynolds, A., E. M. Anderson, et al. (2006). "Induction of the interferon response by siRNA is cell type– and duplex length–dependent." RNA **12**: 988-993.

Rhodes, M. M., K. Reblova, et al. (2006). "Trapped water molecules are essential to structural dynamics and function of a ribozyme." Proceedings of the National Academy of Sciences. USA **103**(36): 13380–13385.

Richman, D. D., D. M. Margolis, et al. (2009). "The challenge of finding a cure for HIV infection " Science **323**: 1304-1307.

Richter, S., H. Cao, et al. (2002). "Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation." Biochemistry **41**: 6391-6397.

Richter, S., Y. Ping, et al. (2002). "TAR RNA loop: a scaffold for the assembly of a regulatory switch in HIV replication." Proceedings of the National Academy of Sciences. USA **99**(12): 7928–7933.

Rimsky, L. T., D. C. Shugars, et al. (1998). "Determinants of human immunodeficiency virus type 1 gp41-derived inhibitory peptides." Journal of Virology **72**: 986-993.

Robbins, M., A. Judge, et al. (2007). "2'-O-methyl-modified RNAs act as TLR7 antagonists." Molecular Therapy **15**(9): 1663-1669.

Robbins, M. A., M. Li, et al. (2006). "Stable expression of shRNAs in human CD34+ progenitor cells can avoid induction of interferon responses to siRNAs *in vitro*." Nature Biotechnology **24**(5): 566-571.

Robinson, R. (2004). "RNAi therapeutics: how likely, how soon?" PLoS Biology **2**(1): 0018-0020.

Rodriguez, M. S., C. Dargemont, et al. (2004). "Nuclear export of RNA." Biology of the Cell **96**(8): 639-655.

Rosati, M., A. Valentin, et al. (2001). "CCAAT-enhancer-binding protein beta (C/EBP beta) activates CCR5 promoter: increased C/EBP beta and CCR5 in T lymphocytes from HIV-1-infected individuals." Journal of Immunology **167**(3): 1654-62.
Rossi, J. J., C. H. June, et al. (2007). "Genetic therapies against HIV." Nature Biotechnology **25**(12): 1444-1454.

Ruby, J. G., C. H. Jan, et al. (2007). "Intronic microRNA precursors that bypass Drosha processing." Nature **448**(7149): 83-86.

Saayman, S., S. Barichievy, et al. (2008). "The efficacy of generating three independent anti-HIV-1 siRNAs from a single U6 RNA pol III-expressed long hairpin RNA." PloS One **3**(7): 1-10.

Sachdeva, G., J. D'Costa, et al. (2007). "Chimeric HIV-1 and HIV-2 lentiviral vectors with added safety insurance." Journal of Medical Virology **79**: 118-126.

Saetrom, P., B. S. Heale, et al. (2007). "Distance constraints between microRNA target sites dictate efficacy and cooperativity." Nucleic Acids Research **35**(7): 2333-2342.

Salim, A. and L. Ratner (2008). "Modulation of beta-catenin and E-cadherin interaction by Vpu increases Human Immunodeficiency Virus Type 1 particle release." Journal of Virology **82**(8): 3932-3938.

Saliou, J. M., C. F. Bourgeois, et al. (2009). "Role of RNA structure and protein factors in the control of HIV-1 splicing." Frontiers in Bioscience **14**: 2714-2729.

Salzwedel, K. and E. A. Berger (2000). "Cooperative subunit interactions within the oligomeric envelope glycoprotein of HIV-1: functional complementation of specific defects in gp120 and gp41." Proceedings of the National Academy of Sciences. USA **97**(23): 12794-12799.

Samson, M., F. Libert, et al. (1996). "Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR5 chemokine receptor gene." Nature **382**: 722-725.

Sano, M., H. Li, et al. (2008). "Expression of long anti-HIV-1 hairpin RNAs for the generation of multiple siRNAs: advantages and limitations." Molecular Therapy **16**(1): 170-177.

Sarver, N., E. M. Cantin, et al. (1990). "Ribozymes as potential anti-HIV-1 therapeutic agents." Science **247**(4947): 1222-1225.

Sato, M., T. Motomura, et al. (2006). "Novel HIV-1 integrase inhibitors derived from quinolone antibiotics." Journal of Medical Chemistry **49**(5): 1506-8.

Schaeffer, E., R. Geleziunas, et al. (2001). "Human immunodeficiency virus type 1 Nef functions at the level of virus entry by enhancing cytoplasmic delivery of virions." Journal of Virology **75**(6): 2993-3000.

Scherer, L. J., R. Frank, et al. (2007). "Optimization and characterization of tRNA-shRNA expression constructs." Nucleic Acids Research **35**(8): 2620-2628.

Schlee, M., A. Roth, et al. (2009). "Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus." Immunity **31**(1): 25-34.

Scholefield, J., L. J. Greenberg, et al. (2009). "Design of RNAi hairpins for mutation-specific silencing of ataxin-7 and correction of a SCA7 phenotype." PLoS One **4**(9): e1-e11.

Schwartz, J. C., S. T. Younger, et al. (2008). "Antisense transcripts are targets for activating small RNAs." Nature Structural & Molecular Biology **15**(8): 842-848.

Schwarz, D. S., H. Ding, et al. (2006). "Designing siRNA that distinguish between genes that differ by a single nucleotide." PLoS Genetics **2**(9): 1307-1318.

Schwarz, D. S., G. Hutvagner, et al. (2003). "Asymmetry in the assembly of the RNAi enzyme complex." Cell **115**: 199-208.

Scripture-Adams, D. D., D. G. Brooks, et al. (2002). "Interleukin-7 induces expression of latent human immunodeficiency virus type 1 with minimal effects on T-cell phenotype." Journal of Virology **76**: 13077-13082.

Seitz, H. and P. D. Zamore (2006). "Rethinking the microprocessor." Cell **125**(5): 827-829.

Senserrich, J., E. Pauls, et al. (2008). "HIV-1 resistance to the anti-HIV activity of a shRNA targeting a dual-coding region." Virology **372**: 421-429.

Shabalina, S. A. and E. V. Koonin (2008). "Origins and evolution of eukaryotic RNA interference." Trends in Ecology and Evolution **23**(10): 578-587.

Sharova, N., C. Swingler, et al. (2005). "Macrophages archive HIV-1 virions for dissemination in *trans*." Embo Journal **24**(13): 2481-2489.

Sharp, P. M., E. Bailes, et al. (2001). "The origins of AIDS viruses: where and when?" Philosoy Transcripts of the Royal Society of London **356**: 867-876.

Shiota, M., Y. Ikeda, et al. (2006). "The factors that contribute to the long-term expression of siRNA." Nucleic Acids Symposium Series No. 50: 243-244.

Siliciano, J. D., J. Kajdas, et al. (2003). "Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells." Nature Medicine **9**(6): 727-728.

Sims III, R. J., R. Belotserkovskaya, et al. (2004). "Elongation by RNA polymerase II: the short and long of it." Genes & Development **18**: 2437-2468.

Singh, S. K., M. P. Bhadra, et al. (2008). "MicroRNAs – micro in size but macro in function." FEBS Journal **275**: 4929-4944.

Sledz, C. A., M. Holko, et al. (2003). "Activation of the interferon system by short-interfering RNAs." Nature Cell Biology **5**(9): 834-839.

Song, E., S. Lee, et al. (2003). "Sustained small interfering RNA-mediated human immunodeficiency virus type 1 inhibition in primary macrophages." Journal of Virology **77**(13): 7174-7181.

Song, J., A. Giang, et al. (2007). "Multiple shRNA expressing vector enhances efficiency of gene silencing." BMB Reports **41**(5): 358-362.

Spurgers, K. B., C. M. Sharkey, et al. (2008). "Oligonucleotide antiviral therapeutics: Antisense and RNA interference for highly pathogenic RNA viruses." Antiviral Research **78**: 26-36.

Stefani, G. and F. J. Slack (2008). "Small non-coding RNAs in animal development." Nature Reviews Molecular Cell Biology **9**: 219-230.

Stegmeier, F., G. Hu, et al. (2005). "A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells." Proceedings of the National Academy of Sciences. USA **102**(37): 13212-13217.

Sullenger, B. A., H. F. Gallardo, et al. (1990). "Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication." Cell **63**: 601-608.

Summerton, J. E. (2007). "Morpholino, siRNA and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity." Current Topics in Medical Chemistry **7**(7): 651-660.

Sun, D., M. Melegari, et al. (2006). "Multi-miRNA hairpin method that improves gene knockdown efficiency and provides linked multi-gene knockdown." BioTechniques **41**: 59-63.

Sun, F., J. Wang, et al. (2009). "Characterization of function and regulation of miR-24-1 and miR-31 " Biochemical and Biophysical Research Communications **380**(3): 660-665.

Surabhi, R. M. and R. B. Gaynor (2002). "RNA interference directed against viral and cellular targets inhibits human immunodeficiency virus type 1 replication." Journal of Virology **76**(24): 12963-12973.

Sutherland, D. R., M. Keeney, et al. (2003). "Enumeration of CD34+ hematopoietic stem and progenitor cells." Current Protocols in Cytometry **Chapter 6**: Unit 6 4.

Sutton, R. E. (2000). Lentiviral vectors. Viral Vectors. A. Cid-Arregui and A. Garcia-Carranca. Natick, MA, Biotechniques Books Eaton Publishing**:** 485-499.

Suzuki, K., T. Juelich, et al. (2008). "Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region." The Journal of Biological Chemistry **283**(34): 23353-23363.

Suzuki, K., T. Shijuuku, et al. (2005). "Prolonged transcriptional silencing and CpG methylation induced by siRNAs targeted to the HIV-1 promoter region." Journal of RNAi and Gene Silencing **1**(2): 66-78.

Suzuki, Y. and R. Craigie (2007). "The road to chromatin — nuclear entry of retroviruses." Nature Reviews Microbiology **5**: 187-196.

Swayze, E. E., A. M. Siwkowski, et al. (2007). "Antisense oligonucleotides containing locked nucleic acid improve potency but cause significant hepatotoxicity in animals." Nucleic Acids Research **35**(2): 687-700.

Takahasi, K., M. Yoneyama, et al. (2008). "Non-self RNA-sensing mechanism of RIG-I helicase and activation of antiviral immune responses." Molecular Cell **29**(4): 428-440.

Tay, Y., J. Zhang, et al. (2008). "MicroRNAs to *Nanog*, *Oct4* and *Sox2* coding regions modulate embryonic stem cell differentiation." Nature **455**(7216): 1124-1128.

Taylor, B. S., M. E. Sobieszczyk, et al. (2008). "The challenge of HIV-1 subtype diversity." New England Journal of Medicine **358**(15): 1590-1602.

Templeton, A. R., M. G. Kramer, et al. (2009). "Multiple-infection and recombination in HIV-1 within a longitudinal cohort of women." Retrovirology **6**(54): e1-e12.

ter Brake, O. and B. Berkhout (2007). "Lentiviral vectors that carry anti-HIV shRNAs: problems and solutions." The Journal of Gene Medicine **9**: 743-750.

ter Brake, O., K. Hooft, et al. (2008). "Lentiviral vector design for multiple shRNA expression and durable HIV-1 inhibition." Molecular Therapy **16**(3): 557-564.

ter Brake, O., P. Konstantinova, et al. (2006). "Silencing of HIV-1 with RNA interference: a multiple shRNA approach." Molecular Therapy **14**(6): 883-892.

ter Brake, O., N. Legrand, et al. (2009). "Evaluation of safety and efficacy of RNAi against HIV-1 in the human immune system (Rag-2-/-cc -/-) mouse model." Gene Therapy **16**: 148-153.

Ting, A. H., K. E. Schuebel, et al. (2005). "Short double stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation." Nature Genetics **37**(8): 906-910.

Triboulet, R., B. Mari, et al. (2007). "Suppression of microRNA-silencing pathway by HIV-1 during virus replication." Science **315**(5818): 1579-1582.

Trono, D., M. B. Feinberg, et al. (1989). "HIV-1 Gag mutants can dominantly interfere with the replication of the wild-type virus." Cell **59**: 113-120.

Tuaillon, E., M. Gueudin, et al. (2004). "Phenotypic susceptibility to nonnucleoside inhibitors of virion-associated reverse transcriptase from different HIV types and groups." Journal of Acquired Immune Deficiency Syndrome **37**(5): 1543-1549.

Tufarelli, C., J. A. Stanley, et al. (2003). "Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease." Nature Genetics **34**(2): 157-165.

Turner, A. W., J. De La Cruz, et al. (2009). "Mobilization-competent lentiviral vector–mediated sustained transcriptional modulation of HIV-1 expression." Molecular Therapy **17**(2): 360-368.

Tuschl, T. and A. Borkhardt (2002). "Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy." Molecular Interventions **2**(3): 158-167.

Ui-Tei, K., S. Zenno, et al. (2000). "Sensitive assay of RNA interference in Drosophila and Chinese hamster cultured cells using firefly luciferase gene as target " FEBS Letters **479**: 79-82.

UNAIDS (2008). "Report on the global AIDS epidemic." 1-362.

Verani, A., G. Gras, et al. (2005). "Macrophages and HIV-1: dangerous liaisons." Molecular Immunology **42**: 195-212.

Verma, I. M. and N. Somia (1997). "Gene therapy - promises, problems and prospects." Nature **389**: 239-242.

Vermeulen, A., L. Behlen, et al. (2005). "The contributions of dsRNA structure to Dicer specificity and efficiency." RNA **11**(5): 674-682.

Vlassov, A. V., B. Korba, et al. (2007). "shRNAs targeting hepatitis C: effects of sequence and structural features, and comparision with siRNA." Oligonucleotides **17**(2): 223-236.

von Eije, K. J. and B. Berkhout (2009). "RNA-interference-based gene therapy approaches to HIV type-1 treatment: tackling the hurdles from bench to bedside RNA-interference-based gene therapy approaches to HIV type-1 treatment: tackling the hurdles from bench to bedside." Antiviral Chemistry and Chemotherapy **19**(6): 221-233.

von Eije, K. J., O. ter Brake, et al. (2008). "Human immunodeficiency virus type 1 escape is restricted when conserved genome sequences are targeted by RNA interference." Journal of Virology **82**(6): 2895-2903.

von Eije, K. J., O. ter Brake, et al. (2009). "Stringent testing identifies highly potent and escape-proof anti-HIV short hairpin RNAs." The Journal of Gene Medicine **11**(6): 459-467.

Wain, L. V., E. Bailes, et al. (2007). "Adaptation of HIV-1 to its human host." Molecular Biology and Evolution **24**(8): 1853-1860.

Walensky, R. P., A. D. Paltiel, et al. (2006). "The survival benefits of AIDS treatment in the United States." The Journal of Infectious Diseases **194**(1): 11-19.

Wang, J., T. W. Theunissen, et al. (2007). "Site-directed, virus-free, and inducible RNAi in embryonic stem cells." Proceedings of the National Academy of Sciences. USA **104**(52): 20850-20855.

Watanabe, S., S. Ohta, et al. (2007). "Humanized NOD/SCID/IL2R gamma null mice transplanted with hematopoietic stem cells under nonmyeloablative conditions show prolonged life spans and allow detailed analysis of human immunodeficiency virus type 1 pathogenesis." Journal of Virology **81**(23): 13259-13264.

Watanabe, T., M. Sudoh, et al. (2006). "Intracellular-diced dsRNA has enhanced efficacy for silencing HCV RNA and overcomes variation in the viral genotype." Gene Therapy **13**(11): 883-892.

Watanabe, Y., N. Yachie, et al. (2006). "Computational analysis of microRNA targets in *Caenorhabditis elegans*." Gene **365**: 2-10.

Watson, C., S. Jenkinson, et al. (2005). "The CCR5 receptor-based mechanism of action of 873140, a potent allosteric non-competitive entry inhibitor." Molecular Pharmacology **67**: 1268-1282.

Weber, M., I. Hellmann, et al. (2007). "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." Nature Genetics **39**(4): 457-66.

Weinberg, M., L. M. Villeneuve, et al. (2006). "The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells." RNA **12**: 256-262.

Weinberg, M. S., A. Ely, et al. (2007). "Specific inhibition of HBV replication in vitro and in vivo with expressed long hairpin RNA." Molecular Therapy **15**(3): 534-541.

Weinberg, M. S. and K. V. Morris (2006). "Are viral-encoded micrRNAs mediating latent HIV-1 infection?" DNA and Cell Biology **25**(4): 223-231.

Westerhout, E. M. and B. Berkhout (2007). "A systematic analysis of the effect of target RNA structure on RNA interference." Nucleic Acids Research **35**(13): 4322-4330.

Westerhout, E. M., M. Ooms, et al. (2005). "HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome." Nucleic Acids Research **33**(2): 796-804.

Westerhout, E. M., O. ter Brake, et al. (2006). "The virion-associated incoming HIV-1 RNA genome is not targeted by RNA interference." Retrovirology **3**(57): 1-9.

Westerhout, E. M., M. Vink, et al. (2006). "A conditionally replicating HIV-based vector that stabley expresses an antiviral shRNA against HIV-1 replication." Molecular Therapy **14**(2): 268-275.

White, P., F. Anastasopoulos, et al. (2009). "Overcoming biological barriers to in vivo efficacy of antisense oligonucleotides." Expert reviews in Molecular Medicine **11**: e10.

Whitehead, K. A., R. Langer, et al. (2009). "Knocking down barriers: advances in siRNA delivery." Nature Reviews Drug Discovery **8**: 129-138.

Wiethoff, C. M. and C. R. Middaugh (2003). "Barriers to nonviral gene delivery." Journal of Pharmaceutical Sciences **92**(2): 203-217.

Wild, C. T., D. C. Shugrass, et al. (1994). "Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41are potent inhibitors of virus infection." Procedings of the National Academy of Sciences. USA **91**: 9770-9774.

Williams, S. A. F., L. Chen, et al. (2006). "NF-kB p50 promotes HIV latency through HDAC recruitment and repression of transcritional initiation." The EMBO Journal **25**(1): 139-149.

Williams, S. A. F. and W. C. Greene (2005). "Host factors regulating post-integration latency of HIV." TRENDS in Microbiology **13**(4): 137-139.

Wingard, S. A. (1928). "Hosts and symptoms of ring spot, a virus disease of plants." Journal of Agricultural Research **37**: 127-153.

Wiskerchen, M. and M. A. Muesing (1995). "Human immunodeficiency virus type 1 integrase: effects of mutations on viral ability to integrate, direct viral gene expression from unintegrated viral DNA templates, and sustain viral propagation in primary cells." Journal of Virology **69**(1): 376-386.

Woffendin, C., U. Ranga, et al. (1996). "Expression of a protective gene prolongs survival of T cells in human immunodeficiency virus-infected patients." Proceedings of the National Academy of Sciences. USA **93**: 2889-2894.

Wyatt, R., P. D. Kwong, et al. (1998). "The antigenic structure of the HIV gp120 envelope glycoprotein " Nature **393**(6686): 705-711.

Yamagishi, T. Ishida, et al. (2009). "Retroviral delivery of promoter-targeted shRNA induces long-term silencing of HIV-1 transcription." Microbes and Infection **Advanced online publication**.

Yang, W. and W. Paschen (2008). "Conditional gene silencing in mammalian cells mediated by a stress-inducible promoter." Biochemical and Biophysical Research Communications **365**(3): 521-527.

Yekta, S., I. H. Shih, et al. (2004). "MicroRNA-directed cleavage of HOXB8 mRNA." Science **304**(5670): 594-596.

Yekta, S., C. J. Tabin, et al. (2008). "MicroRNAs in the Hox network: an apparent link to posterior prevalence." Nature Reviews Genetics **9**(10): 789-796.

Yeung, M. L., L. Houzet, et al. (2009). "A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication." Journal of Biological Chemistry **284**(29): 19463-19473.

Yi, R., Y. Qin, et al. (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes and Development **17**: 3011-3016.

Ying, S. and S. Lin (2006). "Current perspectives in intronic micro RNAs (miRNAs)." Journal of Biomedical Science **13**: 5-15.

Yoder, K., A. Sarasin, et al. (2006). "The DNA repair genes XPB and XPD defend cells from retroviral infection." Proceedings of the National Academy of Sciences. USA **103**(12): 4622–4627.

Yoshinari, K., M. Miyagishi, et al. (2004). "Effects on RNAi of the tight structure, sequence and position of the targeted region." Nucleic Acids Research **32**(2): 691-699.

Yu, J., S. L. DeRuiter, et al. (2002). "RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells." Proceedings of the National Academy of Sciences. USA **99**(9): 6047-6052.

Yuan, X., C. Liu, et al. (2008). "Clustered microRNAs' coordination in regulating protein-protein interaction network." BMC Systems Biology **3**(65): e1-e10.

Zack, J. A., S. J. Arrigo, et al. (1990). "HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure." Cell **61**: 213-222.

Zamore, P. D. (2001). "RNA interference: listening to the sound of silence." Nature Structural and Molecular Biology **8**(9): 746-750.

Zamore, P. D., T. Tuschl, et al. (2000). "RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." Cell **101**: 25-33.

Zaratiegui, M., D. V. Irvine, et al. (2007). "Noncoding RNAs and gene silencing." Cell **128**(4): 763-776.

Zeng, Y. and B. R. Cullen (2003). "Sequence requirements for micro RNA processing and function in human cells." RNA **9**: 112-123.

Zeng, Y. and B. R. Cullen (2005). "Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences." The Journal of Biological Chemistry **280**(30): 27595-27603.

Zeng, Y. and B. R. Cullen (2005). "Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha." The EMBO Journal **24**: 138-148.

Zeng, Y., E. J. Wagner, et al. (2002). "Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells." Molecular Cell **9**: 1327-1333.

Zhang, H., F. A. Kolb, et al. (2002). "Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP." The EMBO Journal **21**(21): 5875-5885.

Zhang, H., F. A. Kolb, et al. (2004). "Single processing center models for human dicer and bacterial RNase III." Cell **118**: 57-68.

Zhang, J. and C. S. Crumpacker (2002). "Human immunodeficiency virus type 1 nucleocapsid protein nuclear localization mediates early viral mRNA expression." Journal of Virology **76**(20): 10444-10454.

Zhang, R. and T. Nguyen (2008). "Model of human immunodeficiency virus budding and self-assembly: role of the cell membrane." Physical Review **78**: 051903-1 to 051903-12.

Zhen, L., S. Alhousseynou, et al. (2008). "MicroRNA: an emerging therapeutic target and intervention tool." International Journal of Molecular Sciences **9**: 978-999.

Zhou, H., X. G. Xia, et al. (2005). "An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi." Nucleic Acids Research **33**(6): e62-e69.

Zhou, H., M. Xu, et al. (2008). "Genome-scale RNAi screen for host factors required for HIV replication." Cell Host & Microbe **4**: 495-504.

Zhou, M., M. Halanski, et al. (2000). "Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during Human Immunodeficiency Virus Type 1 transcription." Molecular and Cell Biology **20**(14): 5077-5086.

Zuker, M. (2003). "Mfold web server for nucleic acid folding and hybridization prediction." Nucleic Acids Research **31**(13): 3406-3415.

A1 Standard laboratory methods

A1.1 Bacterial growth, storage and transformations

A1.1.1 Bacterial growth and storage

The host bacterial strain used in this work was *Escherichia coli* DH5α (Hanahan, 1983). Cultures were grown with shaking (350 rpm) in Erlenmeyer flasks at 37 ^{o}C in a New Brunswick Series 25 orbital incubator (New Brunswick Scientific Co, Inc, NJ, USA) in Luria Bertani (LB) broth (10 % tryptone powder, 5 % yeast extract, 10 % NaCl (Merck, Darmstadt, Germany) or on LB agar plates [10 % tryptone powder, 5 % yeast extract, 10 % NaCl and 10 % agar powder (Merck, Darmstadt, Germany)] supplemented with 200 µg/mL Ampicillin antibiotic (Sigma-Aldrich, MO , USA). For TA cloning, 40 µL of X-gal (20 mg/ml stock in dimethyl formamide, Sigma-Aldrich, MO, USA) and 4 µL of IPTG (200 mg/ml aqueous solution, Roche, Basel, Switzerland) were spread on Ampicillin positive LB agar plates and left to air dry for 20 minutes at 37 ^{o}C prior to plating. *E.coli* DH5α strains were stored in 33 % glycerol (Merck, Darmstadt, Germany) at -80 ^{o}C.

A1.1.2 Preparation of chemically competent *E.coli*

E. coli cells were grown in 100 mL of LB (see Appendix A1.1.1) until log phase as measured by OD_{600} of 0.4. Cells were chilled for 20 minutes on ice followed by centrifugation for five minutes at 4 ^{o}C and 59.5 x g in a refrigerated centrifuge (Eppendorf 5415D/R, Eppendorf, Hamburg, Germany). The pellet was resuspended in 10 mL of ice cold transformation buffer (60 mM $CaCl_2$, 15 % glycerol, 10 mM PIPES pH 7.0; Merck, Darmstadt, Germany) and incubated for 20 minutes on ice. Followed by centrifugation for 5 minutes at 4 ^{o}C and 59.5 x g. Pellets were resuspended in 1 mL of ice cold transformation buffer and 100 µL aliquots were stored at -80 ^{o}C.

A1.1.3 Bacterial transformations

Plasmid DNA (0.5 µg/µL) was added to a 100 µL aliquot of transformation competent cells (thawed on ice) and incubated on ice for 30 minutes before heat shock for 90 seconds at 42 ^{o}C. Cells were incubated for 60 minutes at 37 ^{o}C without shaking following the addition of 800 µL of

LB, to allow for phenotypic expression of antibiotic resistance genes, prior to plating onto Ampicillin positive LA plates.

A1.2 Plasmid DNA preparation, electrophoresis and extraction

A1.2.1 Plasmid DNA preparation

Plasmid DNA was prepared using three commercially available kits (Qiagen, Hilden, Germany) that differed according to the concentration of DNA required: 'mini-preps' (~150 ng/µL plasmid DNA), 'midi-preps' (~800 ng/µL of plasmid DNA) and 'maxi-preps' (~5 µg/µL of plasmid DNA). Transformed *E.coli* DH5α cultures (three mL for mini-preps, 50 mL for midi-preps and 250 mL for maxi-preps) were grown in LB (see Appendix A1.1.1) until log phase as measured by OD_{600} of 0.4 followed by centrifugation for 20 minutes at 4000 x g (midi and maxi preps) or 16 000 x g (mini preps). Pellets were resuspended in 250 µL (mini prep), 4 mL (midi prep) or 10 mL (maxi prep) of Buffer P1 (50 mM Tris-HCl pH 8.0, 10 mM EDTA and 100 µg/ml RNAse A). An equal volume of Buffer P2 [200 mM NaOH, 1 % SDS (w/v)] was added and mixed by inversion for one minute. An equal volume of pre-chilled Buffer P3 (3.0 M potassium acetate pH 5.5) was added (for midi and maxi preps; 350 µL for mini preps) and mixed by inversion for one minute prior to centrifugation for 20 minutes at 4000 x g for midi and maxi preps or 16 000 x g for mini preps.

For mini preps, the supernatant was pipetted into a QIAprep spin column, centrifuged 60 seconds at 16 000 x g followed by the addition of 750 µL of Buffer PE (wash buffer, composition not provided) and centrifugation for 60 seconds at 16 000 x g. Flow-through was discarded and samples were centrifuged for 60 seconds at 16 000 x g followed by elution into a new microfuge tube with 100 µL of Buffer EB (10 mM Tris-HCl pH 8.5) and centrifugation for 60 seconds at 16 000 x g.

For midi and maxi preps, the supernatant was poured into a QIAtip binding column that had been pre-cleared with 4 mL (midi prep) or 10 mL (maxi prep) of Buffer QBT [750 mM NaCl, 50 mM MOPS pH 7.0, 15 % isopropanol (v/v), 0.15 % Triton® X-100 (v/v)] using gravity flow. The column was washed with 2 x 10 mL (midi prep) or 2 x 30 mL (maxi prep) of Buffer QC [1.0 M NaCl, 50 mM MOPS pH 7.0, 15 % isopropanol (v/v)]. Plasmid DNA was eluted into a new 50 mL tube using 5 mL (midi prep) or 15 mL (maxi prep) of Buffer QF [1.25 M NaCl, 50 mM Tris-HCl pH 8.5, 15 % isopropanol (v/v)] followed by the addition of 3.5 mL (midi prep) or 10.5 mL (maxi prep) of isopropanol (Merck, Darmstadt, Germany) and samples were centrifuged for one hour at 4 ^{o}C and 8000 x g. Pellets were air-dried and resuspended in 250 µL or 500 µL of Buffer TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). Plasmid DNA was subjected to spectrophotometry using a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Inc, MA,

USA) to determine the concentration (at a wavelength of 260 nm, 1 absorbance unit (A_{260}) = 50 µg/mL dsDNA) and purity (A_{260}/A_{280} > 1.8).

A1.2.2 Phenol/chloroform extraction of DNA

The chloroform/phenol method of nucleic acid extraction was used to recover DNA for PCR analysis following immunoprecipitation (ChIP, see section 4.2.5). DNA solutions were brought to a volume of 500 µL with nuclease-free water (Qiagen, Hilden, Germany). A 500 µL volume of 1:1 phenol: chloroform, (1 x volume, Merck, Darmstadt, Germany), was added to the samples which were briefly vortexed for ten seconds to form an emulsion. The mixture was centrifuged for 15 seconds at 12 000 x g. Approximately 500 µL of the aqueous phase was transferred to a new microfuge tube followed by the addition of one volume of chloroform. The samples were briefly vortexed and centrifuged for 15 seconds at 12000 x g. The aqueous phase was removed and 0.1 volumes of 3 M sodium acetate pH 5.2 and 2.5 volumes of 100 % ice-cold ethanol were added. Samples were incubated at -70 ^{o}C for 30 minutes followed by centrifugation for 15 minutes at 13 000 x g and 4 ^{o}C. The pellet was washed with 70 % ethanol and samples were air-dried to remove residual ethanol before being resuspended in 50 µL of nuclease-free water.

A1.2.3 Agarose gel electrophoresis

To make a 0.8 % agarose gel, 0.8 g of agarose powder (Sigma-Aldrich, MO, USA) was dissolved in 100 mL of 1 x TBE buffer (890 mM Tris base, 890 mM Boric acid and 32 mM EDTA, Merck, Darmstadt, Germany) and heated until the agarose had dissolved. Prior to pouring, one µL per 50 mL gel of ethidium bromide (10 mg/mL - Sigma, Germany) was added. DNA samples were mixed with an equal volume of loading dye (1.67 mM Tris-HCl pH 7.6, 0.005 % bromophenol blue, 0.005 % xylene cyanol FF, 10 % glycerol, 10 mM EDTA, Fermentas, WI, USA) and electrophoresed parallel to 2 µL of a DNA molecular weight marker (O'GeneRuler^{TM} Mix, Fermentas, WI, USA) that had been loaded into the wells of a set agarose gel. Gels were subjected to electrophoresis at 100 volts.

A1.2.4 DNA purification from agarose gels

DNA fragments were purified from agarose gels using a commercially available kit (Qiagen, Hilden, Germany). The DNA fragment of interest was excised from the gel using a scalpel blade, weighed and 3 x the volume of a Buffer QG (composition not provided) was added (100 µg = 100 µL) followed by incubation for ten minutes at 50 ^{o}C until melted. One gel volume of

isopropanol (Merck, Darmstadt, Germany) was added, the mixture was inverted ten times and pipetted into a QIAquick spin column followed by centrifugation for 90 seconds at 16 000 x g. Flow through was discarded, 500 µL of Buffer QG was added to the column followed by centrifugation for 90 seconds at 16 000 x g. Flow-through was discarded, 750 µL of Buffer PE (wash buffer, composition not provided) was added to the column followed by centrifugation for 90 seconds at 16 000 x g. Flow-through was discarded followed by centrifugation for 90 seconds at 16 000 x g to dry the column. DNA was eluted into a new microfuge tube with 50 µL of Buffer EB (10 mM Tris-HCl pH 8.5) by centrifugation for 90 seconds at 16 000 x g. Eluted DNA was subjected to spectrophotometry using a NanoDrop Spectrophotometer (Thermo Fisher Scientific, Inc, MA, USA) to determine the concentration (at a wavelength of 260 nm, 1 absorbance unit (A_{260}) = 50 µg/mL dsDNA) and purity (A_{260}/A_{280} > 1.8).

A1.3 RNA extraction and RNA loading dye for Northern blots

A1.3.1 RNA extraction from adherent mammalian cells seeded in 10 cm dishes

Forty eight hours post-transfection mammalian HEK293 or Huh7 cells that had been cultured in 10 cm dishes (Nunc™, AEC-Amersham (Pty) Ltd, South Africa) were washed in PBS (Gibco, UK), resuspended in one mL of TriReagent™ (Sigma, Germany) and transferred to a microfuge tube with 300 µL of chloroform (Merck, Germany) followed by ten seconds of vortexing. Samples were centrifuged for 30 minutes at 16 000 x g and 4 ^{o}C in a refrigerated centrifuge (Eppendorf 5415D/R; Eppendorf, Germany). The upper aqueous phase was carefully transferred to a new microfuge tube, an equal volume (~ 300 µL) of isopropanol (Merck, Germany) was added and samples were centrifuged for 30 minutes at 16 000 x g and 4 ^{o}C. The supernatant was discarded and the pellets air dried before being resuspended in 30 µL RNAse-free de-ionised distilled water (Ambion, Applied Biosystems, USA). RNA was subjected to spectrophotometry using a NanoDrop Spectrophotometer to determine the concentration (at a wavelength of 260 nm, 1 absorbance unit (A_{260}) = 40 µg/mL RNA) and purity (A_{260}/A_{280} > 2).

A1.3.2 RNA extraction from suspension mammalian cells seeded in 6 well dishes

Forty eight hours post-transfection mammalian 1G5 cells that had been cultured in 6 well dishes (Nunc™, AEC-Amersham (Pty) Ltd, South Africa) were pelleted by centrifugation for five minutes at 300 x g, resuspended in 600 µL of Buffer RLT (composition not provided) and transferred to a QIAshredder spin column placed within a two mL collection tube followed by centrifugation for two minutes at 16 000 x g. An equal volume (~600 µL) of 70 % ethanol (Merck, Germany) was

added to the lysate and 700 µL of the sample was transferred to an RNeasy spin column placed in a two mL collection tube followed by centrifugation for 15 seconds at 16 000 x g. Flow-through was discarded and successive aliquots of sample were centrifuged using the same column until all the lysate had been centrifuged. 700 µL of Buffer RW1 (composition not provided) was added to the RNeasy spin column followed centrifugation for 15 seconds at 16 000 x g. Flow-through was discarded and 500 µL of Buffer RPE (composition not provided) was added to the RNeasy spin column followed centrifugation for 15 seconds at 16 000 x g. Flow-through was discarded and 500 µL of Buffer RPE was added followed centrifugation for two minutes at 16 000 x g. The RNeasy spin column was transferred to a new 1.5 mL collection tube and 50 µL of RNAse-free water was added directly to the spin column membrane followed by centrifugation for one minute at 16 000 x g. RNA was subjected to spectrophotometry using a NanoDrop Spectrophotometer to determine the concentration (at a wavelength of 260 nm, 1 absorbance unit (A_{260}) = 40 µg/mL RNA) and purity ($A_{260}/A_{280} > 2$).

A1.3.3 RNA loading dye for Northern Blots

A 2 x stock solution of RNA loading dye contained: 1 mM EDTA pH 8.0 (Sigma-Aldrich, MO, USA), 0.25 % bromophenol blue (Merck, Darmstadt, Germany) and 0.25 % xylene cyanol FF (Merck, Darmstadt, Germany) in 98 % deionised formamide (Merck, Darmstadt, Germany).

A2 Specialised laboratory methods

A2.1 Tissue culture infectious dose (TCID) determination

Total p24 antigen in culture supernatants of cells infected with HIV-1 represented free antigen released from lysed cells and p24 from assembled intact virions. Thus the $TCID_{50}$ assay (50 % tissue culture infectious dose) measured the level of replication competent infectious virus from culture supernatants. The viral supernatant needed to have a p24 value > 30 pg/mL to yield an acceptable $TCID_{50}$ value (p24 ELISA datasheet, Murex Biotech LTD, Kent, UK) thus a dilution series was generated as follows: twenty four hours prior to infection, U87.CD4.CCR5 cells were seeded at 50 000 cells per well in 200 µL of growth media (see section 2.2.4) in flat bottomed 96 well plates (Nunc™, Thermo Fisher Scientific, Inc, MA, USA) The remaining wells were filled with 200 µL each of PBS (Gibco, BRL, UK) to reduce dehydration across the plate. The cells were allowed to settle overnight and the following day, all supernatant from wells seeded with cells was removed. Based on the scheme in Figure A1, 175 µL of growth media was added to all wells from 4^{-4} to 4^{-16} and negative controls. A stock of FV5 was thawed at 37 ^{o}C, immediately

diluted 1:12 in growth media (100 µL FV5 + 1.1 mL of growth media) and 200 µL was pipetted into each 4^{-2} dilution well. Using a multichannel pipette (Eppendorf, Hamburg, Germany), moving left to right across the late and changing tips in between, 25 µL from one row of wells with the same dilution series was transferred to the adjacent row of wells thereby generating a dilution series of FV5 across the plate. Twenty five µL was discarded from each well in the last row (4^{-14}) and cells were left to incubate overnight. The following day (Day 0), the supernatant from each of the 4^{-2} wells was removed and stored at -80 ^{o}C as 'viral input'. The supernatant from all other wells was discarded and all wells were washed three times each in PBS (Gibco, BRL, UK) followed by the addition of 300 µL of fresh growth media to each well. One hundred µL of supernatant from each well was collected and stored at -80 ^{o}C as 'Day 0' samples. Cells were incubated for a further four days when 100 µL of supernatant from each well was collected and stored at -80 ^{o}C as 'Day 4' samples. 100 µL of fresh growth media was added to each well and cells were incubated for a further three days when 100 µL of supernatant from each well was collected and stored at -80 ^{o}C as 'Day 7' samples. Stored supernatants collected from the various days were simultaneously thawed at room temperature and analysed by ELISA (Murex Biotech LTD, Kent, UK) for p24 antigen production (see section 2.2.9).

To determine a $TCID_{50}$ value for FV5, the number of wells that had yielded a negative p24 ELISA value (based on the cut-off value of mean absorbance of negative controls plus 0.050) were totaled (9) and the 50 % endpoint value (M) was calculated according to $M = xk + d.[0.5 – (1/n)(r)]$, where xk = dose of the highest dilution, d = spacing between the dilutions, n = wells per dilution and r = sum of the negative responses. The M value for FV5 was calculated at $4^{-5.5}$ which converted (from base 4 to base 10) to a titre of $10^{3.3}$ and a $TCID_{50}$/mL value of 1×10^{4}.

	1	2	3	4	5	6	7	8	9	10	11	12
A	P	P	P	P	P	P	P	P	P	P	P	P
B	P	P	P	P	P	P	P	P	P	P	P	P
C	P	P	4^{-2}	4^{-4}	4^{-6}	4^{-8}	4^{-10}	4^{-12}	4^{-14}	Neg	P	P
D	P	P	4^{-2}	4^{-4}	4^{-6}	4^{-8}	4^{-10}	4^{-12}	4^{-14}	Neg	P	P
E	P	P	4^{-2}	4^{-4}	4^{-6}	4^{-8}	4^{-10}	4^{-12}	4^{-14}	Neg	P	P
F	P	P	4^{-2}	P	P	P	P	P	P	P	P	P
G	P	P	P	P	P	P	P	P	P	P	P	P
H	P	P	P	P	P	P	P	P	P	P	P	P

Figure A1: A schematic representation of the dilution series used to generated a TCID50/mL value in U87.CD4.CCR5 cells using p24 ELISA. Blocks labeled 'P' denote wells filled with PBS. The dilution series was set up from left to right across the 96 well plate starting with 4^{-2} (green blocks) and ending with 4^{-14} (far right column of blue blocks). The orange blocks represent wells that were seeded with cells but not infected.

A3 Other appendices

A3.1 HBV tiling array

To ensure that the pattern of functional siRNA production from a lhRNA construct was not specific to anti-LTR lhRNAs, a dual luciferase-based tiling array was conducted using the lhRNA HBV cassette targeted to an HBV-specific set of target tiles (see section 2.3.6). A common 3' transcription termination signal within all HBV transcripts resulted in HBV mRNAs sharing a common 3' sequence (Figure A2 upper panel). This region was chosen for the tiling array as inhibition of this target site using shRNAs had previously been shown to be effective in cell culture and *in vivo* studies (McCaffrey, Nakai*, et al.*, 2003; Carmona, Ely*, et al.*, 2006) and provided a useful alternative system to explore lhRNA efficacies. A set of five target tile vectors that overlapped each other by 14 bp and together spanned the common region within HBV (GenBank® accession number V01460) were individually cloned within the 3' UTR of the hRluc gene in psiCheck2 (see section 2.2.3). A complete target comprising the entire region was also separately cloned into psiCheck2 (Figure A2 middle and lower panels).

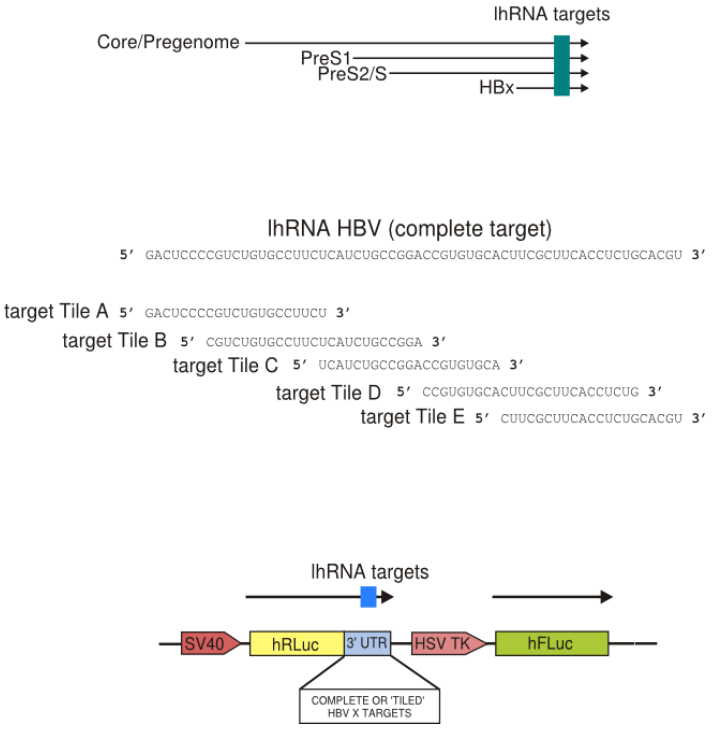


Figure A2: HBV tiling array sequences and cloning strategy. The upper panel has four arrows to depict HBV transcripts that have that same 3' ends due to a common transcription termination sequence and the lhRNA HBV target site common to all transcripts is depicted as a turquoise rectangle. In the middle panel the complete lhRNA HBV target sequence, as well as five overlapping tiled segments (target Tiles A to E), are depicted and the latter are aligned relative to the complete sequence. All six sequences were cloned into the 3' UTR of hRluc within the dual luciferase psiCheck2 vector that expressed hFluc from a separate HSK TK promoter within the same plasmid (bottom panel).

A3.2 Inhibitory efficacies of expressed anti-HIV shRNAs

For a combinatorial RNAi approach against HIV to be effective, a minimum of four target mRNAs had been shown to be effective (ter Brake, Konstantinova*, et al.*, 2006; ter Brake, Hooft*, et al.*, 2008). Many HIV mRNAs had been effectively suppressed using expressed shRNAs and six highly efficacious previously published sequences were cloned and assessed in this work using a wild type subtype C HIV isolate (FV5). The sequences included shRNAs targeted to Gag (Cave, Weinberg*, et al.*, 2006), IN (Nishitsuji, Kohara*, et al.*, 2006), the protease region of Pro (ter Brake, Konstantinova*, et al.*, 2006), Tat (Lee, Dohjima*, et al.*, 2002), Nef (Westerhout, Vink*, et al.*, 2006) and Env (ter Brake, Konstantinova*, et al.*, 2006) and were renamed in this work according to their target site. Specifically, shGag was previously denoted as 'shRNA A' and corresponded to co-ordinates 792 to 812 of the *gag* ORF from HXB2 (GenBank® accession number K03455). shIN corresponded with the N-terminal core domain of the *int* ORF from NL4-3 (GenBank® accession number AF324493). shPro was previously denoted as 'pol-protease shRNA Pol-1 and Pol-2' and correlated with positions 1910 to 1929 of the LAI viral genome (GenBank® accession number AF33819). shTat was previously designated 'site 1' and corresponded to a *tat* transcript located 20 bp downstream of a *rev* translation initiation codon. shNef was previously designated 'wt shNef' and corresponded to positions 8553 to 8571 in the *nef* gene of LAI (GenBank® accession number AF33819). shEnv corresponded to the RRE at positions 7391 to 7412 of LAI (GenBank® accession number AF33819).

A two-step PCR approach as described in section 2.2.1 was used to clone shRNAs targeted to using a universal U6 Forward primer containing a *Spe*I restriction site (5' CTA ACT AGT GGC GCG CCA AGG TCG GGC AGG AAG AGG G 3'; the *Spe*I site is underlined) and the following reverse oligonucleotide primers: shGag Reverse (5' AAA AAA GGG TGC GAG AGC GTC AAT ATT ATG GGT CAG GTA ATA TTG ACG TTC TCG CAC CCG GTG TTT CGT CCT TTC CAC AA 3'), shIN Reverse (5' AAA AAA GCC GGA GAG CAA TGG CTA GTG ATT GGG TCA GGT AAC TAA CCA TTG CTC TCC GGC GGT GTT TCG TCC TTT CCA CAA 3'), shPro Reverse (5' AAA AAA GAC AGG AGC AGA TGA TAC AGT TGG GTC AGG AAT ATA TCA TCT GCT CCT GTC GGT GTT TCG TCC TTT CCA CAA 3'), shTat Reverse (5' AAA AAA GCG GAG ACA GCG ACG AAG AGC TGG GTC AGG GCT CTT CAT CAC TAT CCC CGC GGT GTT TCG TCC TTT CCA CAA 3'), shNef Reverse (5' AAA AAA GTG CCT GGC TAG AAG CAC AAG TGG GTC AGG ATT GTA CTT CTA GCC AGG CAC GGT GTT TCG TCC TTT CCA CAA 3') and shEnv Reverse (5' AAA AAA CAG CAG GAA GCA CTA TGG GCG TGG GTC AGG AAC ACA TAG TGC TTC CTG CTG CGG TGT TTC GTC CTT TCC ACA A 3'). Candidate pTZ57R/T-based shRNA clones were screened using *BamH*I and *Eco*RI as well as sequencing as described in section 2.2.1.

The U6-driven shRNAs were expressed as sense-loop-antisense constructs with ~23 bp stem regions and 2 bp 3' overhangs required for Dicer recognition (Figure A3).

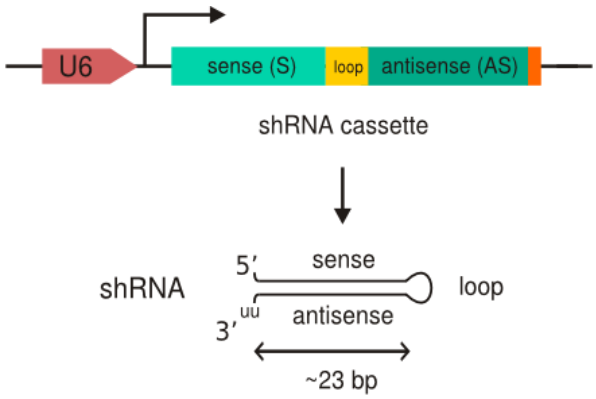


Figure A3: Design of expressed anti-HIV shRNA cassettes. Pol III (U6) cassettes were designed to express individual anti-HIV shRNAs that each folded into a sense-loop-antisense structure with a ~23 bp stem and characteristic 2 bp 3' overhangs.

In addition to the inclusion of previously published shRNAs, an aim of this work was to generate an expressed shRNA targeted to the HIV LTR and to include the guide sequence from this mimic in a combinatorial pri-miRNA mimic construct. Eight target sites within the 5' LTR of HIV were selected using two web-based programs, GeneLink RNAi Explorer (http://www.genelink.com/siRNA/shRNArev.asp) and siRNA Design (http://coh-mfold.sdsc.edu/). These included one control site within the TAR loop and seven sequences within the transcribed U5 region (Figure A4).

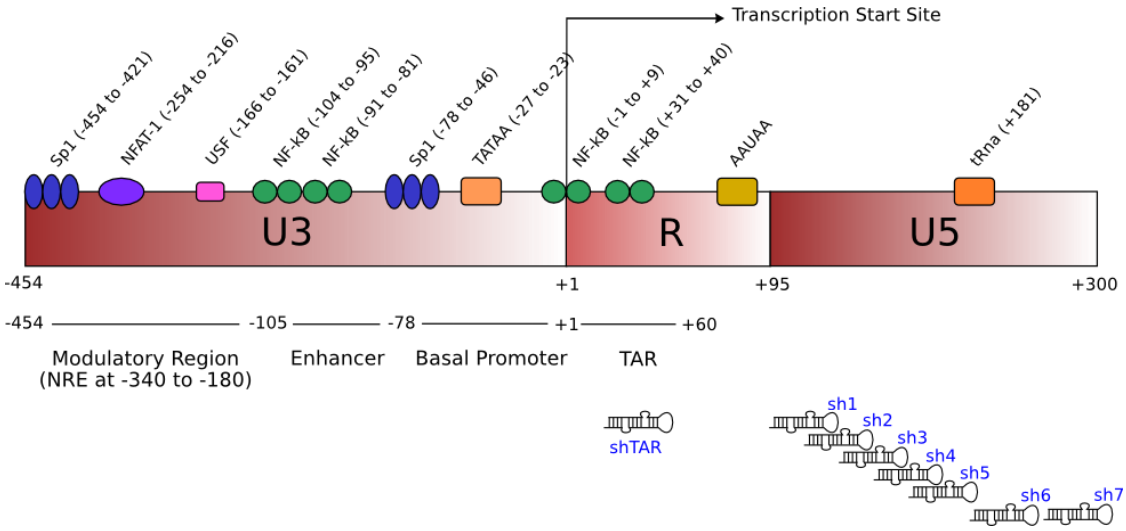


Figure A4: Diagrammatic representation of anti-HIV shRNA target sites within the 5' LTR. Eight shRNAs were designed to target the TAR loop (shTAR) or the transcribed U5 region of the HIV LTR. Positions of the shRNAs are approximate and not to scale.

A two-step PCR approach as described in section 2.2.1 was used to generate eight expressed shRNAs targeted to these regions using a universal U6 Forward primer containing a *Spe*I restriction site (5' CTA ACT AGT GGC GCG CCA AGG TCG GGC AGG AAG AGG G 3'; the *Spe*I site is underlined) and the following reverse oligonucleotide primers: shTAR (5' AAA AAA GGA GCT CTC TGG CTA TCT AGG TGG GTC AGG CGT AGA TAG CCA GAG AGC TCC GGT GTT TCG TCC TTT CCA CAA 3'), sh1 (5' AAA AAA GCC TCA ATA AAG CTT GCC TTG TGG GTC AGG CAA GAC AAG CTT TAT TGA GGC GGT GTT TCG TCC TTT CCA CAA 3'), sh2 (5' AAA AAA GAG CTT GCC TTG AGT GCT CTA AGT GGG TCA GGC TTA GAG CAC TCA AGG CAA GCT TCG GTG TTT CGT CCT TTC CAC AA 3'), sh3 (5' AAA AAA TTG AGT GCT CTA AGT AGT GTG TGT GGG TCA GGC ACA CAC TAC TTA GAG CAC TCA ACG GTG TTT CGT CCT TTC CAC AA 3'), sh4 (5' AAA AAA GCT AAG TAG TGT GTG CCC GTC TTG GGT CAG GAG ACA AGC ACA CAC TAC TTA GCG GTG TTT CGT CCT TTC CAC AA 3'), sh5 (5' AAA AAA GTA ACT AGA GAT CCC TCA GAC TGG GTC AGG GTC TGA GAG ATC TCT AGT TAC GGT GTT TCG TCC TTT CCA CAA 3'), sh6 (5' AAA AAA GCC CTT TGT GTT AGT GTG GAA TGG GTC AGG TTC CAC ACT AAC ACA AAG AGC GGT GTT TCG TCC TTT CCA CAA 3') and sh7 (5' AAA AAA GGG TGC GAG AGC GTC AAT ATT ATG GGT CAG GTA ATA TTG ACG TTC TCG CAC CCG GTG TTT CGT CCT TTC CAC AA 3'). Candidate pTZ57R/T-based shRNA clones were screened using *BamH*I and *EcoR*I as well as sequencing.

Inhibitory efficacies of the eight expressed anti-LTR shRNAs were determined using the LTRc-hFluc HEK293 stable cell line as described in sections 2.2.5 and 2.2.6 although no exogenous Tat was included in the samples so that inhibition of basal mRNA levels could be detected. Luciferase activity measured 48 hours post-transfection revealed highly significant inhibition ($\alpha = 0.99$; $p < 0.01$) for all shRNAs as well as lhRNA LTR 2 that had been included as a positive control (Figure A5). Intriguingly, shTAR mediated 66.23 % ± 1.42 % inhibition of the LTR supporting previous data that showed the TAR loop to be amenable to lhRNA-mediated suppression (see section 2.3.1). The remaining shRNAs mediated varying levels of inhibition with shRNAs 1, 3 and 5 achieving greater than 50 % target suppression (65.70 % ± 1.83 %, 60.23 % ± 5.95 % and 76.63 % ± 6.14 % respectively). As expressed sh5 showed the greatest inhibitory efficacy, it was renamed shLTR and included in the panel of anti-HIV shRNAs.

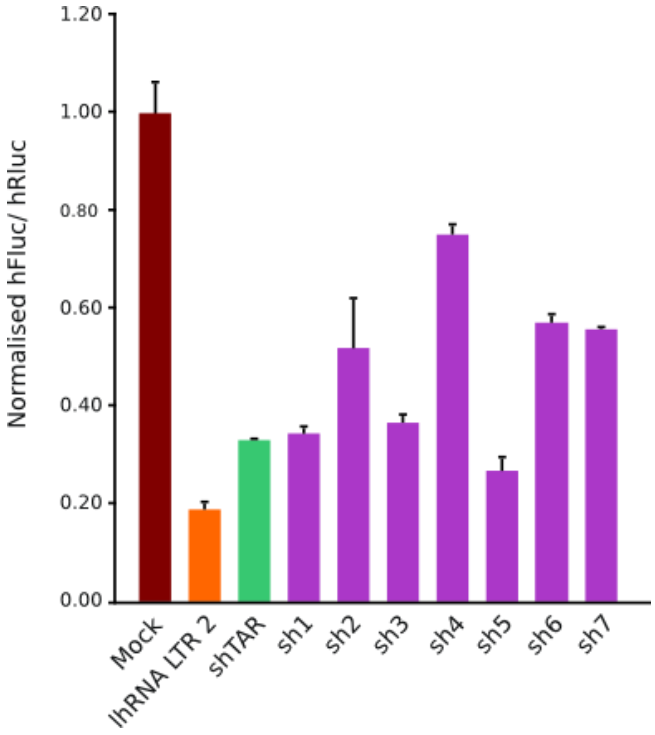


Figure A5: Inhibitory efficacies of anti-LTR shRNAs. Expressed shRNAs targeted to the TAR loop (green bar) or U5 regions (purple bars) of the HIV LTR were co-transfected with pRSV-Rluc in the LTRc-hFluc HEK293 stable cell line. Data are expressed as a ratio of hFluc to hRluc normalised to an empty vector control (mock – red bar) determined 48 hours post-transfection (n = 3, ± SEM). lhRNA LTR 2 (orange bar) was included as a positive control.

Having generated a panel of anti-HIV shRNAs that included one previously unpublished sequence (shLTR), the inhibitory efficacy of each mimic was assessed using the subtype C HIV isolate FV5 as described in section 2.2.9. U87.CD4.R5 cells were transfected in triplicate with either an empty vector control (mock) or one of the anti-HIV shRNAs before being challenged with FV5 using a $TCID_{50}$ of 1000/mL. Inhibition of HIV was assessed 4 days post-infection by determining p24 antigen levels (Figure A6). Cells that received only FV5 or empty vector (mock) treated cells showed no significant differences ($\alpha = 0.99$; $p > 0.05$) indicating that the viral isolate was infectious and the empty vector itself did not induce non-specific inhibition. All seven shRNAs showed highly significant inhibition of HIV replication as compared to mock infected cells ($\alpha = 0.99$; $p < 0.01$) with percentage suppression as follows: shRNA Env = 70.50 % ± 0.07 %, shRNA Nef = 92.00 % ± 0.01 %, shRNA Pro = 97.50 % ± 0.07 %, shRNA IN = 98.00 % ± 0.01 %, shRNA Gag = 98.00 % ± 0.01 %, shRNA LTR = 98.50 % ± 0.07 % and shRNA Tat = 98.50 % ± 0.07 %. From these data it was clear that shRNAs Nef, Pro, IN, Gag, LTR and Tat were effective at inhibiting FV5. This observation was particularly noteworthy as the isolate comprised a pool of non-isogenic viral sequences. As a result of these data the guide sequences

from shRNAs IN, Gag, LTR and Tat were selected for incorporation into a combinatorial pri-miRNA mimic strategy (see section 3.2.1).

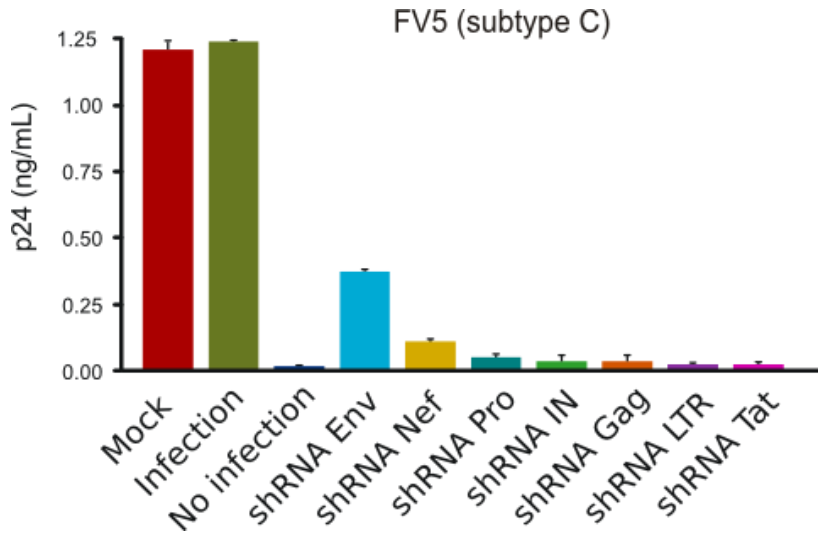


Figure A6: Challenge assay to assess anti-HIV shRNAs against a subtype C viral isolate of HIV. All seven shRNAs as well as an empty vector control (mock- red bar) were transfected into U87.CD4.R5 cells prior to challenge with viral isolate FV5 at a $TCID_{50}$ of 1000/mL. Viral Gag p24 antigen levels were determined 4 days post-infection by ELISA (450 nm). Values were converted to ng/mL using a standard curve supplied with the ELISA kit and data are expressed as p24 antigen levels in ng/mL (n = 3, ± SEM).

A3.3 Sequencing analysis of an ineffective multimeric anti-HIV pri-miRNA mimic

The multimeric pri-miRNA mimic pCi-LTR-Tat did not show levels of inhibition comparable to pCi-Gag-LTR when targeted against psi-LTR or psi-Tat target vectors (see section 3.3.3). Thus sequence analysis was performed as described in section 2.2.1 to assess if there were mutations present within the guide sequences of the mimic. From sequence alignments it was clear that numerous mismatches, insertions and deletions were present within both guide sequences of pCi-LTR-Tat (Figure A7). Disruption of three or more base pairings in the seed region of a guide strand had been shown to have a marked deleterious effect on silencing and single point mutations had been shown to reduce but not obliterate target inhibition (Zeng and Cullen, 2003). Thus the numerous mutations present within pCi-LTR-Tat suggested that poor binding of the guide sequences to their target regions negatively affected the ability of this mimic to inhibit psi-LTR and psi-Tat target reporters.

Figure A7: Alignment of sequenced and expected guide regions from pCi-LTR-Tat. The multimeric pri-miRNA mimic comprised of pri-miR30a LTR and pri-miR122 Tat was sequenced and the guides of each pri-miRNA component were aligned with the expected sequences derived from shLTR and shTat respectively. Mismatches are highlighted in bold with green boxes. Insertions and deletions are indicated with a hyphen.

A3.4 Microarray analysis of LTR-247as+7-mediated off-target gene expression

A total of 185 genes showed altered mRNA expression following 247as+7 treatment relative to GFPas controls in 1G5 cells. A description of each gene as well as the fold change in expression in the presence of LTR-247as+7 is presented in Table A3.4.1.

Table A3.4.1 Genes with significantly altered expression profiles when co-transfected with LTR-247as+7

	Parametric p-value	Fold difference of geom means	Description
1	0.0030111	1.497	myosin phosphatase-Rho interacting protein
2	0.006766	0.642	Hypothetical gene supported by BC040598
3	0.0069841	1.449	methyltransferase like 5
4	0.0089678	1.789	
5	0.0097833	0.654	Hypothetical protein FLJ25715
6	0.0097857	1.359	protein tyrosine phosphatase, receptor type, B
7	0.010494	1.594	transferrin receptor (p90, CD71)
8	0.0107781	0.725	hypothetical protein MGC21644
9	0.0108003	2.103	Helicase, lymphoid-specific
10	0.0110364	1.383	TAK1-binding protein 3
11	0.0111283	1.429	Dystonin
12	0.0112709	0.744	hypothetical protein, clone pT-Adv JuaX22
13	0.012389	1.442	ligand-gated ion channel subunit
14	0.0124899	0.66	RAB30, member RAS oncogene family
15	0.0127279	0.75	A kinase (PRKA) anchor protein 10
16	0.0128103	0.65	Ariadne homolog 2 (Drosophila)
17	0.0128271	1.413	Similar to Laminin receptor 1
18	0.0131841	0.716	acrosomal vesicle protein 1
19	0.0148795	0.621	Cytidine monophosphate-N-acetylneuraminic acid hydroxylase (CMP-N-acetylneuraminate monooxygenase)
20	0.0150231	0.72	ring finger protein 130
21	0.0150762	0.66	signal transducer and activator of transcription 4
22	0.0151328	1.434	thymidylate synthetase
23	0.0151726	0.66	
24	0.0153843	1.355	GTPase, IMAP family member 7
25	0.0156776	1.356	Similar to 40S ribosomal protein S3
26	0.0157431	0.762	Phosphoribosyl pyrophosphate synthetase-associated protein 1
27	0.0157567	1.447	AT rich interactive domain 2 (ARID, RFX-like)
28	0.0159289	1.36	La ribonucleoprotein domain family, member 2
29	0.016885	1.596	Spermidine/spermine N1-acetyl transferase-like 1
30	0.016981	0.703	headcase homolog (Drosophila)
31	0.0173331	0.747	ring finger protein 8
32	0.0173584	1.364	heat shock protein, alpha-crystallin-related, B9
33	0.0177935	0.6	Hypothetical protein FLJ13798
34	0.0177962	0.762	adaptor-related protein complex 3, sigma 2 subunit
35	0.01785	1.419	Stathmin-like 3
36	0.0182503	0.746	E1A-like inhibitor of differentiation 3
37	0.0187037	0.743	chromosome 21 open reading frame 18
38	0.0187222	0.688	DEAD (Asp-Glu-Ala-Asp) box polypeptide 50
39	0.0188578	0.758	PTK2 protein tyrosine kinase 2
40	0.0194334	0.728	Transcribed locus, weakly similar to NP_055301.1 neuronal thread protein AD7c-NTP [Homo sapiens]
41	0.01982	0.722	
42	0.0199389	0.767	tetratricopeptide repeat domain 9C
43	0.0210888	1.478	four and a half LIM domains 1
44	0.0211059	1.426	chromosome 1 open reading frame 188
45	0.0215255	1.292	Low density lipoprotein receptor-related protein 5
46	0.0215853	0.744	golgi autoantigen, golgin subfamily a, 8A /// golgi autoantigen, golgin subfamily a, 8B
47	0.0217775	1.405	WD repeat and SOCS box-containing 2
48	0.0223245	0.689	bruno-like 6, RNA binding protein (Drosophila)
49	0.02246	1.652	Nuclear casein kinase and cyclin-dependent kinase substrate 1
50	0.0224639	1.312	ELOVL family member 5, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)
53	0.0239443	0.679	interleukin 27 receptor, alpha
54	0.0243069	1.291	KIAA1546 protein
55	0.0243548	1.388	coagulation factor C homolog, cochlin (Limulus polyphemus)
56	0.0248244	1.379	Rho GTPase activating protein 21
57	0.0250849	0.76	Poly (ADP-ribose) polymerase family, member 8
58	0.0251804	0.582	cysteine sulfinic acid decarboxylase
59	0.0260157	0.649	cold inducible RNA binding protein
60	0.026539	1.391	Hepatoma-derived growth factor, related protein 3
61	0.0267377	0.746	
62	0.0268886	0.795	amyloid beta (A4) precursor protein-binding, family A, member 2 (X11-like)
63	0.0280254	0.733	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26B
64	0.028477	1.393	
65	0.028652	0.796	caspase recruitment domain family, member 9
66	0.0290164	0.701	Antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5
67	0.0291857	0.694	amine oxidase, copper containing 2 (retina-specific)
68	0.0292046	0.601	hypothetical protein MGC13102
69	0.0297263	0.758	zinc finger protein 302
70	0.0297709	0.723	Full length insert cDNA clone YU26H08
71	0.0301268	1.406	outer dense fiber of sperm tails 1
72	0.030495	0.618	Hypothetical protein DJ328E19.C1.1
73	0.0309534	0.663	spectrin repeat containing, nuclear envelope 1
74	0.0309662	0.769	Transcribed locus
75	0.0312735	0.697	Scm-like with four mbt domains 1
76	0.0316314	0.665	v-jun sarcoma virus 17 oncogene homolog (avian)
77	0.0316373	1.823	nucleoporin 43kDa
78	0.0323231	0.686	v-fos FBJ murine osteosarcoma viral oncogene homolog
79	0.0327273	1.9	Chromosome 5 open reading frame 13
80	0.0329129	1.342	
81	0.0330514	0.69	
82	0.0331589	1.304	amyloid beta (A4) precursor protein-binding, family A, member 1 (X11)
83	0.0332677	0.723	CDNA FLJ25058 fis, clone CBL04608
84	0.0334247	1.383	Chromosome 9 open reading frame 93
85	0.0335291	1.406	Transmembrane protein 63A
86	0.0338019	0.792	Homo sapiens, clone IMAGE:4049873, mRNA
87	0.0339127	0.643	hypothetical protein FLJ43663
88	0.0339264	0.75	Hypothetical protein LOC222070
89	0.0340592	0.79	renal tumor antigen
90	0.0341049	1.333	Arsenic transactivated protein 1
91	0.0346982	0.782	Chromosome 20 open reading frame 111
92	0.0349171	1.739	Phosphoglycerate kinase 1
93	0.0351885	0.723	spastic paraplegia 7, paraplegin (pure and complicated autosomal recessive)
94	0.0357783	0.759	hypothetical protein LOC202347
95	0.0359486	1.278	Thyroid hormone receptor associated protein 2
96	0.0362898	1.38	Zinc finger protein 81 (HFZ20)
97	0.0365124	1.282	nucleosomal binding protein 1
98	0.036595	1.289	class II, major histocompatibility complex, transactivator
99	0.0366169	0.726	Stromal interaction molecule 1
100	0.0368442	0.753	La ribonucleoprotein domain family, member 5
101	0.0368659	0.649	Transcribed locus, weakly similar to XP_518535.1 PREDICTED: similar to dJ108C2.1.4 (MCM3 minichromosome maintenance deficient 3 (S. cerevisiae), variant 4) [Pan troglodytes]
102	0.0369158	1.296	COX11 homolog, cytochrome c oxidase assembly protein (yeast)
103	0.0370055	0.585	
104	0.0370065	0.738	presenilin enhancer 2 homolog (C. elegans) /// protein F25965
105	0.0370295	1.236	disrupter of silencing 10
106	0.0373558	1.455	hypothetical protein FLJ39827
107	0.037559	0.806	chromosome 10 open reading frame 21
108	0.0376322	0.748	Transcribed locus, weakly similar to NP_055301.1 neuronal thread protein AD7c-NTP [Homo sapiens]
109	0.0377213	1.275	hypothetical protein FLJ20209
110	0.0381347	1.268	Solute carrier family 1 (glutamate/neutral amino acid transporter), member 4
111	0.0386032	0.759	Cyclin G2
112	0.038637	0.783	ELG protein
113	0.0389528	0.781	cullin 2
114	0.0395941	0.733	aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III)
115	0.0396401	1.315	frizzled homolog 8 (Drosophila)
116	0.0396465	1.713	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)
117	0.0396706	0.814	stromal interaction molecule 2
118	0.0397029	1.246	hypothetical protein LOC284632
119	0.0400455	0.734	chromosome 10 open reading frame 76
120	0.0401586	0.751	CDNA FLJ11958 fis, clone HEMBB1000996
121	0.0402688	0.705	choline kinase beta /// carnitine palmitoyltransferase 1B (muscle)
122	0.0403161	0.757	ATP-binding cassette, sub-family C (CFTR/MRP), member 5
123	0.040337	1.31	membrane cofactor protein (CD46, trophoblast-lymphocyte cross-reactive antigen)
124	0.0403996	1.265	TAF4b RNA polymerase II, TATA box binding protein (TBP)-associated factor, 105kDa
125	0.0404443	0.778	
126	0.0405492	0.727	GTPase, very large interferon inducible 1
127	0.0405574	1.256	BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast)
128	0.0406652	0.806	CDNA: FLJ23131 fis, clone LNG08502
129	0.0408543	0.746	MRNA; cDNA DKFZp586O1318 (from clone DKFZp586O1318)
130	0.0408773	1.29	G patch domain containing 4
131	0.041044	1.268	kelch domain containing 6
132	0.0410792	0.751	A kinase (PRKA) anchor protein 1
133	0.0412473	0.739	reticulon 1
134	0.0417579	0.781	Muscleblind-like 2 (Drosophila)
135	0.0418332	0.755	SKI-like
136	0.042444	0.719	KIAA1671 protein
137	0.042554	1.356	low density lipoprotein-related protein 12
138	0.0427522	1.37	dystrobrevin, alpha
139	0.0428779	1.366	epoxide hydrolase 2, cytoplasmic
140	0.0431762	0.756	ring finger protein 165
141	0.0436256	0.765	Transcribed locus, weakly similar to XP_209041.2 PREDICTED: similar to KIAA1503 protein [Homo sapiens]
142	0.0439321	0.756	chromosome X open reading frame 24
143	0.0441886	0.745	Hypothetical LOC285711
144	0.0444402	0.781	homeo box B7
145	0.0444805	1.316	Forkhead box O3A
146	0.0446401	0.784	similar to envelope protein
147	0.0447444	1.235	cytochrome P450, family 2, subfamily S, polypeptide 1
148	0.0447516	1.8	voltage-dependent anion channel 3
149	0.0448304	0.651	Peroxisome biogenesis factor 26
150	0.0448889	0.701	KIAA1128
153	0.0453544	1.279	immunity-related GTPase family, Q1
154	0.0453697	1.333	KIAA0513
155	0.0456017	0.792	Hect domain and RLD 6
156	0.046216	1.363	protein tyrosine phosphatase, non-receptor type 11 (Noonan syndrome 1)
157	0.0463012	0.706	hypothetical protein MGC16121
158	0.0464753	0.818	
159	0.0464877	0.812	Homo sapiens, clone IMAGE:5184785, mRNA
160	0.0467609	0.764	elongation protein 4 homolog (S. cerevisiae)
161	0.046944	0.792	myosin XVB pseudogene
162	0.0472286	1.253	MRNA full length insert cDNA clone EUROIMAGE 592473
163	0.0473195	0.77	solute carrier family 25 (mitochondrial carrier; Graves disease autoantigen), member 16
164	0.0473611	1.371	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)
165	0.0477081	0.783	CDNA clone IMAGE:4827117
166	0.0478196	0.771	Fibroblast growth factor 18
167	0.0478546	0.695	Homo sapiens, clone IMAGE:4449283, mRNA
168	0.0479095	0.791	inositol polyphosphate-5-phosphatase, 72 kDa
169	0.0479828	1.259	hypothetical protein FLJ36874
170	0.0482963	0.809	PR domain containing 1, with ZNF domain
171	0.0483703	1.286	pyruvate dehydrogenase kinase, isoenzyme 1
172	0.0483979	1.288	Laminin, alpha 2 (merosin, congenital muscular dystrophy)
173	0.0485947	1.439	Helicase with zinc finger
174	0.0486874	1.298	
175	0.0487384	1.302	nuclear transcription factor Y, beta
176	0.0487826	0.757	RAN binding protein 9
177	0.0489461	0.818	
178	0.0491396	1.305	Nucleoporin 88kDa
179	0.049211	0.77	Formin binding protein 4
180	0.0492159	1.28	Hypothetical LOC400602
181	0.0493224	1.234	solute carrier family 28 (sodium-coupled nucleoside transporter), member 1
182	0.0493455	0.722	Cation channel, sperm associated 2
183	0.0495444	0.75	SMAD, mothers against DPP homolog 7 (Drosophila)
184	0.0495948	1.216	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2
185	0.0498219	0.583	Smu-1 suppressor of mec-8 and unc-52 homolog (C. elegans)