

https://dspace.org/

Inhibition of Human Immunodeficiency virus replication through small RNA-induced gene silencing of HIV-1 Tat specific factor 1

Green, Victoria Andress

2012-02-14
http://hdl.handle.net/10539/11306

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

INHIBITION OF HUMAN IMMUNODEFICIENCY VIRUS REPLICATION THROUGH SMALL RNA-INDUCED GENE SILENCING OF HIV-1 TAT SPECIFIC FACTOR 1

VICTORIA ANDRESS GREEN

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 27 June 2011

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

The HIV-1 pandemic continues unabated. Although treatments exist that can substantially alleviate the morbidity and mortality associated with HIV, there is still a need for improved anti-HIV treatments that reduce toxicities and administration frequency and mediate sustained inhibition of viral replication. Given the high mutability and variability of the virus, a strategy that is garnering increasing focus is the targeting of host factors that the virus requires to replicate, so-called HIV-dependency factors (HDFs). It is hoped this will reduce the emergence of viral drug resistance. A number of genome-wide screens have been performed to identify HDFs, although many remain to be validated, particularly in relevant cells lines.

An objective of this thesis was to validate three host factors as HDFs, in both TZM-bl reporter and T cell-derived cell lines, and to examine their potential as anti-HIV-1 therapeutic targets through exploitation of the cellular gene silencing pathway, RNA interference (RNAi). These were HIV-1 Tat specific factor 1 (HTATSF1), DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked (DDX3X) and SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 (SMARCB1), selected because they had been previously implicated in HIV-1 pathogenesis. The well-characterised HDF, PC4 and SFRS1 interacting protein 1 (PSIP1)/lens epithelium-derived growth factor (LEDGF)/p75, was included in the study as a positive control. Cassettes expressing short hairpin RNAs (shRNAs) targeting the four host proteins were generated, although shRNAs did not suppress endogenous *ddx3x* mRNA levels. The ability of shRNAs to inhibit HIV-1 replication in the reporter cell line, TZM-bl, was examined. These HeLa-derived cells are permissive for R5-tropic HIV-1 infection and contain an integrated luciferase gene driven by the viral promoter. shRNAs mediated a dose-dependent inhibition of luciferase activity in cells infected with a HIV-1 subtype B molecular clone and, although production of the viral protein p24 was unaltered, infectious particle production was decreased in cells treated with a shRNA suppressing HTATSF1. Little effect was observed with a shRNA targeting SMARCB1, suggesting that this may not function as an HDF under these conditions. No effect on infectious particle production was seen with the shRNA targeting PSIP1, which was a result of the long half-life of this protein, highlighting a limitation of using such reporter systems for HDF validation. Importantly, shRNAs were not associated with any cytotoxic effects in TZM-bl cells.

Whether HTATSF1 is a potential therapeutic target was interrogated further in the more relevant T cell-derived SupT1 cell line. Lentiviruses were used to generate populations where >90% had one copy of the integrated shRNA expression cassette. Replication of the subtype B molecular clone p81A-4 was significantly inhibited in the shH1-expressing SupT1 cell line, which targets HTATSF1, for over 14 days post-infection, although inhibition was not as pronounced as

that observed in the shP1-expressing SupT1 cell line, which targets PSIP1. In contrast to a previous report, no change in the ratio of unspliced to singly- or multiply-spliced HIV-1 transcripts were detected in shH1-expressing SupT1 cells, suggesting that HTATSF1 does not function as a splicing cofactor in this system. A slight rebound in p24 levels at 14 days post-infection was accompanied by increased HTATSF1 expression and a decrease in the percentage of cells with transgene expression in the population. In addition, there was a slight decrease in shH1-derived guide strand expression, but no change in transcription rates of the *htatsf1* gene, suggesting that cells within the population with shH1 expression and HTATSF1 suppression may have a growth disadvantage. Thus, although this work demonstrates for the first time that HTATSF1 functions as an HDF in T cell-derived SupT1 cells, it may not constitute a viable therapeutic target.

A second objective of this thesis was to examine the feasibility of transcriptional gene silencing (TGS) of HDFs as an anti-HIV strategy. TGS is a small RNA-induced gene silencing pathway that operates through chromatin remodelling with the potential to mediate long-term silencing of gene expression. Thus, its application may reduce the frequency of drug administration and associated toxicities. Short interfering RNAs (siRNAs) targeting the *htatsf1* promoter were able to reduce target mRNA expression, which was accompanied by decreased *htatsf1* transcription rates in HEK293T cells, suggesting silencing via a TGS mechanism. The *htatsf1* silencing inhibited infectious HIV-1 particle production from TZM-bl cells. This work provides proof of principle that TGS induction at a HDF may inhibit HIV-1 replication.

siRNAs targeting the *ddx3x* promoter did not induce TGS. To examine whether gene susceptibility to TGS may be influenced by promoter architectures, 49 promoter features were examined for enrichment in genes at which small RNA-induced TGS has been reported. Initially, the TGS group was compared to a random set of 2,000 promoters and then all other promoters in the genome. To control for gene activation, two further analyses were performed comparing the TGS group features to those from promoters active in the THP-1 cell line and housekeeping genes. Whilst difficult to ascribe differences between the TGS group and the control groups to anything beyond a variation in the proportion of active genes within each group, there was enrichment for certain promoter features that are independent of activity; the TGS group was characterised by broad transcription start regions, high CpG content and a single expression profile. Moreover, the fraction of promoters with reported non-coding RNA overlap was greater in the TGS group than the control groups. Thus, there is some evidence that a number of promoter features are associated with TGS susceptibility. It is hoped this novel analysis will facilitate selection of future TGS targets, including HDFs. In summary, the work presented in this thesis paves the way for development of improved anti-HIV therapies involving HDF-targeted TGS-based gene therapies that mediate sustained inhibition of the virus.

ACKNOWLEDGEMENTS

This thesis wouldn't have been possible without the guidance of my supervisors, Associate Prof Marco S. Weinberg and Prof Patrick Arbuthnot. Firstly, thank you for taking me on. I appreciate that I could always talk to you both openly and that you continually encouraged me. Thank you for allowing me to take up the opportunities that came my way: to visit Scripps and the NCBI and to attend many meetings. You both taught me a great deal that I will carry forward (if anyone gives me a job to carry things forward to).

I would like to thank Kevin V. Morris for hosting me at The Scripps Research Institute for two months in 2009. I will never doubt the ability of men to multi-task again. Thanks for teaching me so much and giving me the opportunity to do science by the beach and, most importantly, for lending me a tent so I could go see bears.

Thanks to David Landsman for hosting me at the NCBI in May 2010 and collaborating with me since. I am so grateful that someone in your position was so generous, not just by supporting my stay in the US, but also with your time and contacts. Working with you really boosted my confidence, except for when you laughed at my Python book.

Loren Hansen, I cannot thank you enough. You were such a pleasure to work with and should win a patience award for not flinching every time I asked you to retrieve stuff from yet another database. Thanks for all the helpful discussions too.

I also need to thank Drs Melinda Suchard and Elizabeth Mayne for their friendship and rescuing me financially in my first year with a part-time job. Also, of course, for the fact that my only published article has come from my work with you guys and not my thesis!

Thanks too to Dr Samantha Barichievy (Mrs T) and Liam Thompson (Mr T) for advice, both scientific and personal, lots of lovely dinners and keeping my hubby company when I was away. You are my ambassadors of keepin'-it-real when Piled Higher and Deeper.

Finally, I'd like to thank my family. My husband, Chris Green, has been a rock throughout. You are wonderful, especially for trying to understand what I do and persevering with a wife who gave the distinct impression of being a raving lunatic at times. I'd like to thank my parents for doing what they were told – never, ever asking how my PhD is going. That must have been torture for you, especially as I'm so far away, but still you never asked! Most importantly, I have to thank my woogie. Conker, you are the cutest thing in the world and every time I came home from a nasty day, one waggle from you set everything right again. Despite the trials and tribulations, my PhD was always eminently doable with Chris and Conker around.

PREFACE

The following articles and conference presentations arose from work presented in this thesis.

Publications

Green, V. A. and Weinberg, M. S. Small RNA-induced transcriptional gene regulation in mammals: mechanisms, therapeutic applications and scope within the genome. *Progress in Molecular Biology and Translational Science*, in press.

Green, V. A., Arbuthnot, P. A. and Weinberg, M. S. Persistent suppression of HIV-1 replication in CD4+ SupT1 cell lines using lentivirus-delivered short hairpin RNAs targeted to the host HIV-1 dependency factor HTATSF1. Manuscript in preparation.

Green, V. A., Hansen, L. and Landsman, D. Promoter features associated with susceptibility to small RNA-induced transcriptional gene silencing. Manuscript in preparation.

Conference presentations

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Persistent suppression of HIV-1 replication in CD4+ T cells using lentivirus-delivered short hairpin RNAs targeted to the host HIV-1 dependency factor HTATSF1. (2011) EMBO HIV/AIDS lecture course, Stellenbosch, South Africa.

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Persistent suppression of HIV-1 replication in CD4+ T cells using lentivirus-delivered short hairpin RNAs targeted to the host HIV-1 dependency factor HTATSF1. (2010) University of the Witwatersrand Crossfaculty Postgraduate Symposium, Johannesburg, South Africa.

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Inhibition of human immunodeficiency virus type 1 replication through expression of RNAi effectors targeting host dependency factors. (2010) University of the Witwatersrand Faculty of Health Sciences Research Day, Johannesburg, South Africa.

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Promoter architecture and susceptibility to RNA-induced transcriptional gene silencing. (2010) Cold Spring Harbor Laboratory 75th Symposium on Nuclear Organization and Function, Long Island, USA.

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Inhibition of human immunodeficiency virus type 1 replication through gene silencing of host dependency factors. (2010) Second Advanced Summer School in Africa: Molecular mechanisms of viral infection and propagation, Hermanus, South Africa.

Green, V. A., Morris, K. V., Arbuthnot, P. and Weinberg, M. S. Inhibition of human immunodeficiency virus type 1 replication through gene silencing of host dependency factors. (2010) South African Society for Biochemistry and Molecular Biology, Bloemfontein, South Africa.

Green, V. A., Arbuthnot, P. and Weinberg, M. S. Inhibition of human immunodeficiency virus type 1 replication through gene silencing of host dependency factors. (2009) EMBO Annual Meeting, Amsterdam, The Netherlands.

The following articles arose from work not presented in this thesis.

Other publications

Green, V. A., Munshi, S. U., Marakalala, M. J. and Mourão, M. M. Molecular mechanisms of viral infection and propagation: an overview of the second advanced summer school in Africa. (2010) *IUBMB Life* **62(8)**: 573-583.

Suchard, M. S., Mayne, E., **Green, V. A.**, Shalekoff, S., Donninger, S. L., Stevens, W. S., Gray, C. M. and Tiemessen, C. T., FOXP3 expression is upregulated in CD4+ T cells in progressive HIV-1 infection and is a marker of disease severity. (2010) *PLoS One* **5(7):** e11762.

TABLE OF CONTENTS

DECLARATION ii
ABSTRACT iii
ACKNOWLEDGEMENTS v
PREFACE vi
Publications vi
Conference presentations vi
Other publications vii
TABLE OF CONTENTS viii
LIST OF FIGURES xiv
LIST OF TABLES xvi
LIST OF ABBREVIATIONS xvii

CHAPTER 1: General introduction 1
1.1 Introduction 1
1.2 Human immunodeficiency virus 2
1.2.1 HIV origins 2
1.2.2 HIV genome organisation 3
1.2.3 HIV life cycle 4
1.2.4 HIV pathogenesis 8
1.2.5 HIV variability and treatment 9
1.3 RNA interference 11
1.3.1 miRNA biogenesis in mammals 12
1.3.2 Endo-siRNA and non-canonical miRNA biogenesis 13
1.3.3 RNAi mimics 14
1.3.4 RNAi screens for the identification of HDFs 16
1.3.5 RNAi-based validation of HDFs 19
1.3.6 RNAi-based gene therapy 20
1.4 Transcriptional gene silencing 23
1.4.1 Epigenetic changes associated with TGS 24
1.4.2 Mechanism of small RNA-induced TGS in mammals 25
1.4.3 Transcriptional interference 27
1.4.4 Transcriptional gene activation 28
1.4.5 TGS-based gene therapy 30
1.4.6 TGS-based modulation of viruses 31
1.5 HIV-dependency factors 32
1.5.1 HTATSF1 33
1.5.2 DDX3X 34
1.5.3 SMARCB1/INI1/SNF5 35
1.5.4 PSIP1/LEDGF/p75 37
1.6 Thesis objectives 39

CHAPTER 2: Inhibition of HIV-1 replication with host factor-targeting shRNAs in TZM-bl reporter cells 41
2.1 Introduction 41
2.2 Materials and methods 43
2.2.1 Design of cellular factor-targeting shRNAs 43
2.2.2 Cloning expressed shRNAs 43
2.2.3 Cloning luciferase reporter vectors 46
2.2.4 Mammalian cell culture 48
2.2.5 DNA plasmid transfections 48
2.2.6 Dual luciferase reporter assay 48
2.2.7 Quantitative reverse transcription polymerase chain reaction (qRT-PCR) measurement of mRNA concentrations 49
2.2.8 Challenge with HIV-1 molecular clone pNL4-3.Luc.R$^-$.E$^-$ 51
2.2.9 Detection of processed shRNA guide strands using polyacrylamide gel electrophoresis (PAGE) and Northern blot analysis 51
2.2.10 HIV-1 reporter assay in TZM-bl cells 52
2.2.11 Detection of cellular factors by SDS-PAGE and Western blot analysis 53
2.2.12 Detection of early apoptotic TZM-bl cells 55
2.2.13 GFP splicing assay 55
2.2.14 MTT assay 56
2.2.15 Cell viability by fluorescence microscopy 56
2.2.16 qRT-PCR measurement of interferon-β mRNA concentrations 56
2.2.17 Statistical analyses 57
2.3 Results 58
2.3.1 shRNAs suppress cellular targets in HEK293 cells 58
2.3.2 shRNAs can reduce pNL4-3.Luc.R$^-$.E$^-$ luciferase activity 60
2.3.3 shRNAs suppress cognate targets in TZM-bl cells by an RNAi mechanism 62
2.3.4 shRNAs inhibit HIV-1 Tat-induced luciferase activity in TZM-bl reporter cells 63
2.3.5 HTATSF1-targeted shRNA inhibits HIV-1 infectious particle production 66
2.3.6 shRNA-mediated HTATSF1 suppression is rapid but PSIP1 and SMARCB1 have long half-lives 68
2.3.7 shRNA-mediated host factor suppression is not cytotoxic 69
2.3.8 U6-driven shRNA expression in TZM-bl cells is not cytotoxic and does not activate an immune response 71
2.4 Discussion 73

CHAPTER 3: Persistent suppression of HIV-1 replication in a CD4+ SupT1 cell line using lentivirus-delivered shRNAs targeted to HTATSF1 78
3.1 Introduction 78
3.2 Materials and methods 80
3.2.1 Mammalian cell culture 80
3.2.2 Selection of lentivector for shRNA expression cassettes 80

3.2.3 Cloning lentivector-expressed shRNAs 81
3.2.4 Dual luciferase reporter assay 81
3.2.5 Generation of lentiviruses 81
3.2.6 Biological transduction titre of shRNA-expressing lentiviruses 82
3.2.7 Generation of shRNA-expressing SupT1 cell lines 83
3.2.8 Detection of processed shRNA guide strands using PAGE and Northern blot analysis . 83
3.2.9 SupT1 cell proliferation rates by cell counting 84
3.2.10 SupT1 cell proliferation rates by DNA quantification 84
3.2.11 Assessment of p81A-4 replication in shRNA-expressing SupT1 cell lines 85
3.2.12 qRT-PCR measurement of mRNA concentrations 85
3.2.13 Detection of HTATSF1 by SDS-PAGE and Western blot analysis 86
3.2.14 Quantification of GFP+ SupT1 cells in shRNA-expressing cell lines by flow cytometry 86
3.2.15 Nuclear run-on assay to determine *htatsf1* transcription rates in SupT1 cell lines 86
3.2.16 Quantification of p81A-4 transcript classes by qRT-PCR 88
3.2.17 Statistical analyses 89
3.3 Results 90
3.3.1 Selection of the lentivector pLVTH for shRNA expression 90
3.3.2 shRNAs knock down cognate targets when expressed from a lentivector 91
3.3.3 Generation of stable shRNA-expressing SupT1 cell populations 91
3.3.4 Proliferative capacity is unaltered by shRNA expression in SupT1 cells 93
3.3.5 A shRNA targeting HTATSF1 mediates sustained inhibition of HIV-1 replication 94
3.3.6 HTATSF1 levels increase through reduced shH1 expression 96
3.3.7 HTATSF1 suppression does not alter the ratio of HIV-1 transcript classes in SupT1 cells 98
3.4 Discussion 101

CHAPTER 4: Inhibition of HIV-1 replication through transcriptional gene silencing of HTATSF1 106
4.1 Introduction 106
4.2 Materials and methods 108
4.2.1 Mapping the *htatsf1* promoter 108
4.2.2 Design of promoter-targeting siRNAs 109
4.2.3 Generation of siRNAs by *in vitro* transcription 111
4.2.4 Transfection of siRNAs 112
4.2.5 Mammalian cell culture 113
4.2.6 Screening siRNAs for inhibition of target expression 113
4.2.7 Detection of promoter-associated transcripts 113
4.2.8 Examination of the role of Ago proteins 114
4.2.9 Nuclear run-on assay to determine *htatsf1* transcription rates 115
4.2.10 Quantification of p81A-4 infectious particle production 115
4.2.11 Statistical analyses 116
4.3 Results 117

4.3.1 Promoter-targeted siRNAs inhibit target HDF expression 117
4.3.2 siH15-mediated *htatsf1* silencing is attenuated by TSA in HEK293T cells..... 118
4.3.3 Bidirectional promoter-associated transcripts are present at the siH15 target site.... 120
4.3.4 Ago1 and Ago2 suppression does not attenuate siH15-mediated *htatsf1* silencing ... 122
4.3.5 Promoter-targeted siRNAs modulate the *htatsf1* transcription rate..... 124
4.3.6 siH15-mediated HTATSF1 suppression inhibits HIV-1 infectious particle production . 125
4.4 Discussion..... 126

CHAPTER 5: Promoter features associated with susceptibility to small RNA-induced transcriptional gene silencing 130
5.1 Introduction 130
5.2 Materials and methods..... 134
5.2.1 Selection of small RNA-induced transcriptionally-silenced genes and TSR assignment 134
5.2.2 Promoter feature data sources and quantification..... 134
5.2.3 Assignment of TSRs to the THP-1 dataset 137
5.2.4 Assignment of TSRs to the housekeeping gene dataset 137
5.2.5 Classification of *htatsf1* and *ddx3x* promoter features..... 138
5.2.6 Statistical analyses 138
5.2.7 Statement on contributions..... 139
5.3 Results..... 140
5.3.1 Identification of small-RNA induced transcriptionally-silenced TSRs..... 140
5.3.2 The TGS group are enriched for promoter features associated with active transcription compared with 2000 random TSRs 142
5.3.3 The TGS group are enriched for promoter features associated with active transcription compared with all other human TSRs 145
5.3.4 The TGS group show reduced TSR length and number of TSCs per TSR compared with active TSRs in the THP-1 cell line 147
5.3.5 The TGS group are enriched for promoter features associated with reduced activity compared with housekeeping gene TSRs 149
5.3.6 The fraction of TSRs with ncRNA overlap is greatest in the TGS group 151
5.3.7 Classification of *htatsf1* and *ddx3x* promoter features..... 152
5.4 Discussion..... 153

CHAPTER 6: General discussion 157
5.1 Summary of findings..... 157
5.2 Viral escape 158
5.3 Application of HDF-targeted TGS 160
5.4 Delivery of HDF-targeted TGS-based gene therapy..... 162
5.5 Modulation of RNAi by HIV..... 164
5.6 Concluding remarks 165

REFERENCES..... 167

APPENDICES **197**

A1 Standard laboratory practices **197**

A1.1 Bacterial procedures **197**

A1.1.1 Bacterial growth 197

A.1.1.2 Preparation of transformation-competent DH5α 197

A1.1.3 Transformation of DH5α 197

A1.1.4 Storage of transformed DH5α 198

A1.2 DNA procedures **198**

A1.2.1 Small-scale plasmid DNA preparation 198

A1.2.2 Bulk preparation of plasmid DNA 198

A1.2.3 DNA gel electrophoresis 199

A1.2.4 DNA extraction from agarose gels 199

A1.2.5 Calcium chloride method for DNA transfection 200

A1.3 RNA procedures **200**

A1.3.1 RNA extraction from mammalian cells 200

A1.3.2 DNase treatment and reverse transcription 200

A1.3.3 Preparation of radiolabelled Decade™ Marker 201

A1.3.4 Quantitative polymerase chain reaction (qPCR) 201

A1.4 Viral procedures **201**

A1.4.1 Propagation of HIV-1 subtype C isolate FV5 201

A.1.4.2 Generation of HIV-1 subtype B molecular clone p81A-4 202

A.1.4.3 Median tissue culture infectious dose ($TCID_{50}$) from luciferase output 202

A1.4.4 HIV-1 p24 ELISA 203

A.1.4.5 Median tissue culture infectious dose ($TCID_{50}$) from p24 output 203

A1.5 Other standard laboratory procedures **204**

A1.5.1 Bicinchoninic acid (BCA) method for protein quantification 204

A2 Transcriptional gene silencing of HIV-dependency factors **205**

A2.1 Introduction **205**

A2.2 Materials and methods **206**

A2.2.1 Design of target sites within HDF promoters 206

A2.2.2 Cloning of expressed asRNAs 206

A2.2.3 Design and generation of siRNAs 207

A2.2.4 Screening for suppression of target gene expression 208

A2.2.5 Detection of promoter-associated transcripts 209

A2.2.6 Examination of the role of Ago proteins 209

A2.2.7 Chromatin immunoprecipitation (ChIP) assay to examine epigenetic changes at siRNA-targeted regions 210

A2.2.8 Nuclear run-on assay to determine targeted HDF transcription rates 212

A2.2.9 Alignment of siRNA target sites to the genome 212

A2.2.10 Statistical analyses 212

A2.3 Results 213
A2.3.1 asRNAs do not inhibit target gene expression 213
A2.3.2 siRNAs inhibit *psip1* expression 214
A2.3.3 Transcripts overlap siRNA target sites 215
A2.3.4 Ago2 suppression attenuates *psip1* silencing 216
A2.3.5 siP10 does not induce significant epigenetic changes at the siRNA target site 217
A2.3.6 siP10 does not alter the rate of *psip1* transcription 218
A2.3.7 siRNAs do not induce TGS of *ddx3x* 219
A2.4 Discussion 221

Ethics clearance certificate 223

LIST OF FIGURES

Figure 1.1: HIV-1 genome organisation……4

Figure 1.2: The basic steps of the HIV life cycle……6

Figure 1.3: The HIV integrase reaction……7

Figure 1.4: miRNA biogenesis in mammals and RNAi mimics……15

Figure 1.5: Proposed mechanism of TGS……27

Figure 1.6: Potential mechanisms of TGA……29

Figure 1.7: Domains of HTATSF1……34

Figure 1.8: Domain structure of PSIP1/LEDGF/p75 and SMARCB1/INI1/SNF5 and interactions with HIV-1 IN……38

Figure 2.1: Schematic demonstrating the two-step PCR approach used to generate shRNA expression cassettes……44

Figure 2.2: Schematic of generated reporter constructs containing cognate shRNA target sequences for each cellular factor……47

Figure 2.3: Schematic representation of the dual luciferase reporter assay……49

Figure 2.4: Constructed cellular factor-targeting shRNA expression cassettes……58

Figure 2.5: shRNAs significantly knock down their cognate targets in a dual luciferase reporter assay……59

Figure 2.6: Not all shRNAs reduce levels of target mRNAs……60

Figure 2.7: shRNAs may inhibit pNL4-3.Luc.R^-.E^- luciferase activity……61

Figure 2.8: shRNAs knock down target mRNAs by an RNAi mechanism……63

Figure 2.9: Schematic of the HIV-reporter ability of TZM-bl cells……64

Figure 2.10: shRNAs inhibit Tat-induced luciferase activity in a dose-dependent manner……65

Figure 2.11: shRNAs mediate similar effects on different subtypes of HIV-1……66

Figure 2.12: Schematic of the TZM-bl reporter assay for HIV-1 infectious particle production……67

Figure 2.13: Host factor-targeting shRNAs do not inhibit HIV-1 p24 production……67

Figure 2.14: HTATSF1-targeting shRNA inhibits HIV-1 infectious particle production……68

Figure 2.15: The half-lives of targeted host proteins differ……69

Figure 2.16: shRNA-mediated target suppression does not induce apoptosis……69

Figure 2.17: HTATSF1 suppression does not effect splicing of a GFP reporter……70

Figure 2.18: shRNA expression does not alter cell viability……71

Figure 2.19: shRNAs do not induce an immune response……72

Figure 3.1: Schematic showing plasmids used to generate lentiviruses capable of delivering shRNA expression cassettes to SupT1 cells……82

Figure 3.2: SupT1 GFP reporter gene expression is greatest when transduced with pLVTH-derived lentivirus……90

Figure 3.3: shRNAs knock down their cognate targets when expressed from a lentivector……91

Figure 3.4: Biological transduction titre of shRNA-expressing lentiviruses……92

Figure 3.5: Sorting of lentivirus-transduced SupT1 cell lines……93

Figure 3.6: Processing of HDF-targeting shRNAs in SupT1 cells……93
Figure 3.7: SupT1 cell line proliferation rates do not significantly differ……94
Figure 3.8: HTATSF1-targeting shRNA inhibits p81A-4 replication……95
Figure 3.9: HTATSF1 mRNA levels increase in shH1-expressing SupT1 cells over time……96
Figure 3.10: HTATSF1 increases in shH1-expressing SupT1 cell lines over time……97
Figure 3.11: The proportion of GFP+ SupT1 cells in the shH1-expressing cell line decreases over time……97
Figure 3.12: Suppression of HTATSF1 does not alter *htatsf1* transcription but shH1 expression decreases over time……98
Figure 3.13: Schematic of HIV-1 transcript classes and primers used to quantify them……99
Figure 3.14: All classes of p81A-4 transcripts are increased in shH1-expressing SupT1 cells……99
Figure 3.15: HTATSF1 suppression does not alter the ratio of p81A-4 transcript classes……100
Figure 4.1: Schematic of the distribution of various core promoter elements……108
Figure 4.2: Procedure for construction of siRNAs by an *in vitro* transcription method……111
Figure 4.3: Rational design of promoter-targeting siRNAs……117
Figure 4.4: Promoter-targeting siRNAs can suppress *htatsf1* mRNA levels……118
Figure 4.5: Promoter-targeting siRNA-mediated silencing is sensitive to TSA in HEK293T cells……119
Figure 4.6: Transcription is detectable across the TSR but not upstream of the FANTOM TSS……120
Figure 4.7: Sense and antisense transcription is detectable within the TSR……121
Figure 4.8: SMARTpools effectively suppress Ago expression……123
Figure 4.9: Ago suppression has no significant effect on siRNA-mediated silencing……123
Figure 4.10: Promoter-targeting siRNAs may alter *htatsf1* transcription rates……125
Figure 4.11: Promoter-targeting siRNAs can decrease p81A-4 infectious particle production……125
Figure 5.1: Extreme contrasts in promoter architecture……131
Figure 5.2: The transcriptionally-silenced group promoter features compared with those of 2000 random TSRs……143
Figure 5.3: The transcriptionally-silenced group promoter features compared with those of 2000 random TSRs and corrected for multiple testing……144
Figure 5.4: The transcriptionally-silenced group promoter features compared with all other TSRs……146
Figure 5.5: The transcriptionally-silenced group promoter features compared with TSRs assigned to genes expressed in THP-1 cells……148
Figure 5.6: The transcriptionally-silenced group promoter features compared with TSRs assigned to housekeeping genes……150
Figure A2.1: Promoter-targeting with expressed asRNAs……213
Figure A2.2: asRNAs do not suppress expression of target HDFs……214
Figure A2.3: *psip1* promoter-targeted siRNAs can suppress target expression……215
Figure A2.4: Transcription is detectable at siRNA target regions……216
Figure A2.5: Ago2 suppression alleviates siP10-mediated *psip1* silencing……217
Figure A2.6: siRNA-mediated silencing of *psip1* is not accompanied by epigenetic changes at the target region……218
Figure A2.7: *psip1* transcription rates are unaltered by siRNA-mediated silencing……219

Figure A2.8: siRNA target sites do not align to the *psip1* promoter but the 3'UTR……219
Figure A2.9: siRNAs targeting the CSH TRED *ddx3x* promoter do not induce TGS and align to the 5'UTR….220

LIST OF TABLES

Table 1.1: Overview of experimental details of four genome-wide RNAi screens for the identification of HDFs……18
Table 2.1: Regions within host factor transcripts targeted by shRNAs……43
Table 2.2: Oligonucleotide primers used to generate shRNAs targeting cellular factors……45
Table 2.3: Oligonucleotides used to generate shRNA target sequences in the reporter constructs……47
Table 2.4: Primers used in qPCR to determine cellular factor cDNA levels……50
Table 2.5: Sequences of DNA oligonucleotide probes used in Northern blots……52
Table 3.1: Primers used in qPCR to determine levels of HIV-1 transcript classes……88
Table 4.1: Core promoter elements……108
Table 4.2: Oligonucleotide templates used to generate *htatsf1* and *ddx3x* promoter-targeting siRNAs…….110
Table 4.3: Primers used to determine levels of *htatsf1* promoter-associated transcripts……114
Table 5.1: Histone modifications quantified in a 2000 bp window centred on the representative TSS……135
Table 5.2: Transcriptionally-silenced genes included in the study and the assigned TSRs……141
Table 5.3: Fraction of ncRNA overlap with TSRs in the various groups……152
Table 5.4: The Euclidean distance between *htatsf1* and *ddx3x* promoter features and various TSR groups152
Table A2.1: Oligonucleotide primers used to generate asRNAs targeting HDF promoters……207
Table A2.2: Oligonucleotide templates used to generate HDF promoter-targeting siRNAs……208
Table A2.3: Primers used to determine levels of promoter-associated transcripts……209
Table A2.4: siRNAs used to examine the role of Argonaute proteins in siP10-mediated *psip1* silencing……210

LIST OF ABBREVIATIONS

A	adenine
AAV	adeno-associated virus
ac	acetylation
Ago1	Argonaute 1
Ago2	Argonaute 2
asRNA	antisense RNA
ATP	adenine triphosphate
BAF1	barrier to auto-integration factor 1
BCA	bicinchoninic acid
BLAST	basic local alignment search tool
BLT	bone marrow/liver/thymus
bp	base-pair
BRE	TFIIB recognition element
C	cytosine
CA	capsid
CAGE	cap analysis of gene expression
CCD	catalytic core domain
CCR5	chemokine (C-C motif) receptor type 5
CD	cluster designation
CDK9	cyclin-dependent kinase 9
cDNA	complementary DNA
CES	combined epigenetic score
ChIP	chromatin immunoprecipitation
CIBEX	Center for Information Biology Gene Expression Database
CMV	cytomegalovirus
cPPT	central polypurine tract
CRM1	chromosome maintenance region 1
CSDC2	cold-shock domain-containing protein 2
CSH	Cold Spring Harbor
CTD	C-terminal domain
CTL	cytotoxic T lymphocyte
CXCR4	chemokine (C-X-C motif) receptor type 4
DBTSS	database of transcription start sites
DC	dendritic cell
DCE	downstream core element
DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked
DEAE-D	diethylaminoethyl dextran
DEPC	diethylpyrocarbonate
DGCR8	DiGeorge syndrome critical region 8
DIG	digoxigenin
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
ds	double-stranded
DTT	dithiothreitol
EC50	median effective concentration
EDTA	ethylenediaminetetraacetic acid
EF-1α	elongation factor 1 alpha
eGFP	enhanced GFP
EGTA	ethyleneglycoltetraacetic acid
ELISA	enzyme-linked immunosorbent assay
endo-siRNA	endogenous small interfering RNA
Env	envelope
EPD	eukaryotic promoter database

FANTOM	functional annotation of the mammalian genome
FCS	fetal calf serum
fdr	false discovery rate
G	guanine
Gag	group-specific antigen
gDNA	genomic DNA
GFP	green fluorescent protein
gp	glycoprotein
GRC	Genome Reference Consortium
H	histone
HAART	highly active antiretroviral therapy
HBV	hepatitis B virus
HCV	hepatitis C virus
HDF	HIV-dependency factor
HEK	human embryonic kidney
HIV	human immunodeficiency virus
HK	housekeeping
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
HTATSF1	HIV-1 Tat specific factor 1
IBD	IN binding domain
IN	integrase
IL	interleukin
Inr	initiator
IPTG	isopropyl-β-D-1-thiogalactopyranoside
K	lysine
LB	Luria Bertani
LEDGF	lens epithelium-derived growth factor
lhRNA	long hairpin RNA
LNA	locked nucleic acid
LTR	long terminal repeat
MA	matrix
me(1)	methylation
me2	dimethylation
me3	trimethylation
min	minute(s)
mRNA	messenger RNA
miRNA	microRNA
MoMLV	Moloney murine leukaemia virus
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NC	nucleocapsid
ncRNA	non-coding RNA
NCBI	National Center for Biotechnology Information
NDR	nucleosome-depleted region
Nef	negative effector
NF-κB	nuclear factor kappa B
nt	nucleotide
NTP	nucleoside triphosphate
PAK1	p21-activated kinase
PBMCs	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PGR	progesterone receptor
PIC	pre-integration complex
piRNA	Piwi-interacting RNA
PKR	dsRNA-dependent protein kinase
PMA	phorbol 12-myristate 13-acetate
PMSF	phenylmethane-sulfonylfluoride
Pol	polymerase
PR	protease
PSIP1	PC4 and SFRS1 interacting protein 1

pTEF-b	positive transcription elongation factor b
PTGS	post-transcriptional gene silencing
PVDF	polyvinylidene fluoride
qRT-PCR	quantitative reverse transcription PCR
R	arginine
RACE	rapid amplification of cDNA ends
Rev	regulator of viral gene expression
RFV	representative feature vector
RIG-1	retinoic acid-inducible gene 1
RIPA	Radio-Immunoprecipitation Assay
RISC	RNA-induced silencing complex
RITS	RNA-induced transcriptional silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
RNP	ribonucleoprotein
RRM	RNA recognition motif
rRNA	ribosomal RNA
RT	reverse transcriptase
SCID	severe combined immunodeficiency
SDS	sodium dodecylsulphate
sec	second(s)
SEM	standard error of the mean
shRNA	short hairpin RNA
siRNA	small interfering RNA
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
ss	single-stranded
SV40	Simian vacuolating virus 40
T	thymine
TAR	transactivation response region
Tat	transcriptional transactivator
TBP	TATA-binding protein
TBS	Tris buffered saline
$TCID_{50}$	median tissue culture infectious dose
TFBS	transcription factor binding site
TGA	transcriptional gene activation
TGS	transcriptional gene silencing
TI	transcriptional interference
TLR	Toll-like receptors
TRED	transcriptional regulatory element database
tRNA	transfer RNA
TSA	trichostatin-A
TSC	transcription start cluster
TSR	transcription start region
TSS	transcription start site
U	uracil
UCSC	University of California, Santa Cruz
UTR	untranslated region
UV	ultra-violet
VEGF-A	vascular endothelial growth factor A
Vif	viral infectivity factor
Vpr	viral protein R
Vpu	viral protein U
VSV-G	Vesicular stomatitis virus glycoprotein
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside
XCPE1	X core promoter element 1
XCPE2	X core promoter element 2

CHAPTER 1

General introduction

1.1 Introduction

Human immunodeficiency virus (HIV) is the aetiological agent of acquired immune deficiency syndrome (AIDS), which is characterised by immune dysfunction and death approximately 11 years post-infection in the absence of treatment (UNAIDS/WHO 2009). It is currently estimated that 33.3 million people are infected worldwide, with 2.6 million new infections and 1.8 million AIDS-related deaths in 2009 (www.unaids.org/globalreport/). Sub-Saharan Africa bears an inordinate share of the global HIV burden with 68% of infections, with South Africa housing the largest epidemic of any country – 5.6 million in 2009. Although HIV has been the most intensely studied infectious agent since its identification in 1983 (Gallo *et al.* 1983), a preventative vaccine or post-infection cure remains elusive. This may be largely attributed to the high variability of HIV – not only does this facilitate the emergence of drug resistance and escape from the immune system, but it also hampers the design of treatments/vaccines that are effective against all circulating viruses.

Although currently available drugs to treat the infection are highly effective at reducing HIV-related deaths, they require chronic, daily administration to be effective. Treatment success is tightly correlated with patient compliance, which can be poor as a result of toxic side effects of the drugs and accessibility of chronic treatment. This could be dramatically improved through the use of a single intervention therapy that would aid the patient for life. Gene therapy has the potential to provide a solution, although existing single-intervention anti-HIV gene therapy strategies require complex and expensive engraftments. Those not delivered by transplantation are subject to concerns with respect to chronic administration and toxicity. However, a mammalian pathway, small RNA-induced transcriptional gene silencing (TGS), has the potential to mediate durable target silencing. Gene therapies based on TGS may, therefore, reduce administration frequency and toxicity without the need for expensive procedures, which would be of particular benefit to resource-poor settings.

Reduction in the emergence of viral resistance would also improve current treatment regimens. Inhibitors of cellular factors that HIV requires for replication will increase the genetic barrier to viral escape, and so may provide a powerful addition to existing antiviral therapeutics. To this end, there are substantial efforts being employed to identify host factors involved in HIV-1

replication and validate them as therapeutic targets, by showing that their suppression inhibits viral replication without toxic side effects to the host.

The ultimate goal of the work described in this thesis was to contribute to the development of improved HIV therapeutics with reduced toxicity, administration frequency and viral resistance. This was accomplished through validating the therapeutic potential of targeting host factors, by assessing how their suppression affects viral replication and cytotoxicity, and examination of their susceptibility to suppression by small RNA-induced TGS. Thus, this general introduction will examine how the variability of HIV impedes clearance of the virus, emphasising the rationale behind therapies designed to target cellular factors involved in viral replication. The RNA interference (RNAi) pathway will be described and how exploitation of this powerful gene silencing mechanism can facilitate identification of host proteins required for completion of the HIV life cycle and interrogate their therapeutic potential. The capacity of TGS to provide a durable gene therapy will be discussed and, finally, the putative roles in HIV-1 replication of various cellular factors and their potential as anti-HIV therapeutic targets.

1.2 Human immunodeficiency virus

1.2.1 HIV origins

HIV is a member of the *Lentivirus* genus and the *Retroviridae* family and isolates are currently grouped into two types, HIV-type 1 (HIV-1) and -type 2 (HIV-2). HIV-2 is largely restricted to Western and Central Africa, reflecting the geographical distribution of the primate reservoir, the sooty mangabey (Van Heuverswyn and Peeters 2007; Wain *et al.* 2007). In contrast, the HIV-1 primate reservoir was a great ape, the chimpanzee, not a monkey (Wain *et al.* 2007). HIV-1 dominates the global epidemic, which may relate to the apparent decreased virulence of HIV-2 (Whittle *et al.* 1994).

HIV-1 is further subdivided into three major groups, which arose from three independent ape-to-human transmissions: major (M), outlier (O) and non-M/non-O (N). These three groups are as distinct from each other as they are from the lentivirus causing immunodeficiency in chimpanzees, termed simian immunodeficiency virus of chimpanzees (SIVcpz) (Gao *et al.* 1999). Group M includes nine subtypes, or clades, designated A to K based on genomic sequence homology. In Southern and East Africa, HIV-1 subtype C is the most prevalent virus, accounting for 48% of global infections, whilst subtype B is the most represented virus in North and South America, Central and Western Europe and Australia, accounting for 15% of global infections (UNAIDS 2008). Viruses of different subtypes may infect the same cell and share genetic material

resulting in recombinant forms. Those mosaic viruses able to replicate and be transmitted are termed circulating recombinant forms (CRFs), and those that do not spread from their original location, unique recombinant forms (URFs) (Carr *et al.* 2001).

1.2.2 HIV genome organisation

The HIV genome encodes structural proteins and enzymes common to retroviruses. In addition, lentiviruses, such as HIV, encode a number of regulatory and accessory proteins. The HIV virion contains two identical copies of positive single-stranded (ss) RNA molecules containing the genes *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) (Figure 1.1A). These encode structural proteins of the core and matrix, enzymes critical for replication and viral envelope glycoproteins, respectively. Along with host-derived lipoprotein-rich membrane, these components form enveloped viral particles with a diameter of ~120 nm (Figure 1.1B).

The *gag* gene encodes $Pr55^{Gag}$, a polyprotein precursor that is cleaved by the intrinsic viral protease (PR) to produce proteins matrix (MA/p17), capsid (CA/p24) and nucleocapsid (NC/p6 & p7) (Ganser-Pornillos *et al.* 2008). The *pol* gene products are also initially synthesised as a polyprotein precursor, $Pr160^{GagPol}$ that, as the name implies, includes both translated *gag* and *pol* regions following a rare frameshift event during $Pr55^{Gag}$ translation. PR again releases the mature proteins, resulting in the enzymes PR, reverse transcriptase (RT; p66 and p51 subdomains) and integrase (IN) (Moulard and Decroly 2000; Hill *et al.* 2005). Unlike the Gag and Pol precursors, the polyprotein encoded by the *env* gene, gp160, is processed by cellular proteases (Moulard and Decroly 2000; Hill *et al.* 2005). This occurs during trafficking of gp160 to the cell surface, forming gp120, which is exposed on the virion surface and interacts with receptors, and gp41, which constitutes the transmembrane domain and contains elements that are important for membrane fusion. gp120 and gp41 both form trimers that associate non-covalently to form heterodimers.

Reverse transcription of the RNA genome generates proviral DNA with long terminal repeats (LTRs) flanking the coding region. These LTRs are divided into 3 regions: unique, 3' end (U3), repeated (R) and unique, 5' end (U5). Incomplete LTRs, composed of just R/U5 and U3/R are present at the 5' and 3' ends of the RNA genome, respectively (Figure 1.1A). Transcription predominantly initiates from the provirus 5' U3/R junction and, consequently, the U3 region contains a number of enhancer elements, including transcription factor binding sites and a TATA box (Patarca *et al.* 1987; Roebuck and Saifuddin 1999). Basal transcriptional activity is very low, however, but is increased by more than two logs in the presence of the regulatory protein Tat (transcriptional transactivator) (Dayton *et al.* 1986; Fisher *et al.* 1986), which is encoded by a two-exon RNA (Figure 1.1A). Tat acts upon the transactivation response region (TAR) (Berkhout *et al.*

1989), a stem-loop structure encoded in the R region and present at the 5' end of all viral RNAs. Another regulatory gene, *rev* (regulator of viral gene expression), encodes a protein involved in the nuclear export of unspliced and partially spliced viral transcripts that characterise the later stages of viral gene expression (Jeang *et al.* 1991).

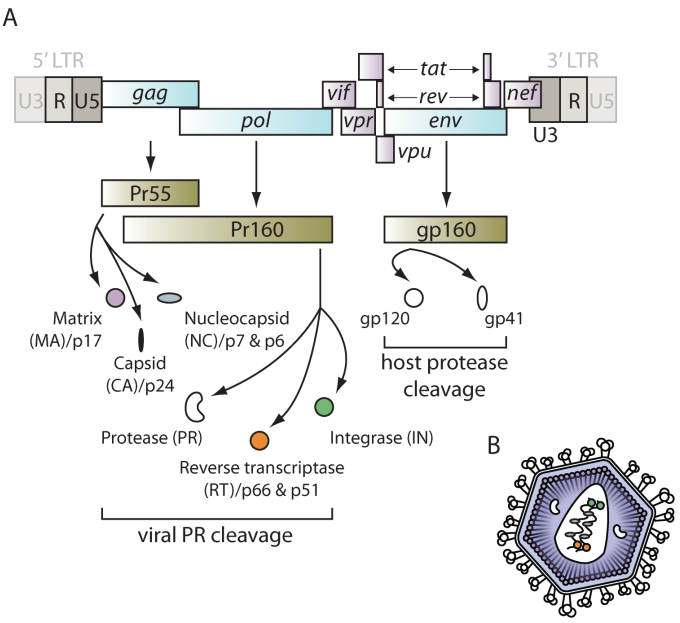


Figure 1.1: HIV-1 genome organisation. A. The HIV-1 genome is composed of nine genes, three of which encode polyprotein precursors. $Pr55^{Gag}$ and $Pr160^{GagPol}$ are processed by the viral protease (PR) to produce structural proteins and enzymes, respectively. The polyprotein product of the *env* gene, gp160, is cleaved by host cellular proteases to produce envelope glycoproteins gp120 and gp41. The genome includes genes encoding two regulatory proteins *tat* and *rev*, and four accessory proteins *vif*, *vpr*, *vpu* and *nef*. LTRs are not present in the RNA genome and only in the proviral DNA following duplication of U3 and U5 sequences during the reverse transcription process. **B.** Schematic of a HIV virion, an enveloped lentivirus of ~120 nm in diameter.

Four accessory proteins are encoded in the HIV-1 genome, including Nef (negative effector/necessary factor), Vif (viral infectivity factor), Vpr and Vpu (viral proteins R and U) (Figure 1.1A). These proteins carry out various functions that enhance infectivity, in particular Nef [reviewed in (Emerman and Malim 1998; Arhel and Kirchhoff 2009; Nomaguchi and Adachi 2009; Dube *et al.* 2010a; Jere *et al.* 2010)]. Of note, Vpr is involved in cell cycle arrest and enables the reverse transcribed viral DNA to access the nucleus, resulting in productive infection, in non-dividing cells such as macrophages (Andersen *et al.* 2008). Neither mature regulatory or accessory proteins require proteolytic cleavage for their production.

1.2.3 HIV life cycle

HIV infects and replicates in human CD4+ T cells and macrophages (Figure 1.2). Viral particles bind to surface CD4, a 58 kDa type I integral cell surface glycoprotein, triggering a

conformational change that exposes a binding surface on gp120 for chemokine receptors (Liu *et al.* 2008). At least 17 chemokine receptors have been identified as functioning as HIV coreceptors (Alkhatib and Berger 2007), although the α-chemokine receptor, chemokine (C-X-C motif) receptor type 4 (CXCR4), and the β-chemokine receptor, chemokine (C-C motif) receptor type 5 (CCR5), appear to predominate *in vivo*. CCR5 is expressed on macrophages, dendritic cells (DCs), brain cells and activated lymphocytes, while CXCR4 is expressed in resting T cells and monocytes (Lusso 2000). This differential distribution of coreceptors is important in determining tropism of viruses: those preferentially using CCR5 are termed M (for macrophage)-tropic, or R5 viruses, and those preferentially using CXCR4, T (for T lymphocyte)-tropic, or X4 viruses (Agrawal *et al.* 2009). Viruses able to use both CCR5 and CXCR4 for entry are termed dual tropic, or R5X4 viruses.

gp120 interaction with a chemokine receptor causes conformational changes allowing gp41 to insert itself into the host cell membrane, effecting membrane fusion (Campbell and Hope 2008). About four to six CCR5 coreceptors are required for viral fusion to occur (Kuhmann *et al.* 2000). The capsid is then injected into the cytoplasm and shuttles toward the nuclear membrane, during which time reverse transcription takes place (Arhel *et al.* 2007). Reverse transcription is a complex process catalysed by the viral RT, a heterodimer of two subunits, p51 and p66. p51 is encoded by the same sequence as p66 but lacks the C-terminal RNaseH domain (through PR activity), yet adopts a very different structure to p66 (Seckler *et al.* 2009). Reverse transcription is initiated by a transfer RNA (tRNA) primer and proceeds in a series of steps including DNA synthesis and RNA degradation that utilise a number of *cis*-elements in the viral genome. A ~100 nt DNA 'flap' within the plus-strand DNA is generated during the final stage of reverse transcription which is critical for uncoating and nuclear import of the viral pre-integration complex (PIC) (Arhel *et al.* 2007).

The mechanism by which HIV overcomes the nuclear pore diffusion size limit and the DNA concentration gradient to import the PIC into the nucleus remains unclear. Nevertheless, the central polypurine tract (cPPT), IN, Vpr and a nuclear localisation signal in MA, have all been shown to play roles in import. Importin 7 has been shown to mediate PIC import (Fassati *et al.* 2003) in a HIV-specific manner that is dependent on IN binding (Zaitseva *et al.* 2009). In addition, tRNAs were shown to promote nuclear import (Zaitseva *et al.* 2006), although the host factors involved in this process have yet to be identified.

Once the PIC reaches the nucleus, with the aid of the viral enzyme IN, the DNA molecule is incorporated into human DNA (Figure 1.3). The main DNA cleaving and joining steps are the initial removal of a terminal dinucleotide from each end of the reverse transcribed cDNA (3' processing) and strand transfer. The latter is a single-step transesterification reaction in which

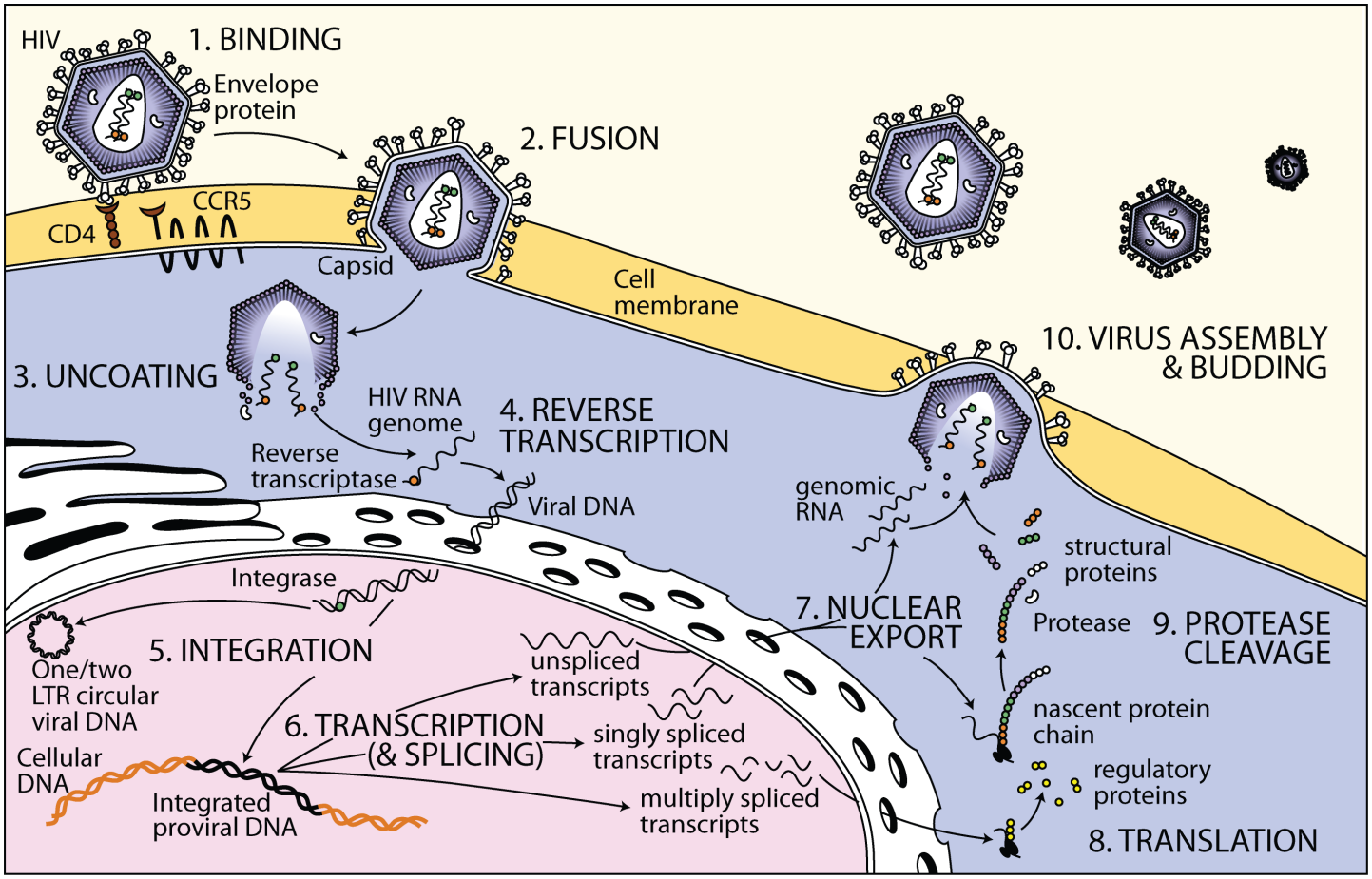


Figure 1.2: The basic steps of the HIV life cycle. Following binding (1) of the virus to the cell surface via its receptor CD4, association with a coreceptor, typically CCR5 or CXCR4, leads to virus and host cell membrane fusion (2). Uncoating (3) and reverse transcription (4) of the RNA genome precedes nuclear import of a dsDNA molecule that is integrated (5) into the host genomic DNA. The integrated proviral DNA then hijacks the host cellular transcription machinery for expression of different classes of RNA transcripts, determined by the degree of splicing (6). Nuclear export (7) of these transcripts allow their translation (8) and, where required, protease cleavage (9). Unspliced transcripts are combined with viral proteins in virus assembly at the host cell membrane with subsequent budding (10) of new virions from the cell surface.

staggered cleavage of opposite target DNA strands provides the energy to join the newly recessed viral 3' ends to them (Brown *et al.* 1987). In the third main step of integration, gap repair, extra nucleotides are trimmed from the 5' ends of the viral cDNA and these are joined to host DNA 3' ends. This step involves host cell DNA repair enzymes (Skalka and Katz 2005). In addition to the integration reaction, non-homologous end joining and homologous recombination cellular pathways are involved in the formation of 1- and 2-LTR circles that are detected in the nuclei of infected cells (Li *et al.* 2001).

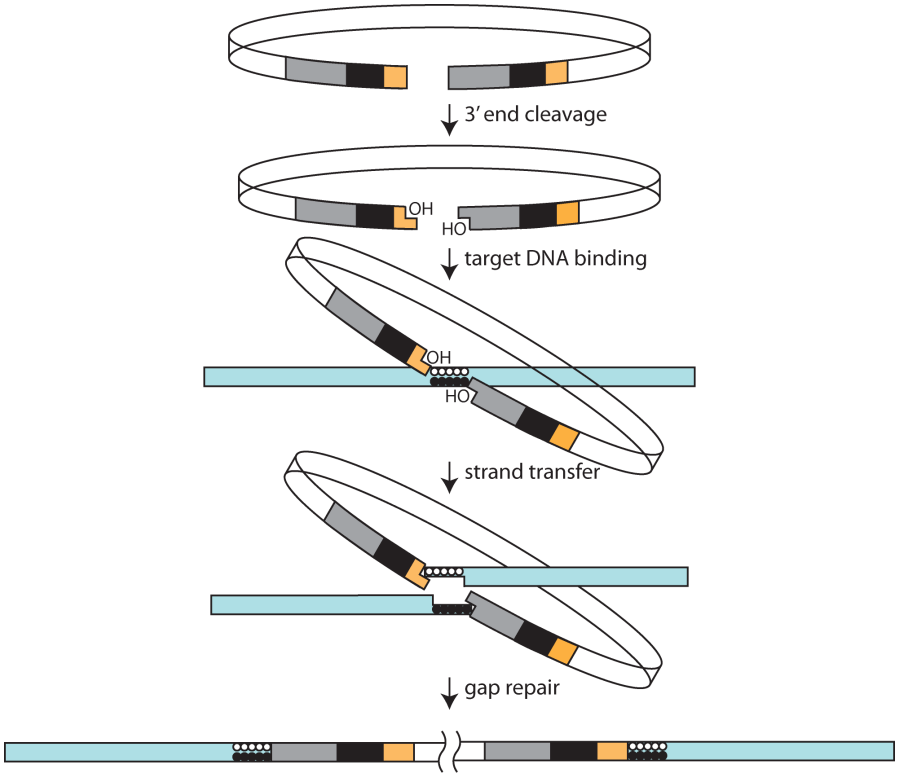


Figure 1.3: The HIV integrase reaction. A blunt-ended, linear viral genome cDNA is the precursor to integration. 3' processing generates a recessed CA at the 3' end. Binding to host DNA is followed by a concerted transesterifcation reaction. The process is completed by the third main step, gap repair, yielding a duplication of host DNA flanking the provirus (ie five bases, shown here by circles). Host cell DNA repair enzymes are involved in this step.

Proviral DNA can remain quiescent for extended periods or become transcriptionally active, particularly where there is inflammation (Staal *et al.* 1990). As previously mentioned, Tat regulates proviral transcription. In the absence of Tat, RNA polymerase (Pol) II is recruited to the LTR through TATA-binding protein (TBP) but productive elongation is largely absent because of a lack of RNA Pol II C-terminal domain (CTD) phosphorylation. The few protein-coding transcripts that are produced during this Tat-independent phase result in sufficient Tat to activate transcription. Tat binds to the TAR loop and recruits positive transcription elongation factor b (pTEF-b) through an association with the cyclin T1 subunit (Richter *et al.* 2002). The cyclin-dependent kinase 9 (CDK9) catalytic domain of pTEF-b is then brought into proximity with RNA Pol

II and can phosphorylate serines 2 and 5 of the RNA Pol II CTD (Parada and Roeder 1996), alleviating the elongation stall. HIV-1 transcription is further regulated through the interaction of transcription factors, including nuclear factor (NF)-κB, with binding sites in the LTR (Williams *et al.* 2007). Binding of such host factors promotes transcription through induction of nucleosome remodelling (He *et al.* 2002).

There are three classes of protein-coding HIV-1 transcripts: (1) ~9 kb unspliced, which function as $Pr55^{Gag}$ and $Pr160^{GagPol}$ precursors and virion genomes; (2) ~4 kb partially spliced, termed singly spliced, which encode Env, Vif, Vpr and Vpu; and, (3) ~2 kb abundantly spliced, often termed fully spliced, which encode Tat, Rev and Nef. The prevailing view is that Rev binding to the Rev responsive element (RRE) shepherds intron-containing unspliced and singly spliced HIV transcripts from the nucleus to the cytoplasm, which would normally be degraded. In doing so, Rev interacts with chromosome maintenance region 1 (CRM1) (Bogerd *et al.* 1998), utilising a shuttling pathway normally employed for export of proteins, small nuclear RNAs (snRNAs) and ribosomal RNAs (rRNAs) (Fornerod *et al.* 1997). The Rev-CRM1 pathway is distinct from that used to export fully spliced HIV-1 messenger RNAs (mRNAs) and cellular mRNAs (Clouse *et al.* 2001). In this way, Rev regulates a key switch between the early pattern of HIV gene expression, in which short, multiply-spliced mRNAs encoding the viral regulatory genes Tat, Rev and Nef are exported from the nucleus, and the late pattern, in which larger, singly spliced and unspliced RNAs encoding the viral structural proteins and the RNA genomes of progeny virions are exported.

Following translation of HIV transcripts in the cytoplasm, viral proteins accumulate at the cell surface. Genomic RNA encapsidation is directed by a *cis*-acting sequence known as the packaging signal, or Ψ-site (Berkowitz *et al.* 1996), which interacts with NC. Initially, immature virions are formed, characterised by incomplete spherical capsids composed of the polyprotein Gag in a hexagonal lattice (Briggs *et al.* 2009). Virion maturation results from sequential cleavage of Gag by PR, which causes a rearrangement of the capsid to a conical structure, composed solely of CA (Wiegers *et al.* 1998; Briggs *et al.* 2003). Prior to formation of complete Gag spheres, cellular factors, including the endosomal sorting complex required for transport (ESCRT) machinery, promote HIV-1 virion budding (Carlson *et al.* 2008) whilst restriction of budding by CD137/tetherin is antagonised by Vpu (Neil *et al.* 2008; Dube *et al.* 2010b). Virion maturation is thought to occur during or shortly after particle release from the plasma membrane.

1.2.4 HIV pathogenesis

Within a few weeks of infection, there is a high level of HIV replication in the blood that can exceed ten million viral particles per microlitre of plasma (Simon and Ho 2003). Concomitant

with this is a decline in the numbers of CD4+ T lymphocytes. However, within a few weeks an immune response to HIV develops that curtails viral replication, resulting in a decline in the viral load and a return of CD4+ T cell numbers to near normal levels [reviewed in (Fanales-Belasio *et al.* 2010)]. The control of viraemia has largely been attributed to the emergence of cytotoxic T lymphocytes (CTL) and, to a lesser extent, the development of a neutralising antibody response (Streeck and Nixon 2010). As a result of these immune responses, individuals can remain asymptomatic for many years. Kinetic studies have shown that, during the asymptomatic period, up to a billion HIV particles and two billion CD4+ T cells are destroyed and produced each day (Simon and Ho 2003). Thus, the virus continues to replicate, causing a gradual drop in CD4+ T cell numbers (Mellors *et al.* 1996) resulting in individuals becoming susceptible to various opportunistic infections (Brooks *et al.* 2009) and aggressive neoplasms (Broder 2010), signalling the onset of AIDS.

There are several reasons why antiviral immunity is unable to eradicate the HIV infection. A functional impairment of immune responses can be detected, indicating that the virus induces a dysfunction of immune cells very early after infection (Rosenberg *et al.* 1998; Douek 2003; Lange *et al.* 2003; Lichterfeld *et al.* 2004). Disease progression is accompanied by destruction of the lymphoid tissue architecture, which further hampers immune responses. Circulating viruses display a high degree of variability, which enables escape from immune recognition. This is facilitated by the persistence of integrated virus in, amongst others, macrophage and resting CD4+ T cell reservoirs. Such latent infections are characterised by low expression of viral antigens preventing eradication even in the presence of effective antiviral treatment.

Some individuals display efficient control of HIV infection – so-called 'elite controllers' or 'long-term non-progressors'. Despite extensive study of these cohorts, no correlates or mechanism of protection against HIV infection have been definitively identified. Vaccine development has been further hampered by the extreme variability of HIV, both within the host and between clades, groups and types.

1.2.5 HIV variability and treatment

Variability underscores the ability of HIV to escape host immunity and restricts development of a vaccine that can elicit cross-clade protection. Variability within the host is a product of at least three features: (1) a highly erroneous reverse transcription process that introduces, on average, one substitution per genome per replication (Sarafianos *et al.* 2009); (2) rapid viral replication that generates around 10^{10} virions per day in an infected individual (Ho 1997); and, (3) recombination of HIV strains in an individual. Consequently, envelope proteins, the

major targets of vaccine-induced antibodies, can differ in sequence by up to 38% within the same HIV-1 subtype (Korber and Gnanakaran 2009). Vaccines designed to elicit CTL responses encounter the same difficulties, as viruses within the same HIV-1 subtype differ by up to 20% (McMichael *et al.* 2002).

Although an effective vaccine has remained elusive, a range of antiretroviral small molecule drugs has brought about a substantial decrease in HIV-associated mortality. Most of these target the HIV enzymes RT, PR and IN, and in combination form highly active antiretroviral therapy (HAART) that has changed HIV infection from a rapidly lethal disease to a chronic, manageable condition. More specifically, antiretroviral intervention extends the life span of a typical HIV-1-infected patient in the USA by about 14 years (Walensky *et al.* 2006).

The HIV/AIDS pandemic is still far from stabilised, however. HAART must be undertaken for a lifetime as its discontinuance is accompanied by viral rebound (Mata *et al.* 2005; Chun *et al.* 2010), yet chronic antiretroviral therapy can bring about cardiac and metabolic side-effects, which increase the risk of developing heart disease and type 2 diabetes (Filardi *et al.* 2008; Sabin *et al.* 2008; Silverberg *et al.* 2009; Williams *et al.* 2009). In addition, Sub-Saharan Africa is the epicentre of the pandemic but antiretroviral therapy is available for too few, or reaches those who need it too late, in this resource-poor region. Other limitations of HAART include an inability to eradicate the virus because of the inducible retroviral reservoir present in memory T cells that can persist for long periods (Marcello 2006). HAART ultimately drives the emergence of drug-resistance and regimen selection is complicated by the ability of new HIV-1 strains to emerge and migrate quickly (Taylor *et al.* 2008). This is of particular relevance to the use of drug-resistance algorithms, which employ subtype B as the consensus sequence (Hirsch *et al.* 2008). Of concern is that there is some indication that the prevalence of CRFs and URFs is growing rapidly (Taylor *et al.* 2008), although the implications of this trend are unknown. New cross-species transmissions from established retroviral reservoirs in apes and Old World monkeys are also possible, particularly as logging in previously inaccessible forests and the demand for bushmeat continues [reviewed in (Broder 2010)]. Indeed, a Cameroonian living in Paris was recently documented as being infected with a HIV strain originating in gorillas (SIV_{gor}), designated group P (Plantier *et al.* 2009). Whether new viruses that arise from recombination or transpeciation events will be eminently susceptible to current antiretroviral therapies is unknown. Therefore, the search for safer and more effective therapies remains of critical importance.

Ideally, novel HIV therapies will present a high genetic barrier to the development of drug resistance and fit with the realities of resource-poor nations. Targeting host proteins required for replication, so-called HIV-dependency factors (HDFs), would increase the genetic barrier to viral escape. Cellular targets are virtually constant relative to the variability of viral targets within a

patient and across the pandemic. Moreover, viral escape would require adaptation to the use of a different cofactor (Zhou *et al.* 2004b; Nair *et al.* 2005). Although this is possible in principle, particularly given the redundancy typical of cellular pathways, it has not been observed experimentally. Thus, a low-intervention gene therapy targeting a HDF presents an attractive anti-HIV therapeutic strategy.

The major concern with HDF-targeting is the potential for adverse effects on the host. Nonetheless, a naturally occurring 32-base pair (bp) deletion (Δ32) in the CCR5 gene was discovered that renders homozygous individuals resistant to R5 virus infection (Huang *et al.* 1996; Liu *et al.* 1996). The deletion results in premature truncation and a lack of CCR5 expression on the cell surface (Samson *et al.* 1996), yet no associated physiological problems have been described. Indeed, some of the most promising anti-HIV treatments already approved or in clinical trials target, not a viral protein, but CCR5. Individuals heterozygous for the Δ32 allele progress to disease more slowly and have lower viral loads. Thus, complete abolition of HDF gene expression may not be necessary for inhibition of viral replication. Additional HDFs may exist, therefore, whose depletion would be detrimental to the virus without compromising host fitness. Consequently, the identification of such HDFs has become a focus in recent years. A number of groups has performed genome-wide screens to this end [reviewed in (An and Winkler 2010)], although most of the novel HDFs identified have yet to be confirmed. The validation of HDFs as potential therapeutic targets is, therefore, a major objective of this thesis. Genome-wide identification of HDFs and their subsequent validation often exploits the cellular pathway RNA interference (RNAi) for specific inhibition of host gene expression.

1.3 RNA interference

RNAi is a process conserved throughout eukaryotes involving small non-coding RNAs (ncRNAs) that suppress gene expression through an association with mRNAs of complementary sequence. The first small RNA, lin-4, was identified through genetic screens in nematodes (Lee *et al.* 1993; Wightman *et al.* 1993), and it was subsequently shown that double-stranded (ds) RNAs were the precursors to such silencing events (Fire *et al.* 1998). Since these seminal findings, RNAi has proved an indispensible cellular pathway and a powerful experimental technique.

The number of eukaryotic silencing small RNA classes continues to expand with improvements in next generation sequencing technologies [reviewed in (Ghildiyal and Zamore 2009)]. Although it is becoming increasingly apparent that different small RNA-mediated silencing pathways are intertwined [reviewed in (Siomi and Siomi 2009)], two classes fall under the

umbrella of RNAi in humans as both originate from dsRNA precursors that are processed by RNase III enzymes: microRNAs (miRNAs) and endogenous small interfering RNAs (endo-siRNAs). Both are involved in post-transcriptional regulation of RNAs and associate with the Argonaute (Ago) subfamily of the Ago family of proteins. They differ in their biogenesis, however, with miRNAs originating from hairpin precursors and endo-siRNAs from long, dsRNA precursors. miRNA biogenesis is the best understood and exploited RNAi pathway, perhaps reflecting the ubiquitous expression of miRNAs and their importance in regulating specific biological processes in specific tissues at specific times.

1.3.1 miRNA biogenesis in mammals

The transcription of miRNA genes is typically mediated by RNA Pol II (Cai *et al.* 2004; Lee *et al.* 2004), although some associated with Alu repeats may be transcribed by RNA Pol III (Borchert *et al.* 2006). About half of mammalian miRNA loci are clustered and transcribed from a polycistronic unit (Lee *et al.* 2002). Their genomic location can be within both non-coding and coding transcripts. In the latter, the majority are intronic, although miRNAs have been reported in protein-coding exons that are subject to alternative splicing (Kim *et al.* 2009). The primary transcripts (pri-miRNAs) generated contain hairpin motifs composed of an ~33 bp stem region, a flexible terminal loop and flanking single-stranded (ss) sequences. Processing of pri-miRNAs, termed cropping, is carried out by the microprocessor complex in the nucleus cotranscriptionally and pre-splicing (Kim and Kim 2007; Morlando *et al.* 2008; Pawlicki and Steitz 2008). The microprocessor complex contains the RNase III enzyme Drosha and the RNA-binding cofactor DiGeorge syndrome critical region 8 (DGCR8) (Denli *et al.* 2004; Gregory *et al.* 2004; Han *et al.* 2004; Landthaler *et al.* 2004). Recognition of the pri-miRNA leads to its cleavage ~11 bp from the junction between the stem and flanking ss regions (Zeng and Cullen 2005; Han *et al.* 2006), releasing a short hairpin termed the precursor miRNA (pre-miRNA) (Lee *et al.* 2002).

The 60-70 nt pre-miRNA contains an ~16-18 bp stem and an ~2 nt 3' overhang that are recognised by exportin 5 (Zeng and Cullen 2004), through which the pre-miRNA is shuttled to the cytoplasm (Yi *et al.* 2003). Here, the hairpins are processed again by another RNase III enzyme, Dicer, near the terminal loop releasing ~22 nt miRNA duplexes (Hutvagner *et al.* 2001; Lee *et al.* 2002). These duplexes are loaded onto one of four Ago proteins (Ago1-4), generating the RNA-induced silencing complex (RISC) (MacRae *et al.* 2008). Dicer interacts with two proteins, TAR RNA-binding protein (TRBP) and PACT, which may contribute to the formation of RISC (Chendrimada *et al.* 2005; Lee *et al.* 2006).

One strand remains associated with Ago as a mature miRNA, or guide strand, whilst the other, miRNA* or passenger strand, is degraded. Strand selection is determined by relative thermodynamic stabilities at the ends of the miRNA duplexes – the strand with the most unstable 5' end is preferentially retained (Schwarz *et al.* 2003; Ma *et al.* 2005; Parker *et al.* 2005), although some pre-miRNAs may produce miRNAs from both strands at similar frequencies (Wei *et al.* 2009). Ago2 has endonuclease, or slicer, activity and thus can cleave the passenger strand (Matranga *et al.* 2005; Rand *et al.* 2005). However, most miRNA duplexes contain mismatches, which prevent cleavage of the passenger strand. Similar to Ago2, Ago1 also mediates passenger strand cleavage of perfect dsRNA duplexes but only strand dissociation activity of miRNA duplexes (Wang *et al.* 2009a). Ago3 and Ago4 are capable of neither strand-dissociation nor cleavage activities. In the absence of passenger strand cleavage, a RNA helicase activity may facilitate unwinding of the duplex and passenger strand removal (Kim *et al.* 2009).

An Ago-sorting mechanism based on the duplex structure exists in *Caenorhabditis elegans* (Jannot *et al.* 2008) and in *Drosophila melanogaster* (Tomari *et al.* 2007), the latter also dependent on the identity of the guide strand's first nucleotide (Ghildiyal and Zamore 2009). In humans, there is little difference between the identity of miRNAs bound to each Ago protein (Czech and Hannon 2011), although there is some evidence that mismatched duplexes are preferentially loaded into Ago1 and perfect dsRNA duplexes into Ago2 (Gonzalez *et al.* 2008). The mechanism by which a miRNA regulates its target depends on the Ago into which a small RNA is loaded and the degree of complementarity between the guide strand and its target mRNA, usually in the 3' untranslated region (UTR). Those few miRNAs fully complementary to their targets direct Ago2-mediated endonucleolytic cleavage of the mRNA (Figure 1.4A). Unlike Ago2, Ago1 cannot cleave target mRNAs, indicating that the slicer activity of passenger strands and target mRNAs is mechanistically distinct (Wang *et al.* 2009a). The most common mechanism of target mRNA suppression is translational suppression and target degradation in P bodies (Figure 1.4B) (Guo *et al.* 2010). This is directed by miRNAs pairing with their targets only through a region corresponding to nucleotides 2-8 at the 5' end, termed the seed region (Bartel 2009; Guo *et al.* 2010). The small size of the seed means that one miRNA can regulate the expression of several hundred genes simultaneously.

1.3.2 Endo-siRNA and non-canonical miRNA biogenesis

Endo-siRNAs are ~21 nt and, in mammals, have so far been identified in mouse oocytes (Tam *et al.* 2008; Watanabe *et al.* 2008) and adult mouse hippocampus (Smalheiser *et al.* 2011). They originate from transcription of intergenic repetitive elements, pseudogenes and endo-siRNA

clusters. The resulting long, dsRNA precursors are processed in a manner dependent on Dicer, but not Drosha (Tam *et al.* 2008; Watanabe *et al.* 2008). The endo-siRNAs identified in mouse oocytes associated with Ago2, suggesting that they may inhibit gene expression through target cleavage.

Other Drosha-independent mechanisms for the production of Dicer substrates that feed into the standard miRNA pathway have been described, including mirtrons (Berezikov *et al.* 2007; Okamura *et al.* 2007; Ruby *et al.* 2007), small nucleolar RNAs (snoRNAs) (Ender *et al.* 2008) and tRNA fragments (Cole *et al.* 2009; Lee *et al.* 2009; Bogerd *et al.* 2010). In addition, a Dicer-independent mechanism for the production of miRNAs has been identified that relies on the catalytic activity of Ago2 (Cifuentes *et al.* 2010; Yang and Lai 2010; Yang *et al.* 2010).

1.3.3 RNAi mimics

Given that the endogenous miRNA pathway accommodates intermediates from numerous sources, it is unsurprising that the introduction of exogenous RNAi mimics into cells has proved so successful at suppressing gene expression. Small interfering RNAs (siRNAs) were the first RNAi mimics shown to be capable of mediating suppression of target gene expression (Elbashir *et al.* 2001). siRNAs are duplexes, typically 21-23 bp in length with 2 nt 3' overhangs that resemble miRNA duplexes (Figure 1.4C). siRNAs may be readily purchased pre-synthesised or generated by *in vitro* transcription. They can be transfected into cells for direct incorporation into RISC. Whilst highly effective, siRNAs induce silencing only transiently: up to seven days in cultured cells with a rapid turnover rate, which is a result of their dilution or degradation over time (Bartlett and Davis 2006).

The use of expressed RNAi mimics provides an option for more prolonged gene silencing. These may be designed to enter the pathway as Drosha or Dicer substrates. Short hairpin RNAs (shRNAs) with stems ≥20 bp mimic pre-miRNAs (Harborth *et al.* 2003; Ge *et al.* 2010) and so enter the pathway at the level of Dicer processing (Figure 1.4D). These siRNA precursors comprise a sense region that generates a passenger strand, a loop sequence, an antisense region that generates the guide strand and a 2 nt 3' overhang to permit recognition by exportin 5 and Dicer (Zeng *et al.* 2002; Boden *et al.* 2004; Liu *et al.* 2007). shRNAs may be transcribed from DNA plasmids or genome-integrated expression cassettes for more prolonged expression. Although RNA Pol II promoters were investigated (Ren *et al.* 2007; Giering *et al.* 2008), shRNA transcription is typically from RNA Pol III promoters, which generates the defined 5' and 3' ends critical for shRNA efficacy (Paul *et al.* 2002; Tuschl and Borkhardt 2002; Yu *et al.* 2002; Cheng and Chang 2007). RNA Pol III promoters, such as U6 and H1, are constitutively active leading to high levels of

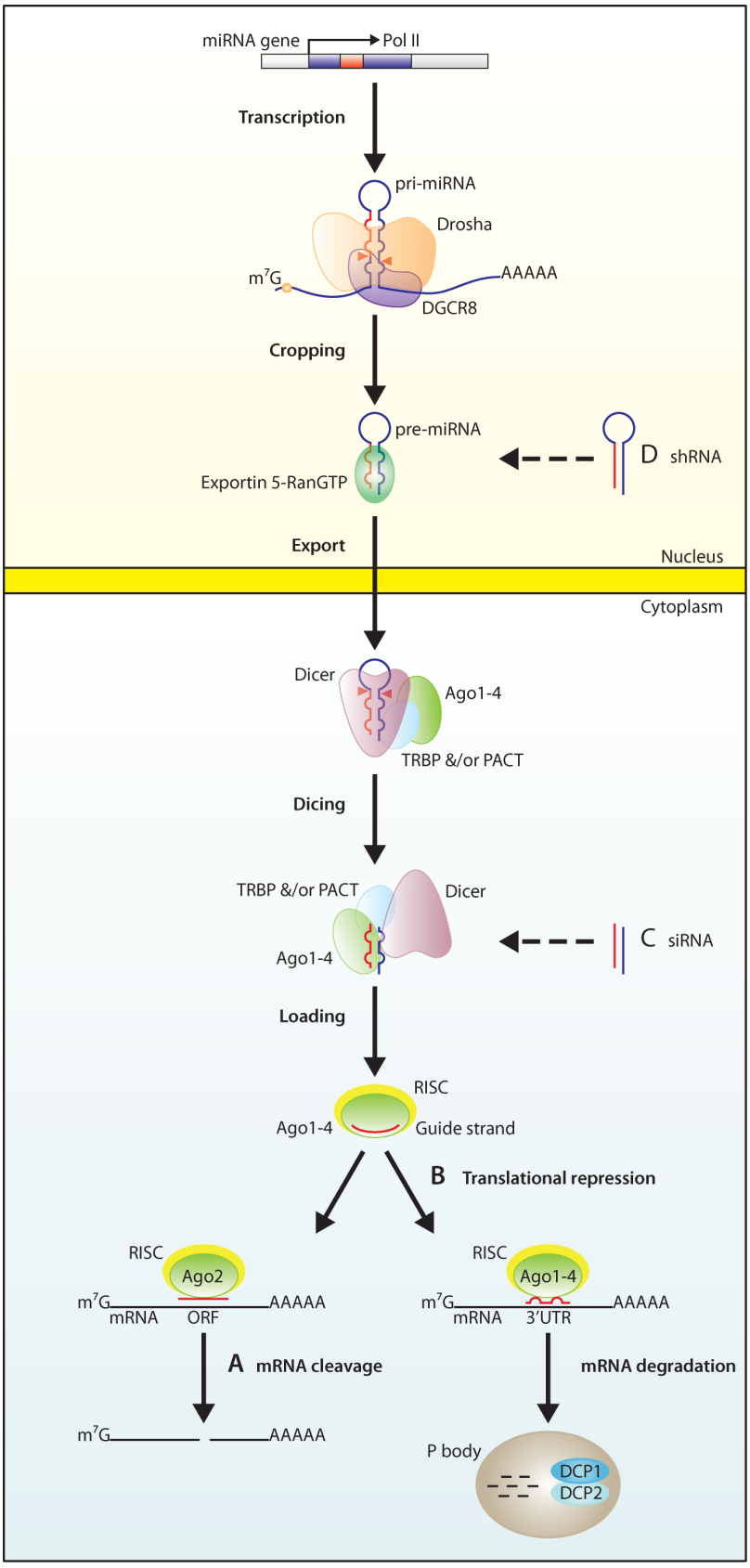


Figure 1.4: miRNA biogenesis in mammals and RNAi mimics. miRNA genes are typically transcribed by RNA Pol II generating pri-miRNAs that are processed in the nucleus by the microprocessor complex, which contains Drosha and DGCR8. The resulting pre-miRNA is exported to the cytoplasm by exportin 5 where Dicer-mediated processing generates miRNA duplexes. Duplexes are loaded into Ago proteins, which may be facilitated by TRBP and/or PACT. Functional RISC is formed on removal of the passenger strand, which can mediate mRNA cleavage where miRNAs are

perfectly complementary to their targets (A). More commonly, miRNAs mediate translational suppression of target mRNAs (B) and their degradation in P bodies. Exogenous siRNAs can be introduced into cells that may be incorporated into RISC (C). DNA plasmids may also be introduced resulting in the expression of shRNAs, which imitate pre-miRNAs, entering the pathway at the level of Dicer processing (D).

shRNA expression (Geiduschek and Kassavetis 2001; Schramm and Hernandez 2002; Boudreau *et al.* 2008). RNA Pol II promoters have the advantage of conferring greater transcriptional control, with the option of introducing tissue-specific and inducible systems. They are more amenable to the production of pri-miRNA mimics, where the length of ss sequences flanking the hairpin does not need to be precise, but >40 bp (McManus *et al.* 2002; Zeng *et al.* 2002).

The small size of the seed region means that RNAi mimics require careful design to avoid off-target effects. G:U wobbles may be introduced to increase the likelihood of desired strand selection (Khvorova *et al.* 2003; Schwarz *et al.* 2003). Both siRNAs and shRNAs are designed to be perfectly complementary to their targets and so typically are loaded into Ago2 and mediate target mRNA cleavage. Substrate requirements must be met and, for this reason, shRNAs and pri-miRNA mimics are usually designed by using endogenous pre- or pri-miRNAs as scaffolds, replacing the guide strand sequences with an artificial sequence that is complementary to the desired target (McManus *et al.* 2002; Zeng *et al.* 2002).

Although RNA Pol II promoters offer the most control and generate mimics less likely to saturate the endogenous RNAi pathway from those produced from RNA Pol III U6 and H1 promoters (Grimm *et al.* 2006), their successful design requires more knowledge of endogenous sequences and good predictions of the RNA mimic's folding. Thus, although RNA Pol II-driven pri-miRNA mimics offer great therapeutic potential as a gene therapy, RNA Pol III-driven shRNAs and siRNAs remain the preferred mimics for potent gene silencing and therefore, in determining protein function. Consequently, validation of a role for a protein in HIV-1 replication typically uses siRNAs or shRNAs, and these mimics have also been employed in genome-wide screens for the identification of HDFs (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b).

1.3.4 RNAi screens for the identification of HDFs

A number of approaches for the identification of host factors required for HIV-1 replication is available. However, candidate gene association studies are limited by what is known, genome-wide association studies by the requirement for large sample sizes, and immunoprecipitations coupled to mass spectrometry by the difficulty in obtaining quality preparations for every HIV-1 protein (An and Winkler 2010; Raghavendra *et al.* 2010). Genome-wide RNAi screens are not necessarily hypothesis-driven and do not require genetic variation for

HDF detection. Together with the availability of genome-wide si/shRNA libraries and advances in high-throughput screening technologies, genome-wide reverse genetics approaches have become increasingly widespread in the investigation of host-virus interactions [reviewed in (Hirsch 2010)].

Four genome-wide RNAi-based screens for the identification of host factors of critical importance for the replication of HIV-1 were reported (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b) as well as a meta-analysis of the first three screens (Bushman *et al.* 2009). The screens differed in their approach (Table 1.1), in particular with respect to cell lines, virus, RNAi library, assay length and readout, and filtering criteria.

Surprisingly, little overlap was observed between these first three studies (<7% for any pairwise combination) (Bushman *et al.* 2009). Bushman *et al*. attributed the lack of overlap to experimental noise, differences in timing of sampling and in the hit filtering criteria. Analysis of a screen duplicate (Konig *et al.* 2008) revealed that experimental variance was such that only 150 of 300 hits would be expected if the screen were repeated under identical experimental conditions. It is perhaps unsurprising, therefore, that the overlap was small given that the screens used different, albeit at times similar, conditions. Indeed, the small overlap was shown to be statistically significant (Goff 2008; Bushman *et al.* 2009). Bushman *et al*. also demonstrated that the screens were more robust at identifying enriched networks than individual proteins involved in HIV-1 replication. Using the Database for Annotation, Visualisation and Integrated Discovery (DAVID) (Huang da *et al.* 2009) they identified gene ontology groups enriched in all three screens, including the nuclear pore, ubiquitin/proteasome, mediator complex, transcription and transcriptional elongation, RNA binding/splicing, mitochondrial function and chaperones.

These first three screens used cells chosen for experimental convenience and are not the natural targets of the virus. Thus, genes required for replication in T cells but not in the cells used in the screens would be missed. Conversely, genes found to be significant in the screens may be of little import in T cells. The first three screens also all employed transfection of siRNAs as the mechanism of gene silencing. Given the time frame of the assays used, this approach may have been insufficient to decrease stable targets. This is likely to account for the notable absence of a well-characterised partner of HIV-1 IN, PC4 and SFRS1 interacting protein 1 (PSIP1)/lens epithelium-derived growth factor (LEDGF)/p75, from all three screens. The fourth study by Yeung *et al*. attempted to address these issues by performing a screen in Jurkat cells, a T cell-derived line, with stable expression of RNA Pol III-driven shRNAs delivered by lentiviral vector (Table 1.1). 252 genes were identified as putative HDFs, again displaying little overlap with the first three screens at the level of individual proteins but enrichment for similar cellular processes (Yeung *et al.* 2009b).

Table 1.1: Overview of experimental details of four genome-wide RNAi screens for the identification of HDFs.

	Brass *et al*. (2008)	**König *et al*. (2008)**	**Zhou *et al*. (2008)**	**Yeung *et al*. (2009)**
Scope	Early and late stage	Early stage (following entry)	Early and late stage	Early and late stage
RNAi library	siRNA targeting 21,121 genes 4 siRNAs/gene, 4 siRNAs/well	siRNA targeting 19,628 genes 4-6 siRNAs/gene, 2 siRNAs/well	siRNA targeting 19,709 genes 3 siRNAs/gene, 3 siRNAs/well	shRNA targeting 54,509 transcripts lentivirus delivery, pooled library
Replicates	Duplicates	Duplicates, triplicates	Single values	n/a
Virus	HIV-IIIB	HIV NL4-3Δenv-Δnef	HIV-HXB2	HIV NL4-3
Cells	TZM-bl (HeLa-derived)	HEK293T	P4/R5 (HeLa-derived)	Jurkat
siRNA concentration	50 nM/well, 12.5 nM/siRNA	25 nM/well, 12.5 nM/siRNA	53 nM/well, 17.6 nM/siRNA	n/a
Assay length (transfection to readout)	120/144 h	72 h	72/120 h	3 weeks selection + 1/4 weeks infection
Assay readout	Tat-dependent β-gal reporter cell line, % p24+ cells	Luciferase reporter virus	Tat-dependent β-gal reporter cell line	Barcoded shRNAs by microarray
Normalisation	Plate mean	Plate mean	Negative control	n/a
Hit selection	≥2 x SD below plate mean	Redundant siRNA analysis, decision matrix	≥2 x SD below negative control	n/a
Toxicity assay	Cell count	ATP content	Reducing capacity	n/a
Validation criteria	≥1 siRNA/gene confirmed	≥2 siRNA/gene confirmed	≥2 siRNA/gene confirmed	n/a
# of filtered hits	273	295	224	252

Another screen used an approach similar to that of Yeung *et al*., in that Jurkat cells were screened using stably expressed shRNAs. However, this screen was not genome-wide, but rather the library of shRNAs was enriched for those targeting human kinases and phosphatases, since these are druggable proteins with critical roles in the regulation of eukaryotic cell signalling pathways (Rato *et al.* 2010). By reducing the scope of the screen, this study successfully reduced the information load whilst increasing information density as experimental noise was decreased. However, reducing the scope was also the source of this study's greatest limitation.

Since experimental differences have such an important influence on screen results, guidelines have been produced by different groups, in accordance with the Minimum Information About an RNAi Experiment (MIARE) guidelines (www.miare.sourceforge.net), to aid researchers in designing RNAi screens for the identification of HDFs (Borner *et al.* 2010; Pache *et al.* 2010). Nevertheless, RNAi screens are always going to have their limitations. The cells used are not the natural targets of HIV-1 replication. This is of particular importance when it is appreciated that enhanced expression of HDFs augments HIV-1 replication in neonatal mononuclear cells (Sundaravaradan *et al.* 2010). Thus, differences in gene expression between different cell types will affect HIV-1 replication kinetics and, possibly, the relative import of HDFs. Although some success has been obtained with *in vivo* virogenomics studies (Aich *et al.* 2007; Cogburn *et al.* 2007), this is not readily applicable to HIV-1. Other limitations of RNAi screens include that HDFs with redundant functions will be overlooked, as may those with a long half-life, such as PSIP1/LEDGF/p75 (Llano *et al.* 2006a). In addition, current RNAi libraries do not interrogate ncRNAs. Off-target effects of siRNAs may give rise to false positives and where the RNAi mimics do not induce sufficient suppression, false negatives. Additional false negatives may arise from filtering processes discarding HDFs whose depletion is toxic and an inability to identify HDFs with purely an extracellular role in HIV-1 replication. Moreover, RNAi screens provide no quantification of the population variation in HDF gene expression. This is notable given that genes with the greatest within-population expression variability are enriched for chemokine signalling and HIV-1 interacting proteins, suggesting that variations in the expression of HDFs may account for differential HIV-1 susceptibility (Li *et al.* 2010). The complexity of virus-host interactions suggests a combination of platforms may need to be exploited for greater clarification of the role of individual cellular factors. Despite these limitations, RNAi screens offer huge promise when used as a primary genetic screen and when it is acknowledged that the design of specific sets of experiments are crucial to translating the collected genomics data into palatable knowledge.

1.3.5 RNAi-based validation of HDFs

Clearly, hits generated in RNAi screens for the identification of HDFs require validation. Brass *et al*. were aware of this and so confirmed the requirement for four of their hits, Rab6, Vps53, TNPO3 and Med28, in HIV-1 replication by generating cells with stable shRNA expression and performing additional experiments (Brass *et al.* 2008). Admittedly, validation would have been preferable in cell types that more closely resemble the natural targets of HIV-1 infection.

RNAi has been used for some time to validate HDF roles for cellular factors. One of the first host factors shown to inhibit HIV-1 replication on siRNA-mediated suppression was CD4 (Novina *et al.* 2002). Numerous studies have demonstrated a resistance to HIV-1 replication on RNAi-mediated suppression of coreceptors CCR5 and CXCR4 (Martinez *et al.* 2002; Anderson *et al.* 2003a; Anderson *et al.* 2003b; Qin *et al.* 2003; Song *et al.* 2003; Anderson and Akkina 2005a; Anderson and Akkina 2005b; Shimizu *et al.* 2009; Shimizu *et al.* 2010). RNAi-mediated silencing has confirmed many other cofactors, including those involved in integration, such as barrier to auto-integration factor 1 (BAF1), emerin (Jacque and Stevenson 2006) and PSIP1/LEDGF/p75 (Maertens *et al.* 2003; Llano *et al.* 2006a), transcription, such as NF-κB (Surabhi and Gaynor 2002), p21-activated kinase (PAK1) (Nguyen *et al.* 2006), cyclin T1 and CDK9 (Chiu *et al.* 2004), and envelope protein maturation, such as furin (Nguyen *et al.* 2006). The examples given have all been well-characterised HDFs for some time. Few of the novel HDFs identified in screens have been validated and those studies which do attempt to do so rarely progress to a T cell-derived line, presumably because of the technical difficulties associated with their transfection necessitating the generation of stable lines through lentiviral delivery of RNAi mimic expression cassettes. One notable exception emerged recently from the Berkhout group: they selected 30 putative HDFs and examined their ability to inhibit HIV-1 replication in T cell-derived lines with stable shRNA expression (Eekels *et al.* 2011). Only half were less susceptible to HIV-1 infection, determined by CA/p24 levels 7 days post-infection, demonstrating the importance of screen hit validation. The anticipated flurry of publications on HDF confirmation following the publication by Brass *et al*. (Brass *et al.* 2008) has not materialised, with instead the global focus seemingly on refining RNAi-based genome-wide screens for HDF identification. Clearly, there is still a great deal of work for molecular biologists as other host factor targets as promising as CCR5 may still be awaiting discovery. Validation of HDFs and examination of their potential as therapeutic targets was, therefore, examined in this thesis through exploitation of the RNAi pathway in T cell-derived lines.

1.3.6 RNAi-based gene therapy

As well as the identification and validation of HDFs, exploitation of the RNAi pathway has been used as a mode of gene therapy. The majority of anti-HIV RNAi-based gene therapies have been directed against the virus itself [reviewed in (Barichievy *et al.* 2009; Berkhout 2009)]. Unfortunately, virus-targeted RNAi is likely to be hampered by the same obstacle as HAART, that of viral escape. Furthermore, the sequence specificity of RNAi may favour more rapid viral escape than treatment with small molecule enzyme inhibitors (Boden *et al.* 2003), the latter requiring several mutations to confer amino acid changes that result in drug resistance. Consequently, there has been a movement toward combinatorial RNAi gene therapies, with estimations that simultaneous targeting of four viral targets is necessary to severely delay viral escape (Berkhout 2004; Ter Brake and Berkhout 2005). A number of combinatorial RNAi approaches has been developed, including shRNA multimers, polycistronic miRNA shuttles, bifunctional shRNAs and long hairpin RNAs (lhRNAs), which aim to produce more than one siRNA from a single expression cassette [reviewed in (Grimm and Kay 2007)]. However, therapies may be simplified and reduce viral escape still further by introduction of an element targeting a HDF.

The potential of HDF-targeting gene therapy was dramatically demonstrated by the apparent cure of a HIV-infected patient who had leukaemia in addition to AIDS. This patient received a bone marrow transplant in which donor stem cells were homozygous for the CCR5 Δ32 allele. CD4+ T cells were successfully reconstituted in the patient without signs of HIV-1 infection, despite discontinuation of antiretroviral therapy (Hutter *et al.* 2009; Allers *et al.* 2010). This approach is certainly not applicable to all HIV-infected individuals, however, as bone marrow transplantation is a high-risk treatment (ie only 10-30% of recipients survive).

Engraftment of autologous cells treated *ex vivo* with an HDF-targeting RNAi mimic introduced by a viral vector may be a more broadly applicable strategy. This may be possible with isolated peripheral blood mononuclear cells (PBMCs), although in this case the modified T cells would have a limited life span, such that repeat infusions would be necessary. The modification of haematopoietic stem cells would be more durable, though more complicated, leading to T cells, monocytes, macrophages and DCs less susceptible to HIV infection. The presence of an anti-HDF RNAi mimic would be expected to confer a survival advantage to modified cells over those unmodified in HIV-infected individuals. Proof of principle has been demonstrated in non-human primates that received blood stem cells treated with an SIV lentiviral vector expressing a shRNA against CCR5 (An *et al.* 2007). These primates exhibited normal haematopoietic reconstitution and T cells were more resistant to SIV infection. A similar approach was used in the humanised bone marrow/liver/thymus (BLT) mouse model, which received human haematopoietic stem cells modified by lentiviral delivery of anti-CCR5 shRNA. These mice showed decreased HIV-1 replication and no defects in T cell development (Shimizu *et al.* 2010). Another approach, which

avoids technically difficult engraftments, is the delivery of RNAi mimics *in vivo*. Obtaining RNAi mimic delivery is an ongoing focus of research in the field of gene therapy with many approaches being developed, many of which address the issue of tissue-specificity [reviewed in (Tiemann and Rossi 2009; Lares *et al.* 2010)].

Even if delivery problems are solved, there are a number of concerns related to RNAi-based gene therapies, whether targeting viral or HDF transcripts. Exogenous siRNAs and shRNAs produce guide strands similar to miRNAs and, therefore, have the potential to silence hundreds of off-target genes through their seed regions (Jackson *et al.* 2003; Jackson and Linsley 2004; Jackson *et al.* 2006). Although alignment algorithms and microarray analyses can be used to predict and monitor off-target effects, respectively, this is hampered by continued expansion in the number of known functional ncRNAs.

High levels of shRNA expression from RNA Pol III promoters can lead to toxic effects through competition with endogenous miRNAs and saturation of the RNAi pathway (An *et al.* 2006; Grimm *et al.* 2006; Castanotto *et al.* 2007). The fatal side effects observed in mice treated *in vivo* with shRNAs were attributed to saturation of exportin 5 (Grimm *et al.* 2006). Saturation has also been observed at the level of Ago proteins (Grimm *et al.* 2010), suggesting that siRNAs could also mediate toxicity through competition with the endogenous pathway.

RNAi-based gene therapies are also complicated by the potential for mimics to activate an immune response. Exogenous RNAs may stimulate Toll-like receptors (TLRs), leading to a signalling cascade that activates the interferon pathway (Takeda and Akira 2005). Indeed, a major setback to the field of RNAi was the discovery that vascular endothelial growth factor A (VEGF-A)-targeting siRNAs alleviated macular degeneration through sequence- and target-independent effects on TLR3 (Kleinman *et al.* 2008). Admittedly, this raises the issue of whether innate immune activation by RNAi mimics is good or bad. In some settings it is certainly favourable: interferon induction has been used synergistically with siRNAs for the destruction of tumour cells (Pastor *et al.* 2010). However, in the context of HIV-1 infection, destruction of cells expressing anti-HDF RNAi mimics would be an unwanted side effect. Expressed shRNAs are less prone to interferon pathway activation as they are not detected by cell surface receptors [reviewed in (Marques and Williams 2005)]. Nevertheless, these may activate a type I interferon response through cytoplasmic receptors dsRNA-dependent protein kinase (PKR) and retinoic acid-inducible gene 1 (RIG-1), which can lead to apoptosis (Karpala *et al.* 2005; Garcia-Sastre and Biron 2006; Hornung *et al.* 2006; Nallagatla *et al.* 2007). Some sequence motifs have been implicated (Pebernard and Iggo 2004), and thus this innate response may be prevented if these are avoided, along with PKR and RIG-1-activating 5' triphosphates, which are present in RNA Pol III-derived shRNAs. Capped RNA Pol II-derived transcripts and tRNA-shRNA fusions offer safer alternatives in this regard

[reviewed in (Rossi 2008)], but their use is limited by the difficulty associated with their successful design. Expression of shRNAs in haematopoietic stem cells does not induce interferon responses (Robbins *et al.* 2006), although this returns to the difficulties associated with modified stem cell introduction (see above).

One of the major drawbacks of RNAi-based gene therapy is the requirement for continual administration/expression of the RNAi mimic for sustained target suppression. This is a result of the targets being mRNAs whose transcription is unaffected by the mechanism of gene silencing. Continuous administration of RNAi mimics and complex engraftments are not conducive to resource-poor regions, however. Thus, a more favourable gene therapy regimen would entail reduced intervention frequency, which would be practically easier to administer to patients on a large scale. In addition, reduced intervention would decrease the exposure of cells to toxicities associated with off-target effects, pathway saturation and innate immune response activation. Fortunately, another arm of the RNAi pathway has become evident in mammals with the potential to mediate long-term gene silencing through modulation of transcription.

1.4 Transcriptional gene silencing

Small RNA molecules bound to Ago proteins have been shown to modulate chromatin and affect gene expression by TGS. First described in plants and yeast [reviewed in (Matzke and Birchler 2005; Buhler and Moazed 2007; Grewal and Elgin 2007)], TGS in mammalian cells was initially met with controversy; partly a result of the paucity of endogenous effectors identified. Guide strands may be loaded into a nuclear form of RISC, termed the RNA-induced transcriptional silencing complex (RITS) (Verdel *et al.* 2004). Like RISC, Ago proteins are the core components of the complex and targeting is mediated by complementary base-pairing of the RNA guide strand. Except for budding yeast, small RNA-mediated heterochromatin regulation is conserved throughout eukaryotes.

The first report of mammalian TGS utilised exogenous siRNAs to silence expression of an integrated GFP reporter driven by the eukaryotic translation elongation factor 1 alpha (EF-1α) promoter (Morris *et al.* 2004). Morris *et al*. confirmed silencing at the transcriptional level by performing nuclear run-on analysis, the gold standard for differentiating silencing effects mediated by TGS from those of the canonical RNAi pathway that involves mRNA targeting and post-transcriptional gene silencing (PTGS). The same strategy was employed to obtain TGS at the endogenous E-cadherin promoter a year later (Ting *et al.* 2005). Both reports implicated a role for epigenetic changes in RNA-induced gene silencing, which have the capacity to be inherited

through multiple rounds of cell division, underscoring the potential of TGS to induce sustained target suppression.

1.4.1 Epigenetic changes associated with TGS

The formation of chromatin through the binding of histones to DNA both restricts access to the genetic code and creates a mechanism by which access is regulated. The nucleosome is the basic unit of chromatin, around which DNA is coiled (Kornberg and Klug 1981). Each nucleosome consists of a core of two copies of histones H2A, H2B, H3 and H4. Approximately 147 bp of DNA is wrapped around each nucleosome (Luger *et al.* 1997). Between core units lies the linker histone, H1, which may be involved in higher-order compaction of DNA. The histones are subject to post-translational modifications, including acetylation and methylation, particularly on N-terminal tails that protrude from the core, mediated by histone acetylases/deacetylases (HATs/HDACs) and histone methyltransferases/demethylases (HMTs/HDMTs), respectively (Strahl and Allis 2000). Histone variants, such as H2A.Z and H3.3, provide additional variety to the nucleosome (Verdel *et al.* 2004). These modifications and variants can recruit specific proteins and also alter the binding affinity between histones and DNA. Binding affinities affect chromatin compaction and, thus, access of the transcription machinery to the DNA. Therefore, histone modifications mediate epigenetic regulation of transcription and may be silencing and activating, depending on the residue and type of modification.

Another common type of epigenetic regulation involves DNA methylation, particularly at CpG dinucleotides. Most CpG-rich regions (CpG islands) overlap promoters, where DNA methylation is associated with gene silencing (Suzuki and Bird 2008). Whilst DNA methylation is largely maintained by DNA methyltransferase (DNMT) 1 (Weber *et al.* 2007), a number of proteins mediate *de novo* DNA methylation, including DNMT3A and DNMT3B (Okano *et al.* 1999).

Identification of the protein complexes involved in the epigenetic changes that accompany mammalian TGS is ongoing. In the seminal TGS article (Morris *et al.* 2004), the ability of a promoter-targeted siRNA to induce TGS was reversed by treatment with trichostatin A (TSA) or 5-azacytidine (5-azaC), implicating roles for histone deacetylation and DNA methylation, respectively. Ting *et al* went further by demonstrating that TGS was accompanied by enrichment for histone 3 lysine 9 (H3K9) dimethylation (Ting *et al.* 2005), a chromatin modification associated with inactive promoters. Many studies have shown a role of core histone post-translational modifications in TGS (Ting *et al.* 2005; Kim *et al.* 2006; Weinberg *et al.* 2006; Han *et al.* 2007; Wang *et al.* 2007; Weinberg *et al.* 2007; Gonzalez *et al.* 2008; Kim *et al.* 2008; Lim *et al.* 2008; Morris *et al.* 2008; Suzuki *et al.* 2008; Hawkins *et al.* 2009; Hong *et al.* 2009; Tan *et al.* 2009b;

Turunen *et al.* 2009; Palanichamy *et al.* 2010; Yue *et al.* 2010). Enzymes responsible have been elucidated over the years through their enrichment at targeted promoters, as determined by chromatin immunoprecipitation (ChIP). These include the Polycomb Group Protein EZH2 (Kim *et al.* 2006; Kim *et al.* 2008; Hawkins *et al.* 2009), which trimethylates histone 3 lysine 27 (H3K27), HDAC1 (Suzuki *et al.* 2008; Hawkins *et al.* 2009) and possibly the H3K9 methyltransferase G9a (Hawkins et al. 2009). Such experiments have also shown that the roles of proteins differ in the establishment and maintenance of TGS (Hawkins *et al.* 2009).

Although Ting and coworkers observed induction of silent-state histone modifications at the targeted E-cadherin promoter, DNA methylation was not found at the targeted region (Ting *et al.* 2005). Thus, it was surprising that the first epigenetic modifier described associating with a TGS-inducing RNA was flag-tagged DNMT3A (Weinberg *et al.* 2006). The controversy associated with the requirement for DNA methylation in TGS has been fuelled by several conflicting reports (Castanotto *et al.* 2005; Janowski *et al.* 2005; Pulukuri and Rao 2007; Napoli *et al.* 2009). An appealing hypothesis, suggested by two studies that employed stable shRNA expression, is that DNA methylation follows histone modifications where guide strand expression is sustained (Kim *et al.* 2007; Hawkins *et al.* 2009). Given that most TGS studies employ transient expression of siRNAs, it is not surprising that DNA methylation is not always observed. Nevertheless, several observations question the importance of DNA methylation in small RNA-induced transcriptional regulation: (1) a DNMT3A association with endogenous short RNA guides has not been demonstrated using immunoprecipitation techniques (Park *et al.* 2010); (2) periodic, strand-specific cyclical DNA methylation at active promoters has been described, which might introduce artifactual observations (Metivier *et al.* 2008); and, (3) it is possible to reactivate aberrantly silenced tumour suppressor genes without loss of DNA hypermethylation (Pruitt *et al.* 2006). These findings bring into question the role of DNA methylation as a 'marker' of sustained TGS.

Despite the controversies surrounding the role of DNA methylation, the epigenetic changes that accompany TGS underlie one of the advantages of this approach over PTGS as a therapeutic strategy – it may preserve silencing even after removal of the promoter-targeted small RNA. Hawkins *et al* have shown that only 3 days of promoter-targeting shRNA induction in dividing cells, and 2 days in serum-starved, non-dividing cells, is required to achieve long-term silencing of ubiquitin C (>30 days), in a manner dependent on DNA and histone methylation (Hawkins *et al.* 2009). Thus, providing silent-state chromatin is attained at the targeted promoter, small RNAs can induce sustained target suppression.

1.4.2 Mechanism of small RNA-induced TGS in mammals

The observation that siRNA-induced silencing at the EF-1α promoter was sensitive to the RNA Pol II inhibitor α-amanitin (Weinberg *et al.* 2006), led to the proposal of two models to explain the mechanism of TGS: (1) the siRNA binds to DNA, facilitated by the opening of the DNA duplex by the transcription machinery; and, (2) the siRNA binds to nascent promoter-associated RNA. Work by Han *et al* favoured the latter after revealing a requirement for a promoter-associated transcript (Han *et al.* 2007). The antisense strand of the siRNA was shown to associate with a low-copy EF-1α-coding transcript with an extended 5' UTR. In addition, the association was inhibited by RNase A but resistant to RNase H treatment, implicating an RNA-RNA interaction in TGS.

Other studies have also demonstrated a requirement for promoter-associated transcripts for small RNA-mediated gene silencing. Some of these studies have shown, as at the *EF-1α* promoter, that silencing requires sense transcripts (Hawkins *et al.* 2009; Tan *et al.* 2009b; Yue *et al.* 2010), whilst others have found antisense transcription through the targeted region is necessary (Gonzalez *et al.* 2008; Schwartz *et al.* 2008). Several reports have shown the presence of both sense and antisense promoter transcripts (Gonzalez *et al.* 2008; Napoli *et al.* 2009), although their detection is not always accompanied by confirmation of their necessity in TGS (Kim *et al.* 2008; Napoli *et al.* 2009; Palanichamy *et al.* 2010). The relative contributions of sense and antisense promoter transcripts to RNA-induced TGS requires further clarity, particularly as their relative levels may affect outcome (Morris *et al.* 2008; Schwartz *et al.* 2008).

Despite these complications, the findings support a model for TGS in which small RNA and Ago-containing RITS complexes are guided to their targets by complementary base pairing with low-copy, promoter-associated RNA Pol II transcripts. This serves as a scaffold for the recruitment of epigenetic-modifying protein complexes that favour heterochromatin formation and, therefore, decreased transcription of the targeted gene (Figure 1.5). Experiments in *Schizosaccharomyces pombe* support the ability of nascent RNA transcripts to direct heterochromatin formation; remodelling occurred where RNAi-associated protein complexes were tethered artificially to nascent transcripts (Buhler *et al.* 2006). The model has gained general acceptance, facilitated by characterisation of endogenous miRNA TGS effectors (Gonzalez *et al.* 2008; Kim *et al.* 2008; Tan *et al.* 2009b). However, a different mechanism for TGS has also been proposed that does not involve any epigenetic modifications at all (Janowski *et al.* 2005; Janowski *et al.* 2006; Schwartz *et al.* 2008; Napoli *et al.* 2009).

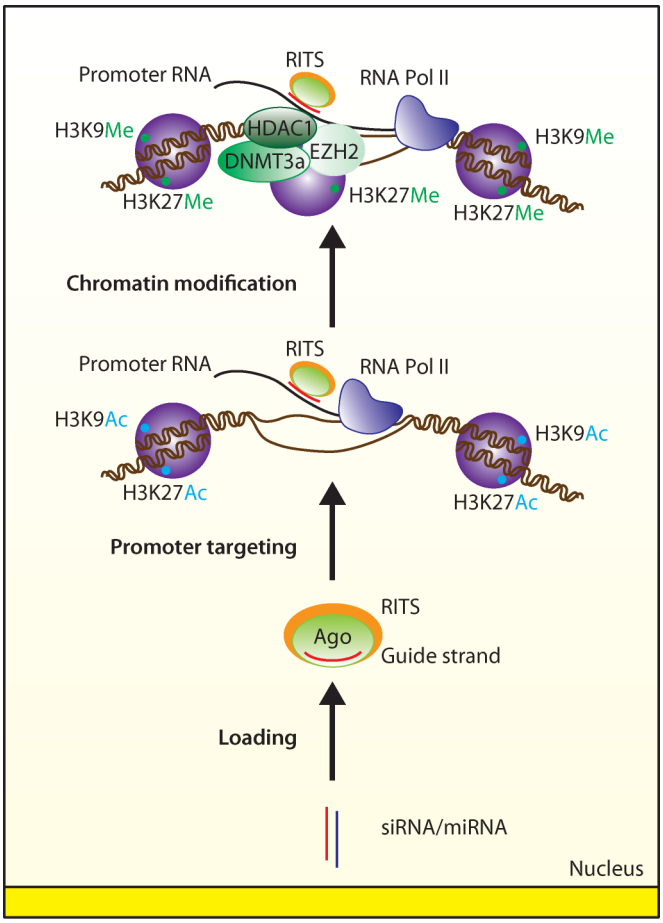


Figure 1.5: Proposed mechanism of TGS. siRNA or miRNA duplexes are loaded into Ago-containing RITS. These complexes are directed to low-copy promoter-associated RNA Pol II transcripts through complementary base pairing, leading to the recruitment of epigenetic modifying proteins. These increase silent-state histone modifications and, where long-term silencing is established, DNA methylation. The resulting increase in nucleosome occupancy and chromatin compaction reduces the accessibility of the promoter to the transcription machinery and, consequently, gene transcription.

1.4.3 Transcriptional interference

The Corey group has focused on targeting the region surrounding the transcription start site (TSS), as opposed to upstream regions, which were targeted in preceding descriptions of mammalian TGS (Morris *et al.* 2004; Ting *et al.* 2005). They reasoned that an open complex is generated at the TSS by the binding of RNA Pol II, rendering it more susceptible to hybridisation with oligonucleotides (Milne *et al.* 2000). The TSS-targeting small duplex RNAs were referred to as antigene RNAs (agRNAs), rather than siRNAs, to differentiate them from those targeting mRNAs (Janowski *et al.* 2006). Corey's group designed agRNAs targeting the progesterone receptor (PGR) TSS (Janowski *et al.* 2005). Surprisingly, the observed silencing in T47D breast cancer cells was potent and not sensitive to either 5-azaC (Janowski *et al.* 2005) or TSA (Janowski *et al.* 2006), suggesting that the inhibition obtained was mediated by a mechanism that was independent of epigenetic changes, since termed transcriptional interference (TI). The mechanism of TI is unclear

but some studies have implicated RNA Pol II occlusion mediated by RNA-DNA-DNA triplex formation (Martianov *et al.* 2007; Jain *et al.* 2010).

Variations in epigenetic outcome between classical TGS (ie accompanied by epigenetic changes) and TI may depend on the Ago protein involved. However, this seems unlikely given that Ago1-3 have been associated with small RNA-induced heterochromatin formation (Gonzalez *et al.* 2008; Hawkins *et al.* 2009; Tan *et al.* 2009b; Turner *et al.* 2009; Turunen *et al.* 2009; Giles *et al.* 2010; Hawkins and Morris 2010), and Ago1 and Ago2 also with TI (Janowski *et al.* 2006; Napoli *et al.* 2009; Chu *et al.* 2010). Indeed, whether a distinct TI mechanism exists is still debateable, particularly as the most recent report from the Corey group does demonstrate increased H3K27 trimethylation at the PGR promoter (Yue *et al.* 2010), in contrast to their previous work, despite using the same TSS-targeting agRNA and cell line. Further studies are required to elucidate whether there is a universal TGS mechanism and, if not, how RITS recruits different proteins to initiate different events and their functional consequence. Regardless, therapeutic application of TGS is likely to focus on mechanisms that result in heterochromatin formation since the potential for long-term silencing with TI is unknown.

1.4.4 Transcriptional gene activation

The apparent ability of promoter-targeting small duplex RNAs to associate with different proteins may underlie the observation that they can also direct transcriptional gene activation (TGA). The initial report of TGA showed increased expression of E-cadherin, p21 and VEGF-A on transfection of promoter-targeting siRNAs into cultured cells (Li *et al.* 2006), which could be prevented by suppressing Ago2 expression. However, these findings require further validation, since gene activation may also be the result of an off-target PTGS effect on a transcriptional repressor, a phenomenon first described by Weinberg *et al*. In this study, microarray analysis demonstrated that transcriptional activation with an antisense RNA targeting the HIV-1 LTR promoter was accompanied by decreased expression of a transcript that functioned indirectly in its regulation (Weinberg *et al.* 2007).

A different PTGS off-target effect was shown to cause the previously observed TGA at the p21 gene (Li *et al.* 2006). In this case, a suppressive antisense transcript overlapping the promoter was the unintended target of the TGA-inducing siRNAs (Morris *et al.* 2008). Whether the PTGS of the suppressive ncRNA occurred in the nucleus or cytoplasm is unknown, but it is interesting that nuclear PTGS has been confirmed in plants (Hoffer *et al.* 2011). Despite these controversies, reports of TGA continued, including identification of endogenous small RNA activators; transcriptional activation of E-cadherin and cold-shock domain-containing protein 2 (CSDC2) was

observed following over-expression of miR-373 (Place *et al.* 2008). However, there was no reciprocal E-cadherin and CSDC2 silencing on endogenous miR-373 knockdown. In addition, computational analysis revealed miR-373 targets within 1 kb of the TSS in >372 genes, but not all responded to miR-373 over-expression. Although the miR-373-mediated TGA may function through a direct interaction of the miRNA with the promoter leading to recruitment of epigenetic modifying complexes, such as HATs and HDMTs (Figure 1.6A), further experimentation is required to exclude off-target PTGS effects on a regulatory ncRNA (Figure 1.6B). Recently, miR-205 was shown to induce TGA of interleukin-24 (IL24) and -32 (IL32) (Majid *et al.* 2010). In this case, a luciferase reporter assay demonstrated that direct promoter targeting mediated the miRNA-induced TGA. It is unknown, however, whether this TGA occurs through direct recruitment of epigenetic modifying complexes (Figure 1.6A) or steric hindrance of transcriptional repressor docking with its cognate binding site in an enhancer region (Figure 1.6C), as has been described at the IL-10 gene (Matilainen *et al.* 2010). Further studies are required to examine the mechanisms of TGA and their significance to gene regulation *in vivo*.

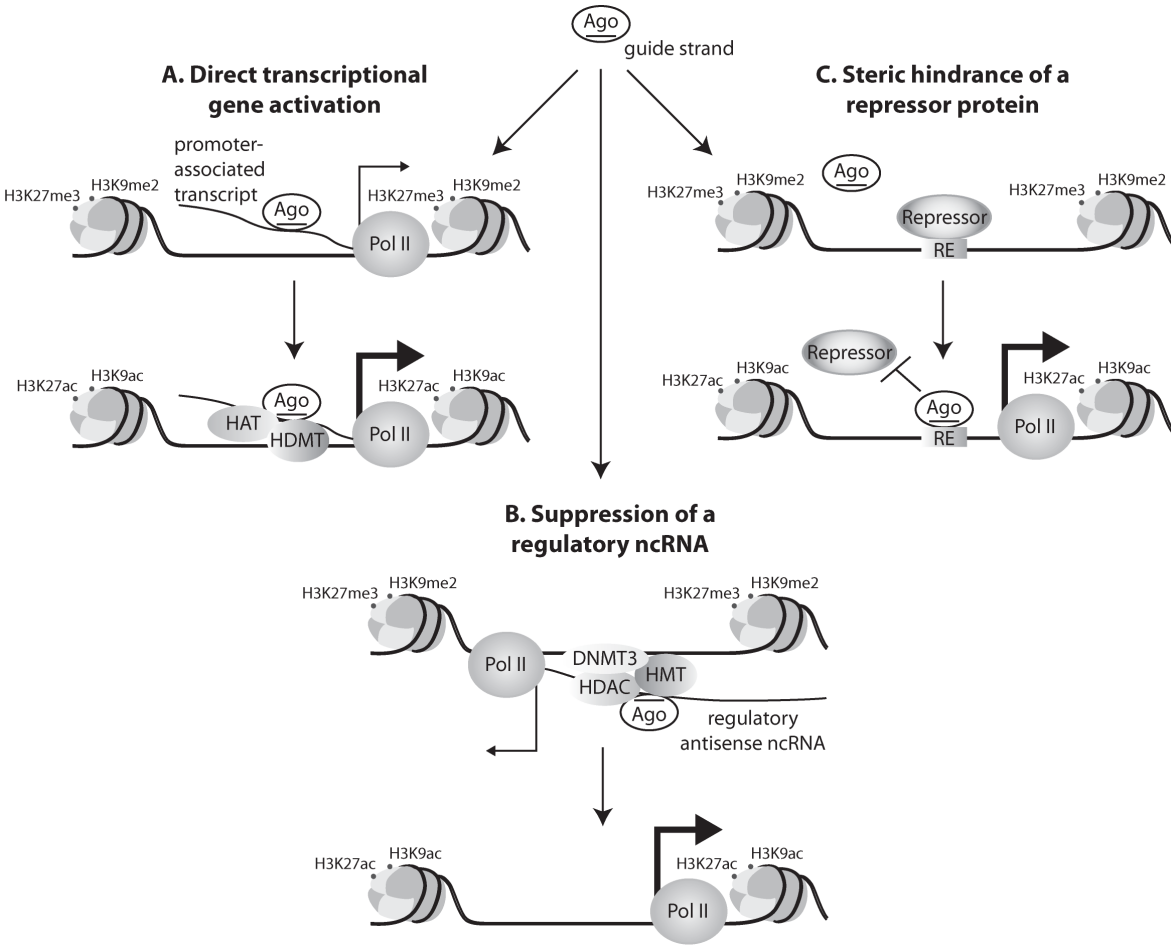


Figure 1.6: Potential mechanisms of TGA. Several mechanisms for the induction of TGA have been described. Ago-bound promoter-targeted small RNAs may induce TGA through direct recruitment of epigenetic modifying proteins (A) or through PTGS of a ncRNA that has an endogenous TGS function (B). Alternatively, TGA may be a result of steric hindrance of a repressor binding site in an enhancer region (C).

Given the number of long ncRNA-mediated gene regulatory pathways already identified, it is perhaps not surprising that different small RNA-mediated transcription modulation pathways exist, particularly as the short RNA transcriptome is probably at least as complex, if not more so, than that of long RNAs (Kapranov *et al.* 2007a). However, it remains unclear how similar small RNAs can induce TGS in one context and TGA in another, particularly as Turunen *et al.* were able to induce TGS and TGA of VEGF-A with promoter-targeting shRNAs within the same cell line. Therapeutic application of TGS and TGA will require a greater understanding of the potential for promoter-targeting RNAs to mediate opposing effects.

1.4.5 TGS-based gene therapy

Although the durability of TGS will vary between genes, the promoter-targeted small RNA will not require continual administration/expression if stable epigenetic changes are induced. Consequently, there is less likelihood of toxicity associated with saturation of the RNAi pathway, a concern with the PTGS strategy (Grimm *et al.* 2006; Grimm *et al.* 2010). In combination with the fact that targets are in limited numbers, as opposed to many copies of mRNA, TGS may provide a more efficient silencing strategy than PTGS.

The therapeutic application of small RNA-induced gene silencing has followed two distinct strategies: introduction of synthetic siRNAs/agRNAs and, secondly, use of gene expression cassettes whose transcription produces RNAi precursors or antisense RNAs. Synthetic agRNAs were recently shown to retain their activity, either silencing or activating transcription, despite introduction of chemical modifications, which may increase their stability and facilitate *in vivo* applications (Watts *et al.* 2010). In contrast to PTGS, siRNAs inducing TGS need to access the nucleus. Some studies suggest that efficient nuclear localisation of exogenous siRNAs requires administration at a high concentration – up to ten times that typically used to obtain PTGS (Janowski *et al.* 2005; Han *et al.* 2007; Hawkins *et al.* 2009). However, the reason for this is unclear, particularly as others have obtained TGS with comparable concentrations of siRNAs to that commonly used for PTGS induction (Janowski *et al.* 2006; Schwartz *et al.* 2008; Yue *et al.* 2010).

Using expressed TGS-inducing small RNAs ensures nuclear delivery. U6 promoter-driven expression of the antisense strand of a siRNA alone was sufficient to mediate TGS at the HIV-1 LTR (Weinberg *et al.* 2006). Although the antisense effector was able to increase Ago1 at the target site (Turner *et al.* 2009), dsRNAs that resemble RNAi pathway precursors may provide increased silencing because of more efficient incorporation into RITS than small antisense RNA alone.

Several groups have induced TGS through expression of RNAi mimics. shRNAs have successfully been used to transcriptionally silence human RASSF1A (Castanotto *et al.* 2005; Kim *et al.* 2006), IL2 (Murayama *et al.* 2006), ubiquitin C (Hawkins *et al.* 2009), VEGF-A (Turunen *et al.* 2009) and the HIV-1 LTR (Yamagishi *et al.* 2009). It is currently unknown exactly how expressed shRNAs are processed: if Dicer processing and Ago loading occurs in the cytoplasm, RITS must undergo nuclear import. Although such import has been described in *Tetrahymena* (Noto *et al.* 2010) and *Caenorhabditis elegans* (Guang *et al.* 2008), in humans the pathway has yet to be identified, although may involve importin 8 (Weinmann *et al.* 2009). Alternatively, processing and loading may occur within the nucleus itself. Ago proteins are present in human nuclei (Tan *et al.* 2009a; Weinmann *et al.* 2009) and nuclear Dicer activity was proposed (Lee *et al.* 2008). In addition, Dicer has been shown to shuttle between the nucleus and cytoplasm in fission yeast, with absence from the former preventing heterochromatin assembly (Emmerth *et al.* 2010) and its association with chromatin mediating gene silencing (Woolcock *et al.* 2011). Moreover, shRNA processing may not be required for TGS given that Ago2 has been shown to form an active RISC with pre-miRNAs in the nucleus (Tan *et al.* 2009a). Clarification on the cellular compartmentalisation of RNAi pathway components will enhance our ability to design TGS-inducing small RNAs and understand their relative potencies. More information on endogenous pathways that allow some miRNAs to modulate chromatin but restrict others to mRNA regulation in the cytoplasm will, hopefully, facilitate this.

1.4.6 TGS-based modulation of viruses

TGS has therapeutic potential in a number of diseases, including cancer (Pulukuri and Rao 2007) and those characterised by dysregulated hetrochromatin [reviewed in (Hahn *et al.* 2010)]. A number of studies have also shown its use in the transcriptional regulation of viruses. TGS has been induced with siRNAs at both the Human papilloma virus (HPV)-16 promoter (Hong *et al.* 2009) and enhancer (Palanichamy *et al.* 2010), leading to decreased production of the oncogenic E6 and E7 viral proteins and increased cell death and senescence. In addition, TGS has successfully been induced at both the SIV (Lim *et al.* 2008) and HIV-1 (Suzuki *et al.* 2008; Turner *et al.* 2009; Yamagishi *et al.* 2009) LTRs. However, similar to RNA-induced TGS at the viral LTR, latent provirus exhibits suppressed expression through heterochromatin formation and associations with heterochromatic chromosome regions in *trans* (Dieudonne *et al.* 2009; Tyagi *et al.* 2010). Thus, despite stable expression of LTR-targeted shRNA sustaining suppressed HIV-1 gene expression for >1 year in a CD4+ T cell line (Yamagishi *et al.* 2009), concerns persist over whether TGS of the HIV-1 promoter would in fact be detrimental to patients by increasing the latent reservoir of virus.

TGA may offer greater therapeutic benefit than TGS at the HIV-1 LTR through reversing latency, rendering replicating virus susceptible to antiviral drugs. Modulation of chromatin to eradicate latent viral reservoirs is not a new idea; an HDAC inhibitor, in combination with intensified antiviral drugs, has been used to deplete latent virus and accelerate clearance of HIV in patients (Lehrman *et al.* 2005). Small RNA-induced TGA may prove more practical than a HDAC inhibitor since effects would be specific to the viral promoter. However, TGA of the viral promoter has not yet been reported, except when inadvertently induced by off-target PTGS (Weinberg *et al.* 2007).

Given these difficulties, transcriptional modulation may be better employed at promoters of host factors required for HIV-1 replication, rather than at the viral promoter itself. TGS has been reportedly induced at CCR5 (Kim *et al.* 2006; Han *et al.* 2007), although the ability of silencing this coreceptor, or any other HDF, by this mechanism to inhibit HIV-1 replication has not been examined. Certainly, TGS of HDFs presents an exciting anti-HIV therapeutic strategy; it simultaneously presents a high genetic barrier to the development of drug-resistance and, providing sustained silent-state chromatin is achieved, reduces therapy administration, which fits with the realities of resource-poor nations and may reduce the toxicities typically associated with RNAi. However, the susceptibility of HDFs, other than CCR5, to TGS has not been investigated, nor the therapeutic potential of silencing HDFs by this mechanism. This thesis, therefore, examines whether HDF-targeted TGS is sufficiently potent to inhibit viral replication and identifies promoter features associated with TGS susceptibility.

1.5 HIV-dependency factors

Four cellular factors were selected for inclusion in this study, for validation as HDFs in a T cell-derived line and/or examination of their susceptibility to TGS. They were selected for different reasons, described in detail below (see 1.5.1-4). Briefly, both HIV-1 Tat specific factor 1 (HTATSF1) and DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked (DDX3X) were hits in genome-wide screens (Brass *et al.* 2008; Zhou *et al.* 2008) and, although their roles as HDFs were confirmed in subsequent studies (Ishaq *et al.* 2008; Miller *et al.* 2009), these also employed reporter cell lines and their potential as therapeutic targets has not been examined in T cell-derived lines. Controversy surrounds the requirement in HIV-1 replication for SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 (SMARCB1) (Bukrinsky 2006) and so this was selected to clarify whether this is a HDF. The fourth cellular factor, PSIP1/LEDGF/p75, is a well-characterised HDF with a role in integration of the provirus (Llano *et al.* 2006a). This was included in the study to serve as a positive control for effects of HDF

suppression on HIV-1 replication. None of the four cellular factors had been previously examined for their susceptibility to TGS.

1.5.1 HTATSF1

Transcription in eukaryotic cells requires the assembly of an elaborate multi-subunit pre-initiation complex, followed by initiation, elongation and termination. Activators of transcription conventionally mediate their regulation by increasing the rate of initiation. In contrast, based on *in vivo* and *in vitro* evidence, the HIV regulatory protein Tat predominantly stimulates transcription at the level of elongation (Kao *et al.* 1987; Laspia *et al.* 1989; Marciniak *et al.* 1990; Feinberg *et al.* 1991) on binding the TAR element (Berkhout *et al.* 1989; Cullen 1993).

Zhou and Sharp originally identified HTATSF1 as a requirement for Tat activation of elongation in a partially reconstituted transcription reaction isolated from HeLa nuclear extract (Zhou and Sharp 1996). These observations were confirmed by several studies: immunodepletion of HTATSF1 from nuclear extracts inactivated Tat stimulation of transcription and complementation with recombinant HTATSF1 restored transactivation (Li and Green 1998; Kim *et al.* 1999; Parada and Roeder 1999). HTATSF1 is a component of an RNA Pol II-containing complex that includes other elongation factors hSPT5 and pTEF-b (Parada and Roeder 1999). CDK9 of pTEF-b, along with phosphorylating RNA Pol II CTD, phosphorylates HTATSF1 (Zhou *et al.* 1998; Zhou *et al.* 2004a), although the functional significance of this is unknown. HTATSF1 also associates with the RAP30 subunit of TFIIF (Kim *et al.* 1999), which is tightly bound to elongating RNA Pol II. Since HTATSF1 also binds hSPT5, which associates with RNA Pol II, HTATSF1 may contribute toward stabilisation of the elongation complex. HTATSF1 is a general transcription elongation factor of cellular genes (Li and Green 1998; Chen *et al.* 2009). Thus, Tat takes advantage of an existing RNA Pol II complex to relieve the elongation block imposed on HIV-1 transcription, although it should be noted that the contribution afforded by HTATSF1 is not completely understood.

Sequence analysis of HTATSF1 revealed a highly acidic carboxyl-terminal half and an amino-terminal half containing two RNA recognition motifs (RRMs) (Zhou and Sharp 1996). In the former, 48% of the last 245 residues consist of glutamate or aspartate. This half also contains many serine residues that are arranged in a short peptide sequence matching consensus sites for phosphorylation, which may contribute additional negative charges to this region. The acidic carboxyl-terminal of HTATSF1 is required for Tat transactivation through an interaction with pTEF-b (Zhou *et al.* 1998).

Primarily through the first RRM domain, HTATSF1 has been shown to interact with spliceosomal U small nuclear ribonucleoproteins (snRNPs), an association that strongly stimulates

Tat transactivation and HIV-1 transcription elongation (Fong and Zhou 2001). This supports the notion that the recruitment of U snRNPs near the elongating polymerase is important for transcription. As the HTATSF1-U snRNP complex also stimulates splicing *in vitro* (Fong and Zhou 2001), HTATSF1 may serve as a dual-function factor to facilitate reciprocal activation of transcription and splicing – a putative mechanism for the requirement of HTATSF1 in HIV-1 transcription elongation. It has also been suggested, in contrast to most studies, that HTATSF1 functions solely in splicing, not elongation (Miller *et al.* 2009).

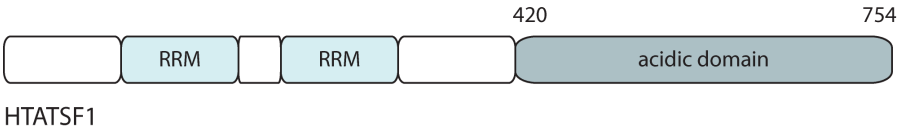


Figure 1.7: Domains of HTATSF1. The amino-terminal half contains two RRMs (light blue) and the carboxyl-terminal half a highly acidic domain (residues 420-754, dark blue).

HTATSF1 was a hit in one of four genome-wide RNAi screens for the identification of HDFs – siRNA knockdown of HTATSF1 resulted in a 69% decrease in p24 production (Brass *et al.* 2008). Together the evidence suggests a requirement for HTATSF1 in HIV-1 replication. However, much of the previous work has focused on immunodepletion *in vitro* and the siRNA screen results, not recapitulated in other screens, have yet to be verified in a relevant cell line. Moreover, the potential toxicity of suppressing this general transcription elongation factor has yet to be assessed. Thus, HTATSF1 still requires validation as an HDF and an anti-HIV therapeutic target.

1.5.2 DDX3X

HIV-1 gene expression is substantially governed post-transcriptionally by Rev-regulated export of unspliced/partially spliced viral RNAs from the nucleus to the cytoplasm (see 1.2.3) (Harris and Hope 2000). DDX3X was first in the family of DEAD box helicases found to play a role in HIV-1 replication through its function in the Rev/RRE-CRM1 export pathway. DEAD box proteins are superfamily 2 RNA helicases involved in many aspects of RNA metabolism, including transcription, mRNA splicing and RNA export [reviewed in (Rajkowitsch *et al.* 2007; Pan and Russell 2010)]. Some, such as DDX3X, use an ATP-dependent cycle of conformational changes to disrupt RNA structure, accelerating structural transitions of RNAs and RNPs, in a manner that bears a strong resemblance to the activities of certain protein chaperones. Consequently, a number of functions have been ascribed to DDX3X in both the nucleus and cytoplasm [reviewed in (Schroder 2010)], including regulation of cell growth through translational control of cyclin E1 (Lai *et al.* 2010).

HIV-1 does not encode a RNA helicase and, thus, hijacks an array of cellular RNA helicases. Studies have previously shown that suppression of DDX3X leads to decreased HIV-1 replication accompanied by a decrease in cytoplasmic RRE-containing viral transcripts (Yedavalli *et al.* 2004). Interestingly, DDX3X may be limiting for HIV-1 replication; CA/p24 production is slightly increased in Jurkat cells infected with a HIV-1 molecular clone in which Nef is replaced with DDX3X (Yedavalli *et al.* 2004). Furthermore, HIV-1 appears to modulate expression levels of DDX3X; expression is enhanced in Tat-expressing samples (Yedavalli *et al.* 2004) and varies between cells with latent and actively-replicating virus (Krishnan and Zeichner 2004). Upregulation of DDX3X for efficient HIV-1 replication suggests this protein is a good therapeutic target, since moderate suppression may not interfere with endogenous cellular pathways, thus minimising toxicity, whilst simultaneously inhibiting viral pathways.

DDX3X was identified as a HDF by two genome-wide siRNA screens performed in HeLa-derived cells (Brass *et al.* 2008; Zhou *et al.* 2008). Screen results were recently validated on demonstration that shRNA-mediated knockdown of DDX3X suppressed replication of an HIV-1 molecular clone without effecting cell viability (Ishaq *et al.* 2008). This supports that DDX3X may not be essential to the cell or, alternatively, that low level DDX3X is sufficient to maintain its putative function for cellular mRNA metabolism while higher levels are required for its function in HIV-1 RNA nuclear export. Unfortunately, this validation was conducted in HEK293T-derived cells. This is of particular importance when the multi-functionality of DDX3X is considered [reviewed in (Schroder 2010)]. At least four different viruses, namely hepatitis B virus (HBV), hepatitis C virus (HCV), HIV and poxviruses, encode proteins that interact with DDX3X. It is required for HCV replication (Angus *et al.* 2010; Sun *et al.* 2010) but, conversely, displays antiviral properties; DDX3X may function as a RIG-1-like receptor in the cytoplasm on recognition of viral RNAs, which leads to a type I interferon response (Mulhern and Bowie 2010; Oshiumi *et al.* 2010b). These antiviral effects are such that both HBV and poxvirus express proteins that interact with DDX3X, abolishing its ability to induce an innate immune response (Kalverda *et al.* 2009; Oda *et al.* 2009; Wang and Ryu 2010). Thus, although DDX3X has proved a prime target for viral manipulation because of it multi-functionality, its suppression may have positive or negative effects on cell proliferation and viral replication depending on the cell type and environmental conditions. Confirmation is required, therefore, of whether RNAi-mediated DDX3X suppression will be effective at inhibiting HIV-1 replication in a T cell-derived line, a model that more closely resembles the situation *in vivo*.

1.5.3 SMARCB1/INI1/SNF5

SMARCB1 is one of four core components of the mammalian SWI/SNF complex that is involved in ATP-dependent chromatin remodelling during transcriptional activation (Wang *et al.* 1996) and nucleotide excision repair (Ray *et al.* 2009). SMARCB1 is also known as SNF5 and integrase interactor 1 (INI1), the latter because it was the first host protein shown to directly interact with HIV-1 IN (Kalpana *et al.* 1994), an interaction that is HIV-specific among retroviruses (Yung *et al.* 2004; Rain *et al.* 2009). SMARCB1 contains three conserved regions including two imperfect repeats, Rpt1 and Rpt2, and a C-terminal coiled-coil domain (Morozov *et al.* 1998). The Rpt domains are involved in protein-protein interactions with both viral and cellular proteins and Rpt1 is necessary and sufficient to bind to HIV-1 IN (Wu *et al.* 1996; Morozov *et al.* 1998; Cheng *et al.* 1999; Lee *et al.* 1999). Residues 106-183 bind DNA non-specifically, whereas residues 183-243 comprise the IN binding domain (IBD), with T213, D224, D226 and S246 critical for IN interaction (Figure 1.8) (Yung *et al.* 2001). A masked nuclear export sequence (residues 266-276) within Rpt2 becomes functional when the C-terminal domain is deleted (Craig *et al.* 2002), facilitating SMARCB1 CRM1-dependent export.

Despite the early discovery of a SMARCB1-IN association, there is still controversy regarding the role of SMARCB1 in HIV-1 replication. Considerable evidence supports that the SWI/SNF complex is directly involved in Tat-mediated activation of HIV-1 transcription through remodelling of the +1 nucleosome (Ariumi *et al.* 2006; Mahmoudi *et al.* 2006; Treand *et al.* 2006), which is positioned immediately downstream of the transcription start site at the viral LTR. SMARCB1 has also been shown to interact with HIV-1 PICs (Turelli *et al.* 2001), RT complexes (Iordanskiy *et al.* 2006) and $Pr160^{GagPol}$ (Yung *et al.* 2001), and to be incorporated into HIV-1 virions (Yung *et al.* 2004; Sorin *et al.* 2009), although the functional significance of these associations, as well as the interaction with IN, is unclear. Recent evidence suggests that SMARCB1 association with the IN portion of $Pr160^{GagPol}$ is necessary for trafficking of the latter and HIV-1 particle assembly and budding (Cano and Kalpana 2010), and with IN as part of the PIC to facilitate initiation of latent or productive transcription programmes soon after integration (Boese *et al.* 2009).

Suppression of SMARCB1, or expression of mutants, has been shown to inhibit HIV-1 replication (Yung *et al.* 2001; Ariumi *et al.* 2006; Boese *et al.* 2009) and infectious particle production (Sorin *et al.* 2006; Cano and Kalpana 2010). In contrast, it has also been suggested that SMARCB1 interferes with early steps of HIV-1 replication (Maroun *et al.* 2006; Sorin *et al.* 2006). Such studies, along with the multiple interactions in which SMARCB1 engages, have compounded efforts to determine the therapeutic potential of targeting this cellular factor. In support of a HDF role, gene expression analyses reveal that, similar to DDX3X, SMARCB1 is up-regulated in actively- versus latently-infected cells (Krishnan and Zeichner 2004). However, several genome-wide RNAi

screens to detect HDFs did not identify SMARCB1 (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b). Given these conflicting reports and the fact that, with the exception of Yeung *et al*., studies examining the role of SMARCB1 in HIV-1 replication have not used a T cell-derived line, clarification of whether SMARCB1 is a HDF is necessary.

1.5.4 PSIP1/LEDGF/p75

PSIP1/LEDGF/p75 is a well-characterised HDF involved in HIV-1 integration. PSIP1 is neither lens-specific nor a growth factor but, rather, an ubiquitously expressed and constitutively nuclear transcription factor-type protein. Two splice variants, 530 amino acid p75 and 333 amino acid p52, possess unique C-termini (Figure 1.8) (Ge *et al.* 1998). Both isoforms function as transcriptional coactivators (Ge *et al.* 1998), although only p75 associates with HIV-1 IN directly (Cherepanov *et al.* 2003; Rain *et al.* 2009; Raghavendra *et al.* 2010).

A number of groups established the role of p75 in chromatin tethering and demonstrated that lentiviral IN proteins are cytoplasmic in the absence of p75 (Cherepanov *et al.* 2003; Maertens *et al.* 2003; Llano *et al.* 2004; Komano *et al.* 2005). Discrete functional modules govern various activities (Figure 1.8); nuclear localisation of p75 is the outcome of both NLS-mediated nuclear import (Maertens *et al.* 2003) and a dynamic scan-and-lock mechanism for chromatin trapping (Hendrix *et al.* 2010). Chromatin linkage of IN is mediated by specific domains located at opposite ends of p75, hence the tether metaphor. The first 198 amino acids of PSIP1/LEDGF contain a modular chromatin-binding ensemble containing charged regions 1, 2 and 3 (CR1, CR2 and CR3) that enhance the dominant chromatin-binding activities of the PWWP (Pro-Trp-Trp-Pro) domain and an AT-Hook pair (Figure 1.8). The C-terminal IBD (residues 347-429) interacts with the IN catalytic core domain (CCD) (Cherepanov *et al.* 2003). The functional unit comprises two asymmetric integrase dimers and two p75 molecules (Michel *et al.* 2009). IN subunit-subunit interactions are highly dynamic, with p75 strongly stabilising interactions and promoting the IN tetramer, the interface of which is important for enzymatic activities and high affinity p75 binding (McKee *et al.* 2008; Michel *et al.* 2009).

Suppression of p75 alters the genomic distribution of HIV-1 integration with three main effects: (1) the lentiviral bias for integrating into active genes was reduced; (2) AT-rich DNA targeting decreased; and, (3) p75-regulated genes were relatively disfavoured (Ciuffi *et al.* 2005). These integration pattern results have been confirmed and extended in two subsequent studies (Marshall *et al.* 2007; Shun *et al.* 2007). Recent profiling of the p75 chromatin interaction in the region analysed by the ENCODE project reveals associations with active chromatin markers, as expected. However, some associations did not correlate with preferred HIV-1 integration sites

indicating that not all p75 complexes on the chromosome are amenable to HIV-1 integration (De Rijck *et al.* 2010). p75 also associates with HIV-1 Rev, which is accompanied by loss of association with IN (Levin *et al.* 2010a; Levin *et al.* 2010b), which may serve to regulate proviral DNA integration and prevent super-infection.

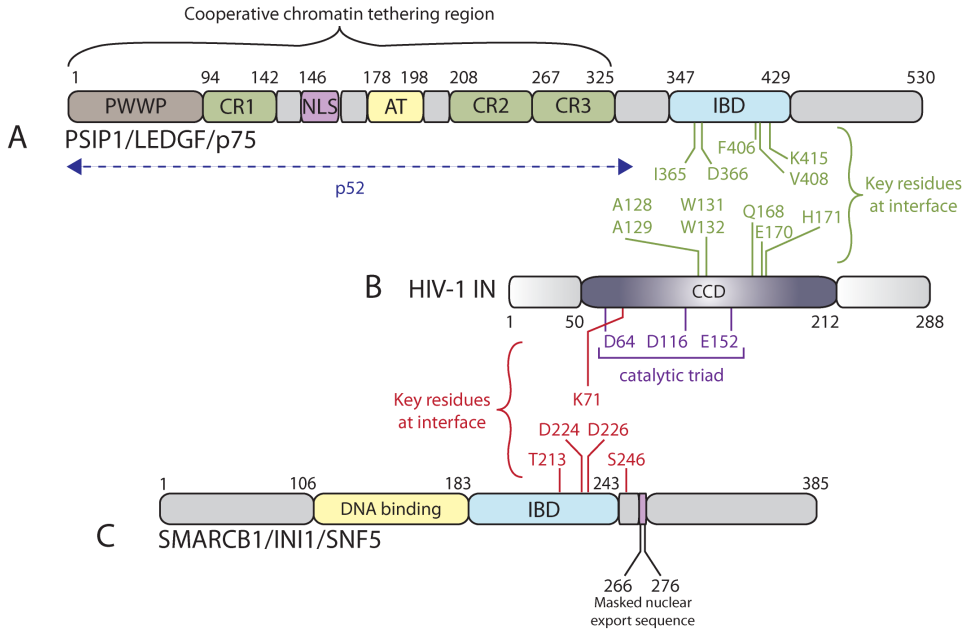


Figure 1.8: Domain structure of PSIP1/LEDGF/p75 and SMARCB1/INI1/SNF5 and interactions with HIV-1 IN. Amino acids at the interfaces between IN and p75 and SMARCB1 are indicated in green and red, respectively. **A.** The N-terminal domain ensemble participating in chromatin binding is primarily the PWWP domain (brown) and the AT Hook domain (yellow), with CR1-3 (green) being relatively charged regions that influence chromatin binding to a lesser extent. In cells, p75 binds tetrameric IN via its IBD (residues 347-429, blue). The nuclear localisation signal is depicted in pink. Alternative splicing can generate a C-termini variant, p52. **B.** IN is generated from $Pr160^{GagPol}$ by viral protease cleavage. IN associates with both p75 and SMARCB1 through its CCD (residues 51-212, purple). The catalytically essential DDE motif of the CCD is depicted in purple. **C.** For SMARCB1, yellow denotes the DNA-binding region, blue the IBD and pink the masked nuclear export sequence.

Several approaches have shown that for HIV-1, p75 is functionally super-abundant and must be stripped from the DNase- and salt-extractable chromatin fraction to produce a substantial block to HIV-1 integration (Llano *et al.* 2006a; Llano *et al.* 2006b; Hombrouck *et al.* 2007; Shun *et al.* 2007). Nevertheless, mouse knockouts reveal that p75 is not essential for cell or organism survival (Sutherland *et al.* 2006), suggesting that p75 is a good anti-HIV therapeutic target. Indeed, a number of approaches are being pursued for the inhibition of the p75-IN interaction, including small-molecule inhibitors (Christ *et al.* 2010; De Luca *et al.* 2010) and inhibitory peptides based on the IBD (Hayouka *et al.* 2010a; Hayouka *et al.* 2010b).

There are concerns over p75-IN targeting inducing a shift toward integration sites more typical of γ-retroviruses, which may increase cancer incidence. In addition, a shift in integration may have consequences for the frequency of post-integration latency, as the integration site of the HIV-1 promoter primarily modulates transcriptional burst size, rather than frequency (Skupsky

et al. 2010). Regardless, shRNA-mediated p75 knockdown in T cell-derived SupT1 cells was shown to inhibit HIV-1 replication for over 30 days (Llano *et al.* 2006a). Consequently, p75-targeting was performed in this study as a positive control for HDF suppression inhibiting HIV-1 replication.

1.6 Thesis objectives

Sustained inhibition of HIV-1 replication may be achieved by targeting cellular factors required for productive completion of the viral life cycle. This is likely to present a higher genetic barrier to viral escape than targeting the virus directly, although care must be taken to avoid toxic side effects. A number of putative HDFs require validation, particularly in cells that are the natural substrates of HIV-1, to provide a clear indication of whether these cellular factors warrant further investigation as anti-HIV therapeutic targets. Strategies that do not require daily administration will be more practical in resource-poor settings and may also reduce therapy-associated toxicities. The application of TGS, which has the potential to induce long-term gene silencing, is an attractive therapeutic option. However, this approach has rarely been explored at HDF gene promoters and the ability of TGS at HDFs to inhibit HIV-1 replication has not been characterised. In addition, it is currently unknown whether all RNA Pol II-transcribed genes are susceptible to small RNA-induced TGS, which hampers target selection. Therefore, the overall aims of this thesis were:

1. To characterise whether the host factors HTATSF1, DDX3X and SMARCB1 are HDFs and potential anti-HIV therapeutic targets.
    a. A panel of shRNAs, designed to target the four cellular factors HTATSF1, DDX3X, SMARCB1 and PSIP1, were tested for their ability to mediate target suppression *in vitro*.
    b. The ability of these shRNAs to inhibit HIV-1 replication was assessed in reporter cell lines, to indicate whether the targeted host factors function as HDFs.
    c. The cytotoxic effects accompanying shRNA-mediated target suppression were examined.
    d. To further validate the therapeutic potential of targeting these cellular factors, T cell-derived lines with stable shRNA expression were generated. These were tested for their ability to support HIV-1 replication.
    e. Proliferation rates of the T cell-derived lines with suppressed host factor expression were determined.
2. To characterise whether TGS could be induced at various HDF promoters and describe promoter features associated with TGS susceptibility.

a. A panel of siRNAs, designed to target HDF promoters, were tested for their ability to mediate target suppression *in vitro*.
b. The mechanism of siRNA-mediated target silencing was characterised, to discriminate between PTGS and TGS.
c. The ability of TGS at HDF promoters to inhibit HIV-1 infectious particle production was determined in reporter cell lines.
d. A genome-wide analysis of promoter features was performed to identify those associated with TGS susceptibility.

Achieving these objectives contributes to the field of HIV therapy development by: (1) validating HDFs and confirming their potential as drug targets; (2) demonstrating whether TGS of HDFs is sufficiently potent to inhibit viral replication; and, (3) facilitating the ability of future researchers to select TGS target genes and, therefore, novel HDF targets. Together, completion of these objectives contributes to the development of improved HIV therapeutics with reduced toxicity, administration frequency and viral resistance.

CHAPTER 2

Inhibition of HIV-1 replication with host factor-targeting shRNAs in TZM-bl reporter cells

2.1 Introduction

Viruses, by definition, are obligate parasites. Encoding only 15 proteins (Frankel and Young 1998), HIV exploits cellular functions of its host to facilitate replication. Given the high mutability of HIV, the targeting of cellular factors required for HIV replication is an attractive anti-HIV therapy since the genetic barrier for viral escape should be significantly higher than when targeting the virus itself (see 1.2.5).

RNAi is a useful method for gene silencing in which dsRNAs induce sequence-specific degradation of homologous mRNAs in plants and animals (Fire *et al.* 1998; Chiu and Rana 2002; McManus and Sharp 2002) (see 1.3.1). Indeed, RNAi has become the method of choice for genome-wide screens for the identification of factors involved in processes, including viral replication (see 1.3.5). Four genome-wide RNAi screens have been performed to identify host proteins involved in HIV-1 replication (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b), so-called HIV-dependency factors (HDFs) (Brass *et al.* 2008) (see 1.3.4). Little overlap was observed between these studies, which is unsurprising given the different conditions used and the variation between identical screens (Bushman et al. 2009). An elegant meta-analysis of the first three screens demonstrated that the analysis of enriched networks was more robust than identification of individual proteins involved in HIV-1 replication (Bushman *et al.* 2009). Thus, the individual proteins identified in these screens require further experimental validation of their role in HIV-1 replication.

HTATSF1 was a hit in the first genome-wide RNAi screen for HDFs (Brass *et al.* 2008) and was previously implicated in HIV pathogenesis as an elongation factor (Zhou and Sharp 1996; Li and Green 1998; Kim *et al.* 1999; Parada and Roeder 1999; Zhou *et al.* 2004a) (see 1.5.1). Increasing evidence supports a general elongation factor role that is not HIV-specific (Li and Green 1998; Parada and Roeder 1999; Chen *et al.* 2009). DDX3X was a hit in two screens (Brass *et al.* 2008; Zhou *et al.* 2008). An RNA helicase, DDX3X is involved in the Rev-CRM1 nuclear export pathway of singly spliced and unspliced HIV transcripts (Yedavalli *et al.* 2004) but has also been implicated in regulation of cell cycle progression (Lai *et al.* 2010) and an antiviral innate immune response (Oshiumi *et al.* 2010a) (see 1.5.2). At the outset of this work, no study had validated the

screen results to show that RNAi-mediated depletion of HTATSF1 and DDX3X in cell culture inhibits HIV-1 replication.

SMARCB1 is one of four core components of the mammalian SWI/SNF complex that is involved in ATP-dependent chromatin remodelling during transcriptional activation (Wang *et al.* 1996) (see 1.5.3). SMARCB1 was the first host protein shown to directly interact with HIV-1 IN (Kalpana *et al.* 1994; Rain *et al.* 2009), yet there is considerable controversy regarding the role of SMARCB1 in HIV-1 replication. The majority of evidence points to SMARCB1 as an HDF, with roles both early and late in the HIV-1 replication cycle. However, the lack of identification of SMARCB1 in several large-scale RNAi screens to detect HDFs (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b), and the numerous conflicting results, necessitate further investigation into the requirement for this protein in viral replication.

The 75 kDa isoform of PSIP1, p75, has a well-characterised role in HIV-1 integration (Llano *et al.* 2006a; Vandekerckhove *et al.* 2006) (see 1.5.4). PSIP1 is believed to function as a general adaptor between chromatin and proteins that must be brought into the proximity of chromatin to function, such as those involved in transcription (Bartholomeeusen *et al.* 2007; Bartholomeeusen *et al.* 2009) and integration. p75 mouse knockouts reveal that p75 is not essential for cell or organism survival (Sutherland *et al.* 2006). In addition, stable knockdown is well tolerated in human cell lines, producing no discernible differences in growth rates, morphology or phenotypes other than lentivirus susceptibility (Llano *et al.* 2004; Ciuffi *et al.* 2005; Llano *et al.* 2006a; Vandekerckhove *et al.* 2006; Marshall *et al.* 2007). This implicates PSIP1 as a good antiviral therapeutic target and, indeed, shRNA-targeting of PSIP1 was shown to effectively inhibit HIV-1 replication in SupT1 cells (Llano *et al.* 2006a).

In this chapter, RNAi was exploited to determine whether three host proteins, HTATSF1, DDX3X and SMARCB1, have potential as novel anti-HIV therapeutic targets. PSIP1 was also included in this study as a positive control for host factor suppression-mediated inhibition of HIV-1 replication. To this end, shRNAs targeting these host factors were expressed from U6-driven cassettes. shRNA-mediated suppression was used to examine the requirement of targeted cellular factors for HIV-1 replication in a reporter cell line. In addition, whether target suppression was accompanied by cytotoxicity was assessed.

2.2 Materials and methods

2.2.1 Design of cellular factor-targeting shRNAs

shRNAs were designed to target *htatsf1* (Entrez gene ID 27336) transcript NM_014500.3, *ddx3x* (Entrez gene ID 1654) transcript NM_001356.3, *psip1* (Entrez gene ID 11168) transcript NM_033222.2 and *smarcb1* (Entrez gene ID 6598) transcript NM_003073.3 (Table 2.1).

Table 2.1: Regions within host factor transcripts targeted by shRNAs.

Target	shRNA	Target region	Target length (nt)
HTATSF1	shH1	378-398	21
	shH2	415-435	21
	shH3	1416-1438	23
DDX3X	shD1	1940-1960	21
	shD2	1154-1175	22
	shD3	1285-1305	21
PSIP1	shP1	1652-1671	20
	shP2	860-880	21
SMARCB1	shS1	342-361	20
	shS2	1052-1072	21

shP2 targets a region common to both the p52 (NM_021144.3) and p75 (NM_033222.2) isoforms whilst shP1 is p75-specific. The shP1 guide sequence has been shown previously to inhibit p75 expression (Llano *et al.* 2006a). Similarly, shS1 was adapted from a siRNA sequence shown to suppress SMARCB1 expression (Maroun *et al.* 2006). Other target sites were selected from sequences suggested by an algorithm developed by The RNAi Consortium, The Broad Institute, (www.broad.mit.edu/genome_bio/trc/rnai.html) that predicts si/shRNA-susceptible sites in all human and mouse RefSeq transcripts based on predicted knockdown efficiency and specificity. All shRNAs were designed to contain a loop sequence derived from miR-31 (Table 2.2). In some shRNAs, G:U mismatches were incorporated at the 3' end of the sense strand to decrease thermodynamic stability and favour antisense strand selection as the guide (Table 2.2) (Schwarz *et al.* 2003).

2.2.2 Cloning expressed shRNAs

Generation of Pol III U6 shRNA expression cassettes has been described (Castanotto *et al.* 2002) and a similar two-step PCR approach was used here (Figure 2.1). The first amplification was carried out with a universal U6 forward primer and the first shRNA reverse primer (R1) that contained sequences complementary to the 3' end of the U6 promoter and encoding the sense and loop regions of the short hairpin (Table 2.2). The DNA plasmid pTZU6+1 (Bertrand *et al.* 1997), which contains the human U6 small nuclear RNA promoter (GenBank® accession number X59362), was used as a template. PCR was performed on a Mastercycler® (BioRad, CA, USA) using Expand High Fidelity^{PLUS} PCR System reagents (Roche, Mannheim, Germany). The PCR reaction included 10 ng template, Expand HiFi^{PLUS} reaction buffer (proprietary information), 1.5 mM $MgCl_2$, 200 μM dNTP mix, 2.5 U of Expand HiFi^{PLUS} Taq polymerase and 10 pmol of each primer in a volume of 50 μl. Thermocycling conditions were 95 °C for 10 minutes (min), 30 cycles of denaturation at 95 °C, annealing at 55 °C and extension at 72 °C, each for 30 seconds (sec), followed by a final extension at 72 °C for 10 min. Amplification specificity was confirmed by resolution of PCR products on a 2% agarose gel (see A1.2.3).

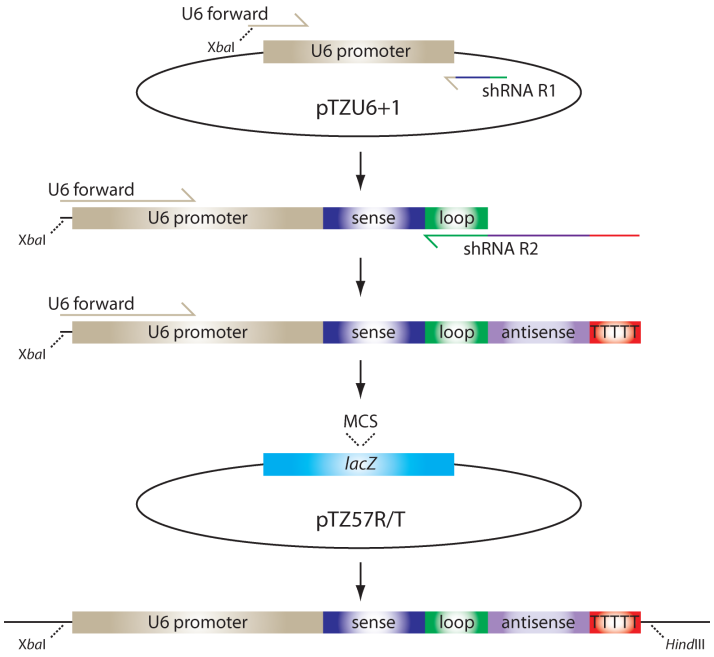


Figure 2.1: Schematic demonstrating the two-step PCR approach used to generate shRNA expression cassettes. The first round of PCR generates an amplicon encoding the sense strand as well as the loop of the shRNA. The second round amplicon encodes the sense, loop and antisense sequences and is cloned directly into the TA cloning vector pTZ57R/T.

Step one PCR products were diluted 500-fold and used as template (~10 pg) in a second PCR reaction. Primers used in the second step were the universal U6 forward primer and a second shRNA reverse primer (R2) that contained sequences complementary to the loop and antisense strand of the shRNA expression cassette, as well as a Pol III termination signal (Table 2.2). Other

reaction conditions, including thermocycling parameters, were the same as those used in the first PCR step.

PCR products of the appropriate molecular weight were ligated directly into pTZ57R/T (Fermentas, WI, USA), a TA cloning vector. The ligation reaction contained a 3:1 molar ratio of insert to vector (180 fmol insert to 60 fmol vector) and 20 U T4 DNA ligase (New England Biolabs®, Inc., MA, USA) in a volume of 20 μl, which was incubated at room temperature overnight. Ten microlitres of the ligation reaction mixture was used to transform 100 μl of chemically competent *Escherichia coli* DH5α (Invitrogen, CA, USA) (see A1.1.1, A1.1.2 and A1.1.3). Transformed cells were plated on Luria Bertani (LB) agar plates containing ampicillin (Gibco, BRL, UK), 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal) (Sigma, MO, USA) and isopropyl-β-D-1-thiogalactopyranoside (IPTG) (Roche, Mannheim, Germany) and incubated at 37 °C overnight (see A1.1.1).

Table 2.2: Oligonucleotide primers used to generate shRNAs targeting cellular factors.

Primer	Sequence (5'-3')
U6 forward	CTAACTAGTGGCGCGCCAAGGTCGGGCAGGAAGAGGG
shH1 R1	TTAT*TGGGTCAGG*ATAATTGGCCTGATATGTAGCGGTGTTTCGTCCTTTCCACAA
shH1 R2	AAAAAA**GCTACATATCAGGCCAATTAT***TGGGTCAGG*ATAA
shH2 R1	GCAA*TGGGTCAGG*TTÂCGGTAGAACTAGATGCGCGGTGTTTCGTCCTTTCCACAA
shH2 R2	AAAAAA**GCGCATCTAGTTCTACCGCAA***TGGGTCAGG*TTÂC
shH3 R1	CGTT*TGGGTCAGG*AACÂCCATTCCAGTTGCAGTGGTGTTTCGTCCTTTCCACAA
shH3 R2	AAAAAA**CTGCAACTGGAATGGCGTT***TGGGTCAGG*AACÂ
shD1 R1	GATA*TGGGTCAGG*TATCTTGTTCGACTATTCTACGGTGTTTCGTCCTTTCCACAA
shD1 R2	AAAAAA**GTAGAATAGTCGAACAAGATA***TGGGTCAGG*TATC
shD2 R1	ACTC*TGGGTCAGG*AAGTTCCTGTTCCAAGCGTTCGGTGTTTCGTCCTTCCACAA
shD2 R2	AAAAAA**GACGGAGTGATTACGATGGCAT***TGGGTCAGG*ATGC
shD3 R1	ACTC*TGGGTCAGG*ÂAGTTCCTGTTCCAAGCGTTCGGTGTTTCGTCCTTCCACAA
shD3 R2	AAAAAA**GAACGCTTGGAACAGGAACTC***TGGGTCAGG*ÂAGT
shP1 R1	CGAA*TGGGTCAGG*TTCÂCTTCCTCATGCTGTCTGGTGTTTCGTCCTTTCCACAA
shP1 R2	AAAAAA**GACAGCATGAGGAAGCGAA***TGGGTCAGG*TTC
shP2 R1	CATT*TGGGTCAGG*AATCTTGACTTCTGTAGCTGCGGTGTTTCGTCCTTTCCACAA
shP2 R2	AAAAAA**GCAGCTACAGAAGTCAAGATT***TGGGTCAGG*AATC
shS1 R1	GGAA*TGGGTCAGG*TTCCAGAGTGAGGGGTATCTCGGTGTTTCGTCCTTTCCACAA
shS1 R2	AAAAAA**GAGATACCCCTCACTCTGGAA***TGGGTCAGG*TCC
shS2 R1	AACT*TGGGTCAGG*AGTTCTCCTTCTCTGACATGTGGTGTTTCGTCCTTTCCACAA
shS2 R2	AAAAAA**CATGTCAGAGAAGGAGAACT***TGGGTCAGG*AGTT

Sequences complementary to the 3' end of the U6 promoter are underlined and to a Pol III termination signal are shaded grey. Those complementary to the antisense guide strand are in bold and to the loop sequence in italics. Introduced G:U mismatches in the sense strand are indicated by a circumflex.

White colonies were screened for the presence of shRNA expression cassettes by culturing in 4 ml LB broth supplemented with ampicillin overnight at 37 °C with shaking (see A1.1.1). Plasmids were prepared using the High Pure Plasmid Isolation Kit (Roche, Mannheim, Germany) (see A1.2.1) and 5 μg digested with 5 U *Xba*I and 5 U *Hind*III (Fermentas, WI, USA) for I hour at 37 °C in a 20 μl reaction containing Tango Buffer (Fermentas, WI, USA). Digestion reactions were resolved on a 2% agarose gel (see A1.2.3). Clones that exhibited the desired digestion products were sequenced with M13 forward (5' GTA AAA CGA CGG CCA G 3') and reverse (5' CAG GAA ACA GCT ATG AC 3') primers using automated cycle sequencing (Inqaba Biotec, Pretoria, South Africa). Those containing correct shRNA expression cassettes were cultured in 300 ml ampicillin-supplemented LB overnight at 37 °C (see A1.1.1) with shaking prior to bulk plasmid DNA preparation using the QIAGEN Plasmid Maxi Kit (QIAGEN, CA, USA) (see A1.2.2).

Several previously developed constructs were used as controls in experiments: a mock pTZU6+1 construct with no shRNA sequence (Bertrand *et al.* 1997); a shRNA negative control, shHBVx, which targets an irrelevant site in hepatitis B virus (HBV) X protein (Carmona *et al.* 2006); and, two positive controls, shU5 and shTAT, which are named after the location of their target sequences within HIV-1 transcripts and were initially developed based on subtype B molecular clone HXB2, GenBank® accession number K03455 (Barichievy *et al.* 2007).

2.2.3 Cloning luciferase reporter vectors

To assess knockdown by shRNAs, four reporter constructs were generated, each containing the multiple shRNA targets for a cellular factor (Figure 2.2). Target sequences were generated from complementary oligonucleotides (target F1 and target R1) that also contained an *Xho*I site at the 5' end and a *Not*I site at the 3' end to facilitate cloning. In addition, an *Eco*RV site was incorporated to enable screening for positive clones. HTATSF1 and DDX3X target sequences, where three shRNA targets were needed (Figure 2.2), were generated from four oligonucleotides (target F1, target F2, target R1 and target R2; Table 2.3).

To enhance ligation, 200 pmol of target F2 and R1 oligonucleotides were treated with 2 U polynucleotide kinase (Promega, WI, USA) in a 20 μl volume that also contained polynucleotide kinase buffer (Promega, WI, USA) and 1 mM ATP at 37 °C for 1 hr. One hundred picomoles of complementary oligonucleotides were annealed in a volume of 20 μl by heating at 80 °C for 5 min and then allowing the samples to slowly cool to room temperature. The resulting dsDNA inserts

were used in either two- (PSIP1 and SMARCB1 reporters) or three-way (HTATSF1 and DDX3X reporters) ligations with psiCheck™-2 (Promega, WI, USA). This plasmid contains a multiple cloning site within the 3' UTR of a SV40 promoter-driven *Renilla* luciferase, as well as a HSV TK-driven firefly luciferase cassette.

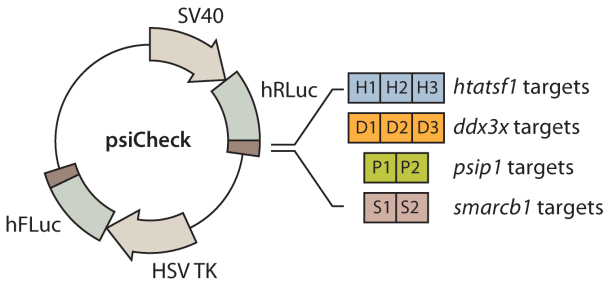


Figure 2.2: Schematic of generated reporter constructs containing cognate shRNA target sequences for each cellular factor. Complementary oligonucleotides were annealed and cloned into psiCheck™-2 such that each cellular factor shRNA target sequence is present sequentially within the 3' UTR of the *Renilla* luciferase gene. *Renilla* luciferase expression reflects, therefore, the ability of a shRNA to inhibit its cognate target. The construct also contains an independently-expressed firefly luciferase gene, whose expression is unaffected by the presence of shRNAs.

Table 2.3: Oligonucleotides used to generate shRNA target sequences in the reporter constructs.

Target oligonucleotide	Sequence (5'-3')
HTATSF1 F1	TCGA**GATATC**GCTACATATCAGGCCAATTATGCGCA
HTATSF1 F2	TCTAGTTCTACCGCAAACTGCAACTGGAATGGCGTT*GC*
HTATSF1 R1	CTAGATGCGCATAATTGGCCTGATATGTAGC**GATATC**
HTATSF1 R2	*GGCCGC*AACGCCATTCCAGTTGCAGTTTGCGGTAGAA
DDX3X F1	TCGA**GATATC**GTAGAATAGTCGAACAAGATAGACGG
DDX3X F2	AGTGATTACGATGGCATGAACGCTTGGAACAGGAACTC*GC*
DDX3X R1	TCACTCCGTCTATCTTGTTCGACTATTCTAC**GATATC**
DDX3X R2	*GGCCGC*GAGTTCCTGTTCCAAGCGTTCATGCCATCGTAA
PSIP1 F1	TCGA**GATATC**AGACAGCATGAGGAAGCGAAGCAGCTACAGAAGTCAAGATT*GC*
PSIP1 R1	GGCCGCAATCTTGACTTCTGTAGCTGCTTCGCTTCCTCATGCTGTCT**GATATC**
SMARCB1 F1	TCGA**GATATC**GAGATACCCCTCACTCTGGAAACATGTCAGAGAAGGAGAACTGC
SMARCB1 R1	*GGCCGC*AGTTCTCCTTCTCTGACATGTTTCCAGAGTGAGGGGTATCTC**GATATC**

*Xho*I target sites are underlined, *Eco*RV target sites in bold and *Not*I target sites in italics.

Prior to ligation, 9 μg of the plasmid was digested in three separate 40 μl reactions, each containing 15 U *Xho*I and 15 U *Not*I and Orange buffer (Fermentas, WI, USA), at 37 °C for 1 hr. Digestion products were resolved on a 1% agarose gel (see A1.2.3) and the digested backbone excised and purified using the MinElute™ Gel Extraction Kit (QIAGEN, CA, USA) (see A1.2.4). Three micrograms of the purified backbone was treated with 5 U Antarctic phosphatase (New England Biolabs®, Inc., MA, USA) in a 10 μl reaction containing Antarctic phosphatase buffer (New England Biolabs®, Inc., MA, USA) at 37 °C for 1 hr, followed by inactivation of the enzyme at 65 °C for 5

min. Sixty femtomoles of the dephosphorylated backbone was used in ligation reactions with 180 fmol of insert(s) at 16 °C overnight. Transformation, colony screening and bulk vector preparation were as described in 2.2.2, with the exception that colonies were screened by *Eco*RV (Fermentas, WI, USA) digestion.

2.2.4 Mammalian cell culture

Cells of the human embryonic kidney (HEK) cell line, HEK293, were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (BioWhittaker, MD, USA) supplemented with 10% heat-inactivated fetal calf serum (FCS) (Delta Bioproducts, Johannesburg, SA) at 37 °C and 5% CO_2 in a Forma Series II 3110 water-jacketed incubator (Thermo Fisher Scientific, Inc., MA, USA). Cervical cancer Henrietta Lacks (HeLa) cells were maintained as HEK293 cells. HeLa-derived TZM-bl cells (NIH AIDS Research and Reference Reagent Program, DC, USA), which express the HIV receptor CD4 and coreceptor CCR5 and contain a luciferase reporter driven by a Tat-inducible LTR promoter derived from pSG3.1 (GenBank® accession number L02317) (Platt *et al.* 1998; Derdeyn *et al.* 2000; Wei *et al.* 2002), were also maintained as HEK293 cells. All cells were propagated in 75 cm^2 flasks and seeded into plates of different well sizes for experiments (Nunc™, Thermo Fisher Scientific, Inc., MA, USA). To determine the correct number of cells for seeding, 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).

2.2.5 DNA plasmid transfections

Cell transfections were carried out using 1 μl Lipofectamine™ 2000 (Invitrogen, CA, USA) per 1 μg DNA. DNA was mixed with a volume of OptiMEM (Gibco, BRL, UK) equal to $1/10^{th}$ of the cell culture medium (eg 50 μl OptiMEM per 0.5 ml culture in a 24-well plate). The same volume of OptiMEM was simultaneously mixed with the appropriate volume of Lipofectamine™ 2000. Both solutions were incubated at room temperature for 10 min prior to mixing and a further room temperature incubation of 20 min. A volume equal to $1/5^{th}$ of the cell culture medium of DNA-Lipofectamine™ 2000-OptiMEM mixture was slowly added to cells (eg 100 μl mixture per 0.5 ml culture in a 24-well plate). Media was changed 4-5 hours post-transfection to minimise transfection reagent-associated toxicity.

2.2.6 Dual luciferase reporter assay

To assess the inhibitory potential of the shRNAs, HEK293 and HeLa cells were seeded at 1.2 x 10^5 and 5.0 x 10^4 cells per well, respectively, in a 24-well plate 24 hours pre-transfection. Each well was transfected with 1 μg of pTZU6+1 shRNA expression construct, 100 ng of the appropriate psiCheck™-2 target reporter construct and 10 ng pCI-eGFP, which expresses enhanced green fluorescent protein (eGFP) from a CMV promoter (Passman *et al.* 2000) to allow estimation of transfection efficiency. Transfections were performed in triplicate as described in 2.2.5.

Forty-eight hours post-transfection, reporter assays were performed using the Dual Luciferase® Reporter Assay System (Promega, WI, USA). Briefly, medium was removed and cells lysed with 100 μl Passive Lysis Buffer for 20 min. Ten microlitres of cell lysate was transferred to a Costar® 96-well flat bottom assay plate (Corning, Inc., BY, USA). Firefly luciferase activity was measured following the addition of 50 μl LARII (Figure 2.3). The subsequent addition of Stop & Glo® solution simultaneously inhibited firefly luciferase and permitted the quantification of *Renilla* luciferase activity (Figure 2.3). All luciferase activities were quantified by a Veritas™ Microplate Luminometer (Turner Biosystems, CA, USA). *Renilla*:firefly luciferase activity ratios were normalised to the average expression ratio of mock (ie pTZU6+1)-transfected samples.

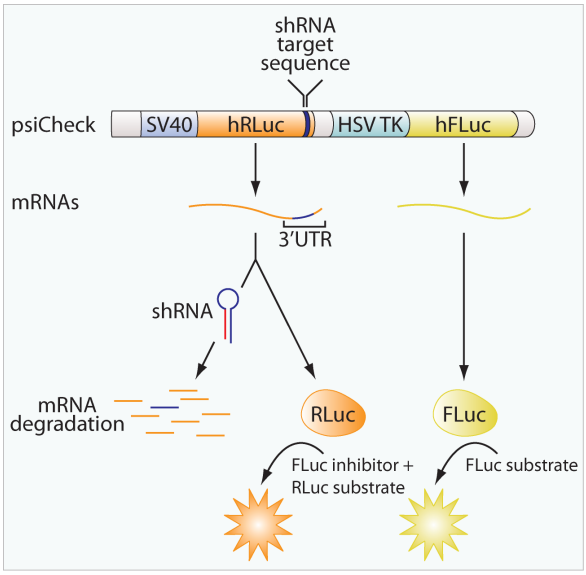


Figure 2.3: Schematic representation of the dual luciferase reporter assay. In the absence of shRNA expression, processing and/or guide strand association with its target, there is *Renilla* luciferase expression. When the shRNA successfully mediates inhibition of its target, *Renilla* luciferase expression is decreased whilst firefly luciferase expression is unaffected. This results in a reduced *Renilla*: firefly luciferase activity ratio.

2.2.7 Quantitative reverse transcription polymerase chain reaction (qRT-PCR) measurement of mRNA concentrations

To assess the suppression of endogenous mRNA targets mediated by shRNAs, HEK293 and TZM-bl cells were seeded at 1.2 x 10^5 and 5.0 x 10^4 cells per well, respectively, in a 24-well plate 24 hours pre-transfection. Cells of each well were transfected with 1 μg of pTZU6+1 shRNA expression construct and 10 ng pCI-eGFP in triplicate, as described in 2.2.5. Forty-eight hours post-transfection, medium was removed and total RNA extracted using TriReagent™ (Sigma, MO, USA) (see A1.3.1). Total RNA (100-200 ng) was subjected to deoxyribonuclease (DNase) treatment and random-primed reverse transcription (RT) using the QuantiTect Reverse Transcription Kit (QIAGEN, CA, USA) (see A1.3.2).

Table 2.4: Primers used in qPCR to determine cellular factor cDNA levels.

Primer	Sequence (5'-3')	Amplicon length (bp)
HTATSF1 forward	AGTGGGACCTGGACAAAAAGG	171
HTATSF1 reverse	GTTCCGGGGCTTTTTCTTGTG	
DDX3X forward	CGTGGAAGTGGATCAAGGGGAA	144
DDX3X reverse	TTGGTGGGAGTGGTTTTGACCA	
PSIP1 (p75 specific) forward	GCTGAACAAAGACAGCATGAGGA	146
PSIP1 (p75 specific) reverse	ATTGCTCTCCCCGTTATGTTGTG	
SMARCB1 forward	ATCAGACAGCAGATCGAGTCCTA	144
SMARCB1 reverse	TCAGGGCAAACTTCTCTGGTGA	
ACTB forward	AGGTCATCACCATTGGCAATGAG	118
ACTB reverse	TCTTTGCGGATGTCCACGTCA	

One microlitre of complementary DNA (cDNA) was analysed for target mRNA expression relative to beta-actin (*actb*) mRNA (Entrez Gene ID 60, transcript NM_01101.2) expression using the Lightcycler® FastStart DNA Master^{PLUS} SYBR Green I Kit (Roche, Mannheim, Germany) in a 20 μl reaction (see A1.3.4). Primers were designed to produce specific 100-200 bp amplicons spanning splice junctions (Table 2.4). The qPCR was performed in a Carousel-Based Lightcycler® V.2 System (Roche, Mannheim, Germany) with the following parameters: denaturation at 95 °C for 10 min, 50 cycles of denaturation at 95 °C, annealing at 60 °C and extension at 72 °C, each for 10 sec. Amplification cycles were followed by melting curve analysis to verify the specificity of the PCR products. No RT controls were included for each sample and no cDNA controls for each primer set. Target:*actb* mRNA concentration ratios were determined by the Lightcycler® V.2 software (Roche, Mannheim, Germany) and normalised to the average expression ratio of mock-transfected samples.

2.2.8 Challenge with HIV-1 molecular clone pNL4-3.Luc.R$^-$.E$^-$

To assess the ability of shRNAs targeting cellular factors to inhibit HIV-1 replication, HEK293 cells, seeded at 1.2 x 10^5 cells per well in a 24-well plate, were transfected with 900 ng shRNA expression construct, 10 ng of pRSV-RLuc (gift from J. J. Rossi), which encodes a CMV-driven *Renilla* luciferase gene and 90 ng pNL4-3.Luc.R$^-$.E$^-$, a HIV-1 subtype B molecular clone (NIH AIDS Research & Reference Reagent Program, MD, USA) (Connor *et al.* 1995; He *et al.* 1995). This clone consists of pNL4-3 (GenBank® accession number M19921) with a firefly luciferase reporter gene fused to the 5' end of the *nef* coding region using *Not*I and *Xho*I sites. Two frame shifts render this clone Vpr- and Env- and, therefore, competent for a single round of replication only.

Forty-eight hours post-transfection, cells were assayed for dual luciferase activity using the Dual Luciferase® Reporter Assay System (Promega, WI, USA), as described in 2.2.6. Firefly luciferase activities were normalised to *Renilla* luciferase activities and then normalised to the average firefly:*Renilla* luciferase expression ratio of the mock-transfected samples.

2.2.9 Detection of processed shRNA guide strands using polyacrylamide gel electrophoresis (PAGE) and Northern blot analysis

TZM-bl cells were seeded at 2 x 10^6 cells per 60 cm^2 Costar® dish (Corning, Inc., NY, USA) and each dish transfected 24 hours later with 20 μg shRNA expression constructs using the calcium chloride method (see A1.2.5). Forty-eight hours post-transfection, total RNA was isolated using TriReagent™ (Sigma, MO, USA) (see A1.3.1). Thirty micrograms of RNA and a radiolabelled RNA molecular weight Decade™ Marker (Ambion, TX, USA) (see A1.3.3) were combined with an equal volume of RNA loading dye (Ambion, TX, USA) and denatured at 80 °C for 5 min. Samples were resolved on a 15% polyacrylamide gel with a 1:19 ratio of *bis*-acrylamide:acrylamide, 8 M urea and TBE buffer (890 mM Tris base, 890 mM boric acid and 32 mM EDTA, pH 8.0). The gel was stained with 10 mg/ml ethidium bromide and visualised on a UV transilluminator (Kodak Gel Logic 200 Imaging System) to check RNA quality and equal loading of samples.

RNA underwent transfer to Hybond-N+ positively charged nylon membrane (GE Healthcare, NJ, USA) using the Hybaid semi-dry blotter (Amersham Biosciences, CT, USA) at 3.3 mA/cm^2 (not exceeding 400 mA) for one hour. RNA was cross-linked to the membrane at 2000 x 100 μJ/cm^2 in a UV cross-linker (UVP, Inc., CA, USA) and the membrane then baked at 80 °C for 1 hour in a vacuum oven (Thermo Fisher Scientific, Inc., MA, USA).

DNA oligonucleotide probes complementary to the processed shRNA guide strands (Table 2.5) were labelled with γ-^{32}P-ATP (6000 ci/mmol) (Perkin Elmer, MA, USA) using T4 polynucleotide

kinase (Promega, WI, USA). The 20 μl reactions, composed of Reaction Buffer A (Promega, WI, USA) (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 probe, 20 pmol γ-^{32}P-ATP and 10 U of T4 polynucleotide kinase, were incubated for 20 min at 37 °C. Addition of 1 μl of 0.5 M EDTA, pH 8.0, and incubation at 75 °C for 10 min, terminated the labelling reaction. Labelled probes were purified by centrifugation at 200 x *g* for 2 min through Sephadex® G-25 columns (Sigma, MO, USA) and then added to membranes pre-hybridised with 10 ml Rapid Hyb solution (Amersham Biosciences, CT, USA) in a rotating oven (Thermo Fisher Scientific, Inc., MA, USA). Hybridisation was permitted overnight at 42 °C.

Table 2.5: Sequences of DNA oligonucleotide probes used in Northern blots.

Oligonucleotide probe	Sequence (5'-3')
Anti-shH1	GCTACATATCAGGCCAATTAT
Anti-shH2	GCGCATCTAGTTCTACCGCAA
Anti-shH3	ACTGCAACTGGAATGGCGTT
Anti-shP1	AGACAGCATGAGGAAGCGAA
Anti-shP2	GCAGCTACAGAAGTCAAGATT
Anti-shS1	GAGATACCCCTCACTCTGGAA
Anti-shS2	ACATGTCAGAGAAGGAGAACT
Anti-U6 snRNA	TAGTATATGTGCTGCCGAAGCGAGCA

Probes were labelled with γ-^{32}P-ATP prior to hybridisation with membrane-bound RNA.

Membranes were subject to a low stringency washes in 50 ml 5x SSC (750 mM NaCl, 65 mM disodium citrate, pH 7.0) and 0.1% sodium dodecylsulphate (SDS) (Sigma, MO, USA) for 20 min at room temperature. Two high stringency washes were then performed at 42 °C, each with 50 ml 1x SSC (150 mM NaCl, 13 mM disodium citrate, pH 7.0) and 0.1% SDS for 15 min. Following a 48 hour exposure to an imaging plate (Fujifilm, Tokyo, Japan), blots were viewed on a FLA-7000 phosphorimager (Fujifilm, Tokyo, Japan). Blots were stripped with 50 ml 1% SDS at 80 °C for 30 min prior to reprobing. A probe complementary to U6 small nuclear RNA was used to verify equal loading of RNA samples (Table 2.5).

2.2.10 HIV-1 reporter assay in TZM-bl cells

TZM-bl cells were seeded at 5 x 10^4 cells per well in a 24-well plate. Cells of each well were transfected 24 hours later with various amounts of pTZU6+1 shRNA expression construct and 10 ng pCI-eGFP in triplicate, as described in 2.2.5. When shRNA dose-effects were being

assessed, the stuffer DNA pUC19 (NEB, MA, USA) was also included in transfections such that the total amount of transfected DNA was constant for all samples at 1.01 μg.

Cells were infected with various R5-tropic HIV-1 at a $TCID_{50}$ (see A1.4.3) of 1000/ml 24 hours post-transfection (primary infection): subtype C envelope-pseudotyped subtype B molecular clone $pSG3^{\Delta env}$.ZM53M (a gift from M. Papathanasopoulos) (Wei et al. 2002; Derdeyn et al. 2004); subtype B molecular clone p81A-4 (NIH AIDS Research & Reference Reagent Program, MD, USA) (Chesebro et al. 1992; Walter et al. 2005) (see A1.4.2); and, subtype C FV5 (GenBank® accession number DQ382363), propagated from a primary isolate originally isolated in 1998 at Johannesburg General Hospital (a gift from M. Papathanasopoulos) (see A1.4.1). Infections were carried out in the presence of 15 μg/ml diethylaminoethyl dextran (DEAE-D) to enhance infection efficiencies.

Cells were washed three times with PBS (Gibco, BRL, UK) 24 hours post-infection. In some experiments, 200 μl of culture supernatant was removed 48 hours post-infection and stored at -80 °C for subsequent analysis of HIV-1 p24 content by enzyme-linked immunosorbent assay (ELISA) (Murex Biotech Ltd, Kent, UK) (see A1.4.4). Another 100 μl of culture supernatant was used to infect additional TZM-bl cells (secondary infection), seeded at 5 x 10^4 cells per well in a 24-well plate the preceding day, in the presence of 15 μg/ml DEAE-D.

Tat-induced luciferase activity was assessed 48 hours post-infection(s) in cell lysates using the Bright-Glo™ Luciferase Assay System (Promega,WI, USA). Briefly, medium was removed and cells lysed with 100 μl Glo Lysis Buffer for 5 min. Twenty-five microlitres of cell lysate was transferred to a Costar® 96-well flat bottom assay plate (Corning, Inc., BY, USA). Luciferase activity was measured in a Veritas™ Microplate Luminometer (Turner Biosystems, CA, USA) before (background reading) and after the addition of 25 μl Bright-Glo™ Assay Reagent. Background was subtracted from each sample.

2.2.11 Detection of cellular factors by SDS-PAGE and Western blot analysis

To assess the kinetics of shRNA-mediated target suppression, TZM-bl cells were seeded at 1.5 x 10^5 cells per well in a 6-well plate. Each well was transfected with 2 μg shRNA expression plasmid and 50 ng pCI-eGFP 24 hours later, as described in 2.2.5. Twenty-four, 48 and 96 hours post-transfection, media was removed and cells washed with PBS. Cells were harvested by trypsin (Gibco, BRL, UK) treatment and centrifugation at 130 x *g* for 4 min and lysed by addition of 500 μl Radio-Immunoprecipitation Assay (RIPA) buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 0.1% SDS,

0.5% sodium deoxycholate and 1% Triton X-100) containing 1:200 dilution of Calbiochem® Protease Inhibitor Cocktail Set III (Merck, Darmstadt, Germany).

Total protein was quantified using the Pierce® bicinchoninic acid (BCA) Protein Assay Kit (Thermo Fisher Scientific Inc, IL, USA) (see A1.5.1). The EasyWestern Protein Standard (GenScript Corp, NJ, USA), composed of IgG-binding proteins ranging from 22 to 120 kD in size, and 80 μg of cellular protein were resolved on a 12% polyacrylamide gel with a 1:19 ratio of *bis*-acrylamide:acrylamide, 0.375 mM Tris, pH 8.8, and 0.2% SDS. Protein was transferred to a polyvinylidene fluoride (PVDF) membrane (Millipore, MA, USA) at 35 V overnight in circulating transblot buffer (25 mM Tris base, 192 mM glycine and 20% methanol), pH 8.3. Following transfer, the membrane was rinsed in Tris buffered saline (TBS) (0.5 M Tris.HCl, 9% (w/v) NaCl), pH 7.5 and then blocked for 1 hour at room temperature with blocking solution (TBS with 5% skimmed milk powder and 0.1% (v/v) Tween 20). Primary antibodies were added to the blocking solution and incubated at room temperature with gentle shaking for a further hour. The dilutions of rabbit polyclonal primary antibodies were as follows: anti-HTATSF1 (a gift from M. Garcia-Blanco) at 1:100, anti-PSIP1 (GenWay Biotech, Inc., CA, USA) at 1:2,000, anti-SMARCB1 (GenWay Biotech, Inc., CA, USA) at 1:500 and anti-β-actin (GenWay Biotech, Inc., CA, USA) at 1:1,000. The latter was used to verify equal loading of samples. Membranes were washed three times with blocking solution for 10 min each before a 1 hour incubation with a horseradish peroxidase (HRP)-conjugated donkey anti-rabbit IgG secondary antibody (GenWay Biotech, Inc., CA, USA) at 1:25,000 in blocking solution. Membranes were washed twice with blocking solution and once with TBS-T (TBS with 0.1% (v/v) Tween 20), each for 10 min. One millilitre of SuperSignal® West Pico Chemiluminescent Substrate (Thermo Fisher Scientific, Inc., IL, USA), composed of enhancer solution and stable peroxide solution in a 1:1 ratio, was added to the membrane and incubated for 5 min. Images were acquired with a G-BOX (Syngene, MD, USA) using an exposure of up to 45 min. Levels of target protein is reported relative to levels of β-actin and normalised to the shHBVx negative control at each time point.

Given the difference in molecular weight of some of the targeted proteins, HTATSF1 (~86 kDa), PSIP1 (~75 kDa) and β-actin (~42 kDa) could be detected simultaneously. SMARCB1 (~44 kDa) and β-actin could not and, therefore, some membranes were stripped between immunoblots. This involved washing the membrane twice with stripping buffer (20 mM Tris.HCl, pH 7.5, 6 M guanidine hydrochloride, 0.2% (v/v) Nonidet P-40 and 0.1 M β-mercaptoethanol) for 5 min each, three times with TBS-T for 3 min each and once with TBS for 5 min, all at room temperature.

2.2.12 Detection of early apoptotic TZM-bl cells

To examine whether target suppression is cytotoxic, TZM-bl cells were seeded at 3 x 10^4 cells per well on CELLocate® microgrid coverslips (Eppendorf, Hamburg, Germany) in a 24-well plate. Each well was transfected with 500 ng shRNA expression plasmid, as described in 2.2.5, in duplicate. Seventy-two hours post-transfection, apoptosis was quantified using the TACS™ Annexin V-FITC Apoptosis Detection Kit (R&D Systems, Inc., MN, USA). Briefly, cells were washed with PBS and then incubated in the dark for 15 min with 100 μl of Annexin V Incubation Reagent, containing binding buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, 1.8 mM $CaCl_2$) and 0.5 μl Annexin V-FITC. Cells were then washed twice with 500 μl binding buffer before inversion of coverslips onto slides with Calbiochem® FluorSave™ Reagent (Merck, Darmstadt, Germany). Cells were treated with 500 nM TSA 16 hours pre-analysis as a positive control for apoptosis induction.

Transmitted light and fluorescence images were acquired for two fields of view per well on an Axiovert 100 M with image capture by AxioVision 2.0.5 software (Carl Zeiss Microimaging GmbH, Germany). Fluorescence was quantified using ImageJ 1.40g (public domain; developed by W. Rasband, National Institutes of Health, MD, USA).

2.2.13 GFP splicing assay

To examine whether suppression of HTATSF1 affects splicing, TZM-bl cells were seeded at 6 x 10^4 cells per well in a 12-well plate. Twenty-four hours later, each well was transfected with 1 μg of shRNA expression plasmid, targeting HTATSF1, or HBVx as a negative control. Cells were transfected a second time 48 hours later with either pEGFP-C1-NAD (a gift from M. Wood), which contains a CMV-driven eGFP ORF interrupted by intron 6 from NADH-coenzyme Q reductase (Entrez gene ID 4719), or a mutant construct (a gift from M. Wood), which resembles pEGFP-C1-NAD but for mutated splice sites. Transfections were carried out as described in 2.2.5 and in triplicate. After another 48 hour incubation, transmitted light and fluorescence images were acquired, as described in 2.2.12.

Cells were then harvested, lysed and subject to protein quantification by the BCA method, as described in 2.2.11, with the exception that 250 μl of RIPA buffer was used (see also A1.5.1). GFP fluorescence in 150 μl of cell lysate was quantified in a VersaFluor™ fluorometer (BioRad, CA, USA), with excitation at 379-401 nm and emission at 505-515 nm, and reported relative to the total amount of protein in the sample.

2.2.14 MTT assay

To examine whether the shRNAs were themselves cytotoxic, TZM-bl cells were seeded at 1 x 10^4 cells per well in a 96-well plate. Twenty-four hours later, wells were either transfected with 100 ng shRNA expression plasmid, as described in 2.2.5, or treated with 10, 100 or 500 nM TSA, all in triplicate. Forty-eight hours post-transfection/-TSA treatment, 0.1 mg of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was added to each well, which is reduced by active reductase enzymes in living cells, such as mitochondrial dehydrogenases. Following incubation at 37 °C for 1 hour, all media was removed and the resulting formazan precipitates resuspended in 200 μl dimethyl sulfoxide (DMSO). The absorbance at 570 nm, with a reference wavelength of 655 nm, was determined in a Model 680 microplate reader (BioRad, CA, USA).

2.2.15 Cell viability by fluorescence microscopy

Viability of cells expressing shRNAs was also assessed by fluorescence microscopy. TZM-bl cells were seeded at 5 x 10^4 cells per well in a 24-well plate 24 hours prior to transfection with 500 ng of shRNA expression plasmid and 100 ng pCI-eGFP per well, as described in 2.2.5. Forty-eight hours post-transfection, transmitted light and fluorescence images were acquired, as described in 2.2.12.

2.2.16 qRT-PCR measurement of interferon-β mRNA concentrations

Induction of interferon-β expression accompanies an immune response to double-stranded RNAs. To determine whether shRNAs induced an immune response, TZM-bl cells were seeded at 3 x 10^4 cells per well in a 24-well plate and each well transfected with 500 ng shRNA expression plasmid and 10 ng of pCI-eGFP, as described in 2.2.5. One microgram of polyinosinic: polycytidylic acid (poly(I:C)) (Sigma-Aldrich, MO, USA), which is a dsRNA construct known to induce an interferon response, was used as a positive control (Loseke et al. 2006; Reynolds et al. 2006). Forty-eight hours post-transfection, total RNA was extracted using TriReagent™ (see A1.2.1) and subject to DNase treatment, reverse transcription and qPCR, as described in 2.2.7 (see also A1.3.2 and A1.3.4). Primers used to amplify interferon-β mRNA were: forward 5' TCC AAA TTG CTC TCC TGT TGT GCT 3'; and, reverse 5' CCA CAG GAG CTT CTG ACA CTG AAA A 3'. Expression was determined relative to β-actin mRNA (Table 2.4), as described in 2.2.7.

2.2.17 Statistical analyses

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. Data was analysed using unpaired, two-tailed t-tests, one-way ANOVAs, followed by a Dunnett's multiple comparison post-test, and two-way ANOVAs, followed by Bonferroni post-tests, where appropriate, calculated using GraphPad Prism 4.0c (GraphPad Software, Inc., CA, USA).

2.3 Results

2.3.1 shRNAs suppress cellular targets in HEK293 cells

Ten shRNAs were designed to target cellular factors HTATSF1, DDX3X, PSIP1 and SMARCB1 (Figure 2.4A). The loop sequence incorporated into each shRNA was derived from miR-31. Nucleotide substitutions were sometimes made in the passenger strand to create G:U mismatches that enhanced thermodynamic asymmetry in the shRNAs and facilitated the desired guide strand selection (see 1.3.1 and 1.3.3). Sequences encoding shRNAs were cloned downstream of a RNA Pol III U6 promoter (Figure 2.4B) to promote high levels of shRNA expression.

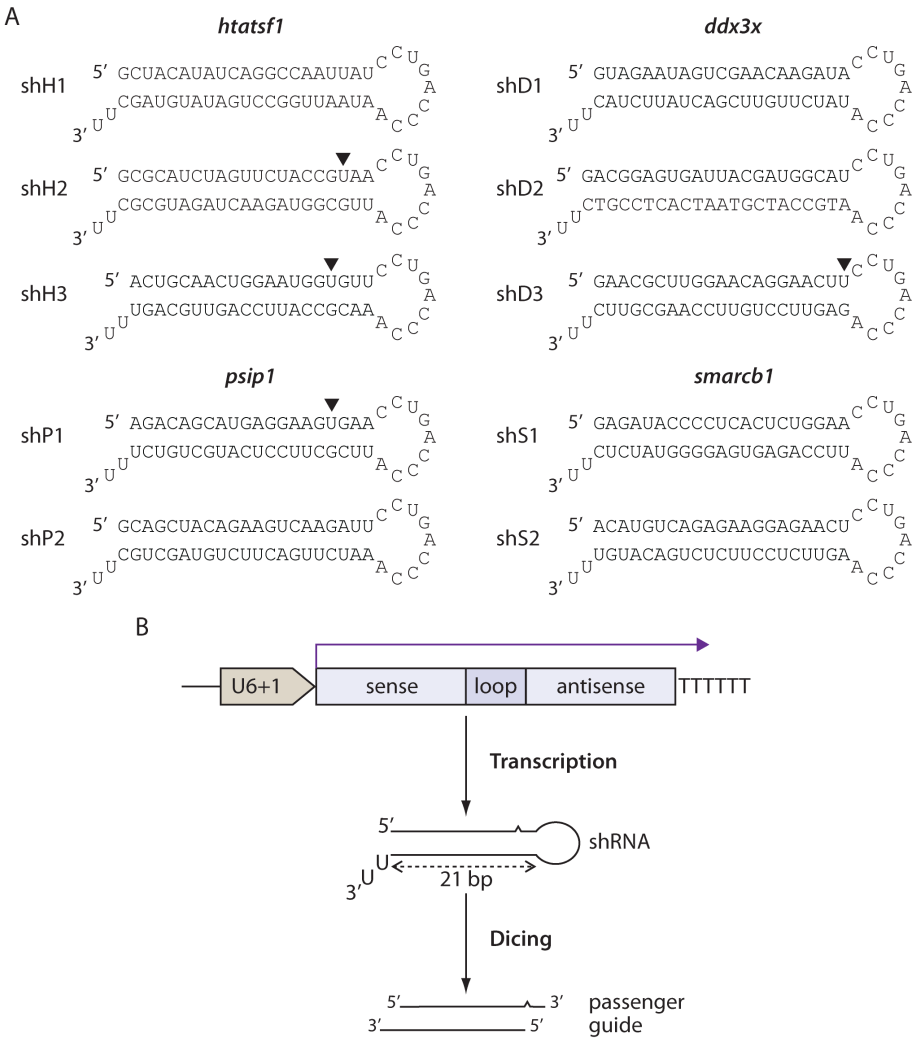


Figure 2.4: Constructed cellular factor-targeting shRNA expression cassettes. A. The cellular factors *htatsf1* and *ddx3x* were each targeted by three shRNAs and *psip1* and *smarcb1* by two shRNAs. All shRNAs contain a loop sequence derived from miR-31. G:U mismatches are indicated by a black arrowhead. **B.** shRNAs were cloned into Pol III expression cassettes by a two-step PCR method. On introduction of the shRNA plasmids into cells, RNAi pathway processing of the expressed shRNAs produced target-specific guide strands.

RNA Pol III transcription has clear start and termination signals, producing shRNAs with well-defined ends. All shRNAs commence with a purine, the preferred initial ribonucleotide for RNA Pol III transcription. The termination signal of >4 thymidines result in a 3' two nt overhang in the shRNA, which is characteristic of Drosha products, ensuring shRNA recognition by exportin 5 and Dicer.

The ability of shRNAs to suppress target expression was initially assessed using a dual luciferase reporter assay, in which shRNA target sequences are within the 3' UTR of the *Renilla* luciferase gene. Following cotransfection of HEK293 cells with cellular factor-targeting shRNA expression cassettes and cognate dual luciferase target constructs, the ratio of *Renilla*:firefly luciferase activity was decreased by at least 90% (p <0.001; Figure 2.5). Cells treated with a control shRNA, shHBVx, which targets a region within the HBV X protein, did not alter the luciferase activity ratio. Similarly, the luciferase activity ratio was unaffected in cells transfected with a pTZU6 plasmid (with no shRNA expression) or dual luciferase target constructs only. This demonstrates that the shRNA expression cassettes were able to specifically inhibit cognate target sequences inserted downstream of the *Renilla* coding region.

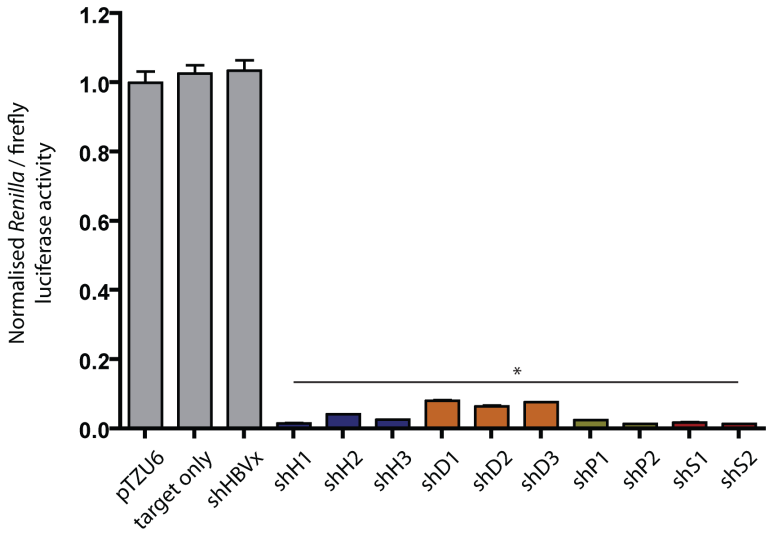


Figure 2.5: shRNAs significantly knock down their cognate targets in a dual luciferase reporter assay. Transfection of HEK293 cells with pTZU6+shRNA expression cassettes and cognate reporter constructs results in a significant inhibition of *Renilla* luciferase activity relative to firefly luciferase activity 48 hours later. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

Such inhibition does not necessarily translate into suppression of endogenous target mRNA and protein, however, as the context of the target sequence within a mRNA may affect shRNA efficacy. For example, target sites within a mRNA may be protected from RNAi by bound

protein or the formation of an obstructive structure (Overhoff *et al.* 2005; Patzel *et al.* 2005; Schubert *et al.* 2005).

Quantification of target mRNA by qRT-PCR revealed that all three shRNAs targeting *htatsf1* transcripts reduced levels relative to β-actin mRNA (*actb*) by at least 55% compared with a pTZU6 mock (p <0.01; Figure 2.6). Both shP1 and shP2 significantly reduced *psip1* transcripts (p <0.01), although shP1 mediated greater knock down (81%) than shP2 (52%). shS2 showed marginally greater efficacy than shS1 (72% and 87% knockdown, respectively; p <0.01). Thus, shRNAs targeting *htatsf1*, *psip1* and *smarcb1* mRNAs were able to suppress their endogenous cellular targets.

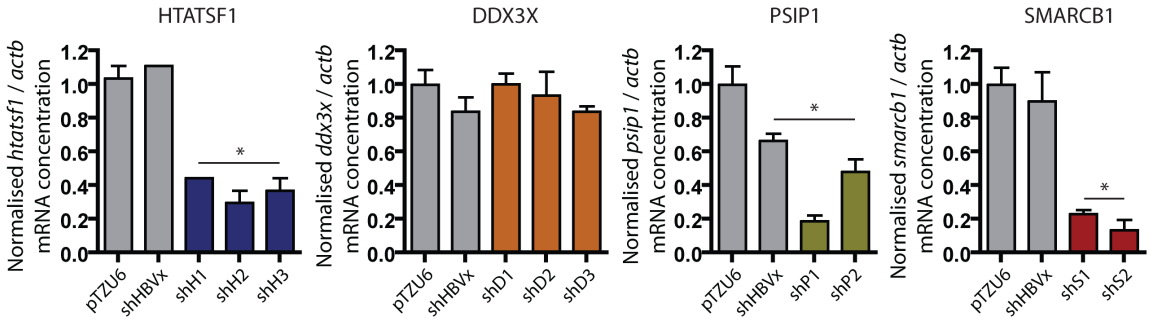


Figure 2.6: Not all shRNAs reduce levels of target mRNAs. Total RNA was analysed by qRT-PCR 48 hours post-transfection of HEK293 cells with shRNA expression plasmids in triplicate. Target mRNA levels are given relative to β-actin mRNA (*actb)*. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

In contrast, treatment of cells with shD1-3, which target *ddx3x* mRNA, did not reduce target transcript levels relative to *actb* mRNA (Figure 2.6). This suggests that target sites for shD1-3 may be inaccessible. Given that, concurrent with this work, DDX3X was validated as an HDF through shRNA-mediated suppression in a reporter cell line (Ishaq *et al.* 2008), shD1-3 were no longer pursued in this study. DDX3X was re-examined as a potential TGS target in Chapter 4.

2.3.2 shRNAs can reduce pNL4-3.Luc.R$^-$.E$^-$ luciferase activity

The effect of cellular factor-targeting shRNAs on the expression of the HIV-1 subtype B molecular clone pNL4-3.Luc.R$^-$.E$^-$, which contains a firefly luciferase reporter, was examined in HEK293 cells. This HIV molecular clone is readily transfectable into HEK cells but replication-defective as a result of two frame shifts rendering it Vpr- and Env-. Therefore, whilst this assay may not identify HDFs with roles in the formation of infectious virions or viral entry, it is a quick and convenient method for identification of proteins with a role in integration and viral gene expression. Cells were cotransfected with shRNA expression cassettes, pNL4-3.Luc.R$^-$.E$^-$ and a *Renilla* luciferase expression plasmid as a control for transfection efficiency. Two previously

developed shRNAs that target the virus itself, shU5 and shTAT (Barichievy *et al.* 2007), were used as positive controls for HIV-1 inhibition (Figure 2.7A). shU5 targets the leader RNA, just downstream of the TAR element, which is found in all HIV transcripts, including pre-genomic RNA. shTAT targets a region common to both *tat* and *rev* mRNAs.

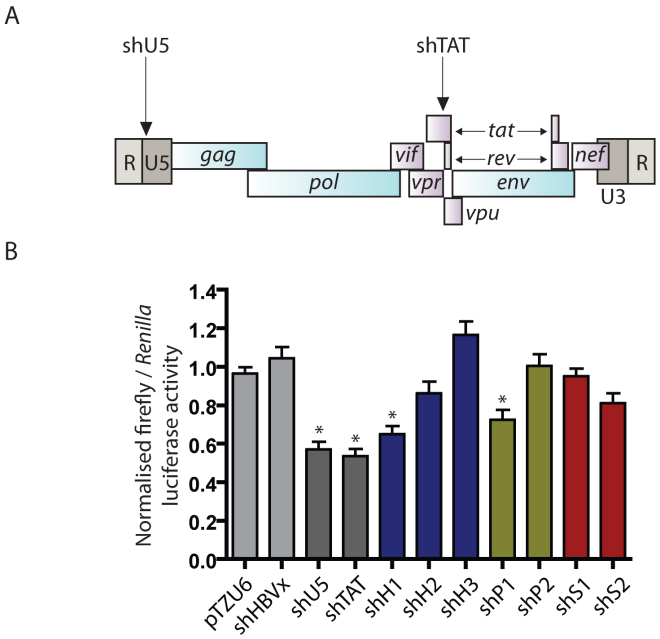


Figure 2.7: shRNAs may inhibit pNL4-3.Luc.R$^-$.E$^-$ luciferase activity. A. Schematic of the HIV-1 genome depicting the target sites of shU5 and shTAT, which were used as positive controls. **B.** HEK293 cells were transfected with shRNA expression cassettes, pRSV-RLuc and the HIV-1 subtype B molecular clone pNL4-3.Luc.R$^-$.E$^-$ and luciferase activities determined 48 hours later. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

Significant inhibition of pNL4-3.Luc.R$^-$.E$^-$ firefly luciferase activity was observed in cells that received shH1, shP1 and the two positive control shRNAs, but not either shRNA targeting SMARCB1 expression (p <0.05; Figure 2.7B). This supports a role for HTATSF1 and PSIP1 in viral gene expression and/or integration, and suggests that shH1 and shP1 may be the most potent shRNAs against their respective targets in the context of HIV-1 infection. The latter is likely to be a result of shP2 targeting a region in *psip1* common to both p52 and p75 isoforms, whereas shP1 is p75-specific, the isoform required for HIV integration. shH1-3 reduced mRNA levels to a similar extent (Figure 2.6), thus it is surprising that shH1 is more effective than shH2 and shH3 in inhibiting luciferase expression from pNL4-3.Luc.R$^-$.E$^-$. This may reflect differences in the kinetics of mRNA degradation: if shH1-derived guide strands were expressed more efficiently, resulting in a more rapid decline in HTATSF1, than those from shH2 and shH3, it may give rise to increased efficacy against pNL4-3.Luc.R$^-$.E$^-$ without reducing *htatsf1* mRNA concentrations 48 hours post-transfection to a greater extent.

Overall, firefly luciferase inhibition did not exceed 50%, which may be a result of the low expression rate of pNL4-3.Luc.R$^-$.E$^-$ and it being competent for a single round of replication only (Connor *et al.* 1995; He *et al.* 1995). This system suffers from additional limitations: transfection of cells with a HIV-1 plasmid is not a good model for HIV-1 infection, particularly when validating the roles of host proteins in the life cycle, as the mechanism of viral entry differs. Similarly, the import of cellular factors that interact with Vpr and Env, which are deleted in pNL4-3.Luc.R$^-$.E$^-$, may be overlooked. To overcome the limitations associated with transfection of pNL4-3.Luc.R$^-$.E$^-$, experiments moved to a different cell line that permitted the use of a virus capable of productive replication and HIV receptor-mediated entry, whilst maintaining the ability to be easily transfected with shRNA expression cassettes.

2.3.3 shRNAs suppress cognate targets in TZM-bl cells by an RNAi mechanism

TZM-bl cells are derived from JC.53, a HeLa cell line that stably express large amounts of the HIV receptors CD4 and CCR5. The TZM-bl cell line was generated from JC.53 cells by introducing separate integrated copies of the luciferase and β-galactosidase genes under control of the HIV-1 promoter (Platt *et al.* 1998; Derdeyn *et al.* 2000; Wei *et al.* 2002). TZM-bl cells are, therefore, permissive to infection with diverse isolates of R5-tropic HIV-1 and are a useful indicator cell line for quantitative analysis of HIV, using either luciferase or β-galactosidase as a reporter.

Before assessing the ability of shRNAs to inhibit HIV-1 replication TZM-bl cells, it was first verified that the shRNAs suppressed their targets in these cells, as shRNA efficacy may vary between cell lines. The ability of shRNAs to inhibit cognate targets in HeLa cells, from which TZM-bl are derived, was verified with a dual luciferase reporter assay. All shRNAs decreased *Renilla*:firefly luciferase activity by greater than 95% (p <0.001; Figure 2.8A), which was similar to that observed in HEK293 cells (Figure 2.5). Suppression of target mRNA levels relative to *actb* mRNA when normalised to a pTZU6 control was also similar in TZM-bl cells to HEK293 cells; all cellular factor-targeting shRNAs significantly suppressed levels of their cognate target transcripts (p <0.01; Figure 2.8B). shP1 and shS2 were more effective at reducing *psip1* and *smarcb1* transcripts, respectively, than shP2 (91% versus 88% knock down) and shS1 (71% versus 61% knock down). shH1 reduced *htatsf1* mRNA levels to a lesser degree than shH2 and shH3 (38% versus 60% knock down).

Guide strand processing was confirmed for all shRNAs by small RNA Northern blot (Figure 2.8C). Conclusions cannot be drawn with respect to the relative levels of guide strands produced, as different probes were required for their detection that differ in their hybridisation efficiency.

Nonetheless, the detection of guide strands from all cellular factor shRNAs supports that the *htatsf1*, *psip1* and *smarcb1* mRNA silencing was mediated by an RNAi mechanism.

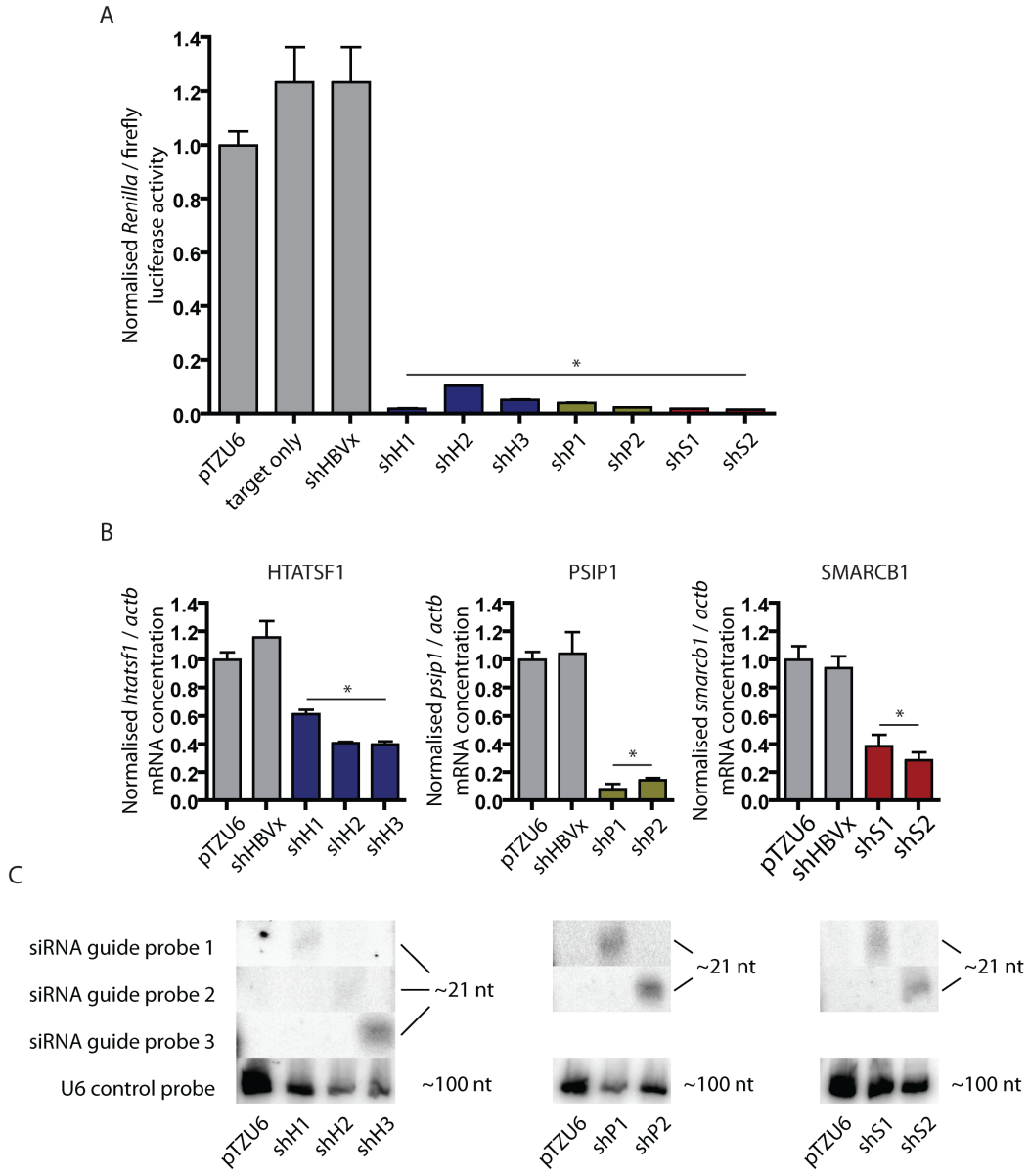


Figure 2.8: shRNAs knock down target mRNAs by an RNAi mechanism. **A.** shRNAs suppress cognate targets in a dual luciferase reporter assay in HeLa cells 48 hours post-transfection. **B.** Endogenous target mRNA levels are reduced by shRNAs in HeLa-derived TZM-bl cells relative to β-actin mRNA (*actb*). **C.** Guide strands, the product of RNAi processing of shRNAs, were detected by small RNA PAGE Northern blot in TZM-bl cells 48 hours post-transfection. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

2.3.4 shRNAs inhibit HIV-1 Tat-induced luciferase activity in TZM-bl reporter cells

Infection of TZM-bl cells with HIV-1 results in integration of the proviral DNA and, subsequently, low-level production of the viral transactivator Tat. Tat binding to the TAR element in the viral promoter, or LTR, leads to a dramatic increase in transcription and thus, viral gene

expression. The integrated luciferase genes in TZM-bl are downstream of viral LTRs such that Tat can also stimulate transcription of luciferase (Figure 2.9). Consequently, assays for luciferase activity in TZM-bl cell lysates provide a simple, quantitative indicator of HIV-1 replication.

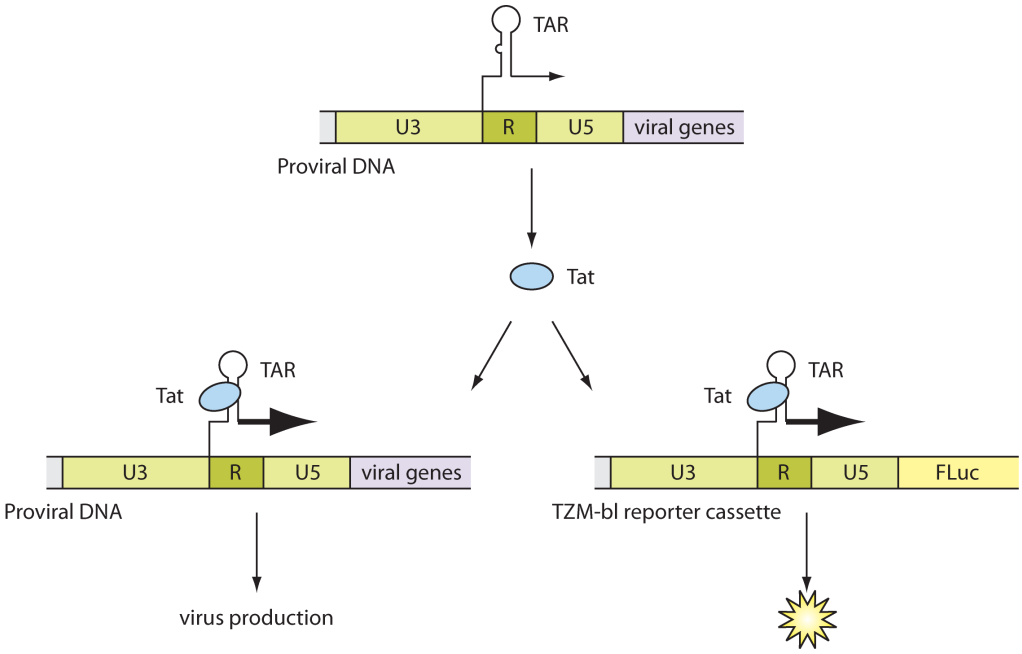


Figure 2.9: Schematic of the HIV-reporter ability of TZM-bl cells. Following HIV infection and integration of the proviral DNA, low levels of transcription produce the viral regulatory protein Tat. Tat can bind the TAR element in nascent RNA transcripts at the viral promoter resulting in a dramatic increase in viral gene expression. TZM-bl cells contain an integrated firefly luciferase gene driven by the HIV promoter. As such, on HIV infection of these cells, Tat also induces luciferase expression.

TZM-bl cells were transfected with various amounts of shRNA expression plasmids prior to infection with the subtype C envelope-pseudotyped subtype B molecular clone pSG3$^{\Delta env}$.ZM53M at a $TCID_{50}$ of 1000/ml. In contrast to pTZU6 and shHBVx negative controls, a shRNA dose-dependent suppression of Tat-induced luciferase activity was observed, with maximum suppression typically when 500 ng of shRNA expression plasmid was transfected (Figure 2.10A). The only exception being shH3, which exhibited the greatest inhibition of luciferase activity at 1000 ng, possibly reflecting differences in the kinetics of guide strand processing. Since the majority of shRNAs exhibited greatest inhibition of Tat-induced luciferase activity at 500 ng, and the lowest effective dose is preferable to minimise toxicity, this amount was used in subsequent experiments. At this dose, the greatest inhibition of luciferase activity was with HTATSF1-targeting shH1 and shH2 (~60%, p <0.01), with the exception of positive controls shU5 and shTAT (~80%, p <0.001; Figure 2.10B). shP1 and shS2 also significantly inhibited luciferase activity (p <0.01), although not to the same extent (43% and 30%, respectively).

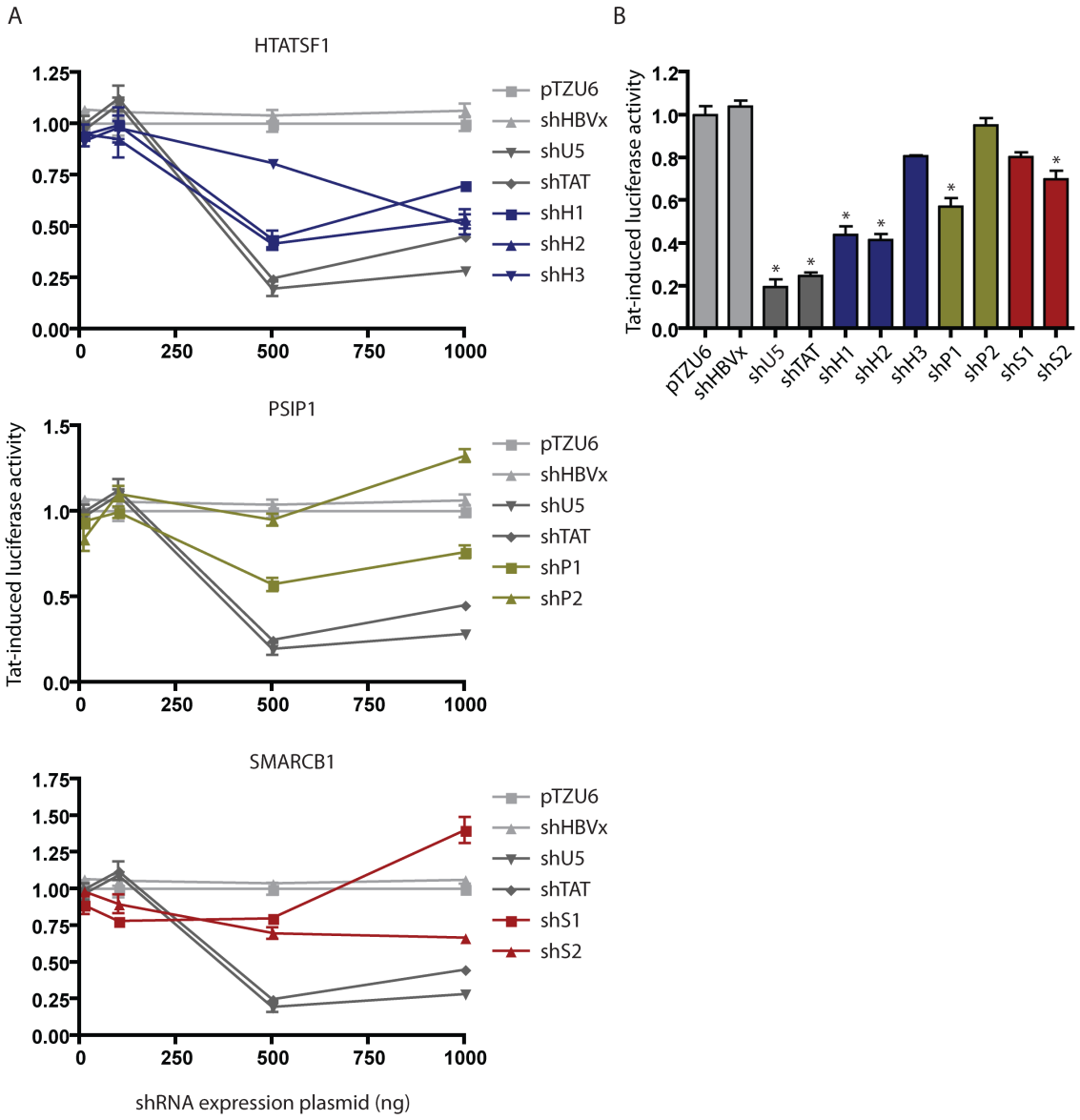


Figure 2.10: shRNAs inhibit Tat-induced luciferase activity in a dose-dependent manner. **A.** TZM-bl cells were infected with pSG3$^{\Delta env}$.ZM53M at a $TCID_{50}$ of 1000/ml 24 hours post-transfection with various quantities of shRNA expression plasmids. Firefly luciferase activity was determined 48 hours post-infection in triplicate. **B.** As (A) but where 500 ng of shRNA expression plasmid transfected only. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

Similar results were obtained when cells were infected with a different subtype B molecular clone, p81A-4, and a virus passaged from a subtype C primary isolate, FV5 (Figure 2.11). In general, the ability of shRNAs to inhibit Tat-induced luciferase activity was reduced in the FV5-infected cells, including positive controls. This may indicate that the propagation rate of FV5 is greater than the subtype B molecular clone-derived viruses used, which has been observed previously (personal communication of M. Papathanasopoulos, S. Barichievy and S. Saayman). Regardless, repeatedly the same shRNAs display the greatest efficacy, irrespective of whether pNL4-3.Luc.R$^-$.E$^-$, pSG3$^{\Delta env}$.ZM53M, p81A-4 or FV5 was used, namely shH1, shP1 and shS2 for each cellular factor targeted, which became the focus of subsequent experiments.

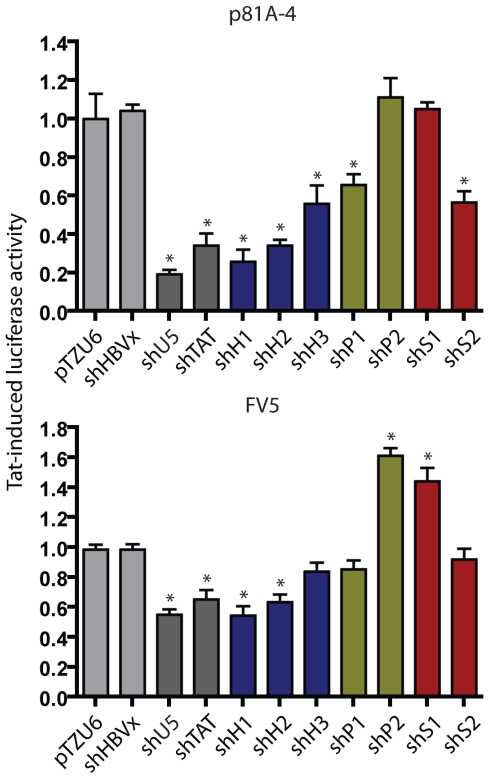


Figure 2.11: shRNAs mediate similar effects on different subtypes of HIV-1. TZM-bl cells were infected with either subtype B molecular clone p81A-4 or virus passaged from subtype C isolate FV5, both at a $TCID_{50}$ of 1000/ml, 24 hours post-transfection. Firefly luciferase activity was determined 48 hours post-infection. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

2.3.5 HTATSF1-targeted shRNA inhibits HIV-1 infectious particle production

The TZM-bl luciferase reporter system is less sensitive for factors affecting the late stages of viral replication, including assembly and budding, compared to those involved in early stages (Brass *et al.* 2008). Therefore, to indicate the import of targeted cellular factors in the entire viral life cycle, whether shRNA-targeting of HTATSF1, PSIP1 and SMARCB1 could inhibit infectious particle production was examined. Culture supernatants from primary-infected cells were assayed for levels of the HIV-1 antigen p24 and the ability to induce luciferase activity in fresh reporter cells, thereby measuring infectious progeny (Figure 2.12).

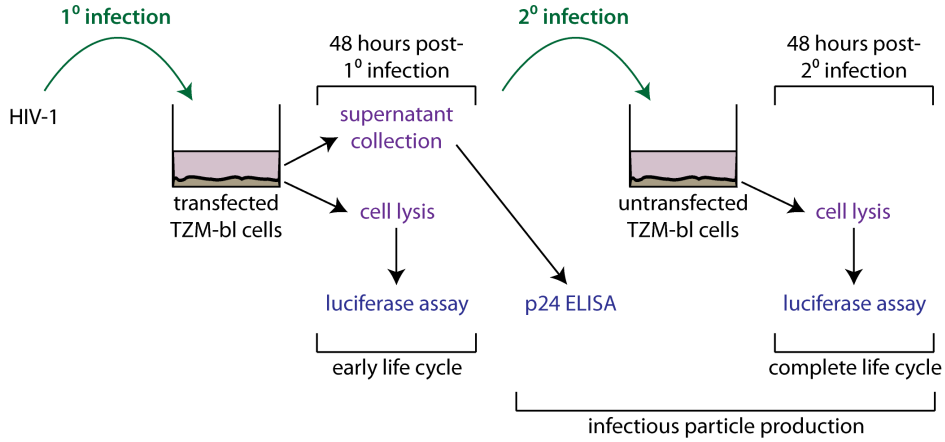


Figure 2.12: Schematic of the TZM-bl reporter assay for HIV-1 infectious particle production. Following a primary infection of shRNA-transfected TZM-bl cells, the HIV-1 antigen p24 was quantified in culture supernatant by ELISA. In addition, some of the culture supernatant was used in a secondary infection of untransfected TZM-bl cells that were analysed for Tat-induced firefly luciferase activity 48 hours later.

None of the shRNAs targeting host factors was able to suppress either p81A-4 or FV5 p24 production, in contrast to positive controls shU5 and shTAT (p <0.05; Figure 2.13). Nonetheless, shH1 did mediate a significant drop in p81A-4 infectious particle production (62%, p <0.01; Figure 2.14), quantified by luciferase activity in cells receiving culture supernatant from shRNA-treated cells. This discord suggests that HTATSF1 may have a role in HIV-1 replication that is required for infectious particle, but not p24, production. shP1 and shS2 had no effect on the amount of infectious progeny produced, whether quantified by p24 levels or secondary luciferase assay.

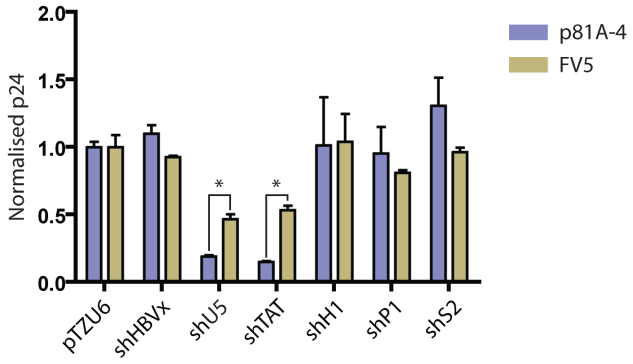


Figure 2.13: Host factor-targeting shRNAs do not inhibit HIV-1 p24 production. TZM-bl cells were transfected with 500 ng shRNA expression plasmids and infected with either p81A-4 or FV5 at a $TCID_{50}$ of 1000/ml 24 hours later. HIV-1 p24 was quantified in cell culture media 48 hours post-infection by ELISA. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

The contrast between the inhibition in primary luciferase activities but unaltered p24 levels with shH1, shP1 and shS2 suggests a bias in the primary luciferase activity assay toward factors influencing early stages of the life cycle, which may be related, in part, to the numerous copies of the integrated LTR-luciferase reporter gene. Nevertheless, the lack of infectious particle inhibition from shP1-treated cells was surprising given the essential role of PSIP1 in HIV-1

integration (Llano *et al.* 2006a). Whether the lack of infectious particle inhibition in cells treated with shP1 and shS2 was a result of insufficient protein suppression was investigated.

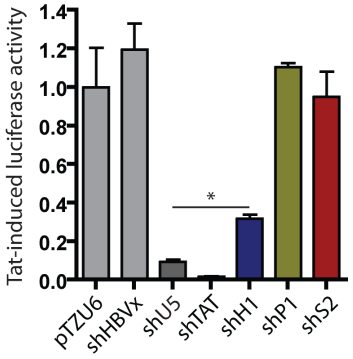


Figure 2.14: HTATSF1-targeting shRNA inhibits HIV-1 infectious particle production. Culture media were collected from shRNA expression plasmid-transfected TZM-bl cells 48 hours post-infection with p81A-4 at a $TCID_{50}$ of 1000/ml and used to infect untransfected TZM-bl cells. Firefly luciferase activity was determined 48 hours post-secondary infection. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

2.3.6 shRNA-mediated HTATSF1 suppression is rapid but PSIP1 and SMARCB1 have long half-lives

The TZM-bl reporter system employed transient expression of shRNAs following transfection of expression cassettes. Primary viral infections followed just 24 hours later, at which time maximal target protein suppression may not have been attained. The kinetics of target protein suppression was determined by Western blot.

Relative to β-actin, HTATSF1 levels, although unclear 48 hours post-shH1 transfection, were significantly reduced to 99.5% of the shHBVx control 72 hours post-transfection (Figure 2.15). In contrast, p75 levels were only marginally reduced at 48 hours post-shP1 transfection and still at 77% of shHBVx levels 96 hours post-transfection. Together with the significant inhibition of mRNA levels 48 hours post-shP1 transfection (>90%, p <0.001; Figure 2.8B), this suggests that p75 has a slow turnover. This may account for the marginal inhibition of Tat-induced luciferase activity following primary infection and the lack of inhibition of HIV-1 infectious particle production. Similarly to HTATSF1, SMARCB1 suppression was maximal at 72 hours post-shS2 transfection, though only dropping to 47.5% of shHBVx (Figure 2.15). Hence, suppression of SMARCB1 may not have been sufficient to have a detrimental impact on viral particle production, as with p75.

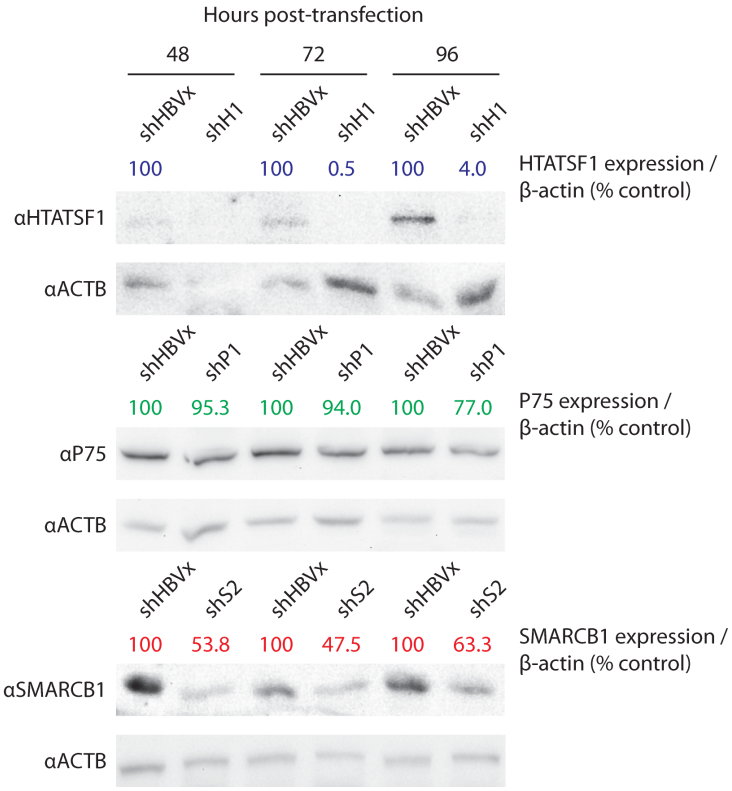


Figure 2.15: The half-lives of targeted host proteins differ. TZM-bl cell lysates were subject to PAGE and Western blot at 48, 72 and 96 hours post-shRNA expression plasmid transfection. Expression is given relative to β-actin and normalised to cells treated with the negative control shHBVx at each time point.

2.3.7 shRNA-mediated host factor suppression is not cytotoxic

Having determined that maximal HTATSF1 and SMARCB1 suppression was attained 72 hours post-transfection (Figure 2.15), the extent of apoptosis induction at this time point was quantified by fluorescence microscopy of FITC-conjugated Annexin V-stained cells (Figure 2.16). Annexin V binds to phosphotidylserine, which becomes exposed on disruption of phospholipid asymmetry during the early stages of apoptosis.

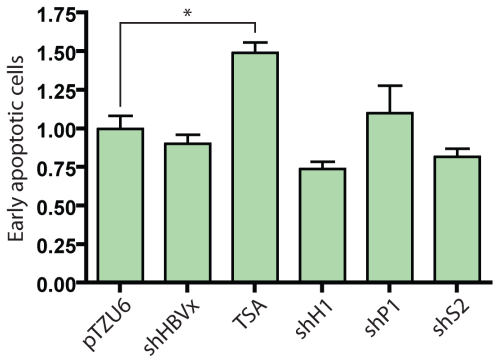


Figure 2.16: shRNA-mediated target suppression does not induce apoptosis. TZM-bl cells were stained with Annexin V-FITC 72 hours following transfection with shRNA expression plasmids. As a positive control for apoptosis induction, cells were treated with TSA 16 hours pre-stain. Four images of FITC and transmitted light were acquired of stained cells per sample and analysed by ImageJ. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

No increase in early apoptotic cells was detected in cells transfected with any shRNA expression plasmids (Figure 2.16). In contrast, a significant increase in early apoptotic cells was seen when cells were treated with 500 nM of TSA, which was used as a positive control for apoptosis induction as it is toxic at high concentrations.

It is important to characterise as fully as possible whether suppression of cellular targets has cytotoxic effects. Recently, it was suggested that HTATSF1, rather than being a general elongation factor, plays a role in splicing of HIV-1 transcripts (Miller *et al.* 2009). This report also proposed that its role in splicing extended to endogenous cellular transcripts, although data in support of the latter was not provided. To interrogate this possibility, the effects of HTATSF1 suppression on general cellular splicing were examined using pEGFP-C1-NAD, a GFP reporter construct with an inserted intron from the NADH-coenzyme Q reductase gene (NAD) (Figure 2.17A). A construct with mutated splice sites was used as a negative control for GFP production.

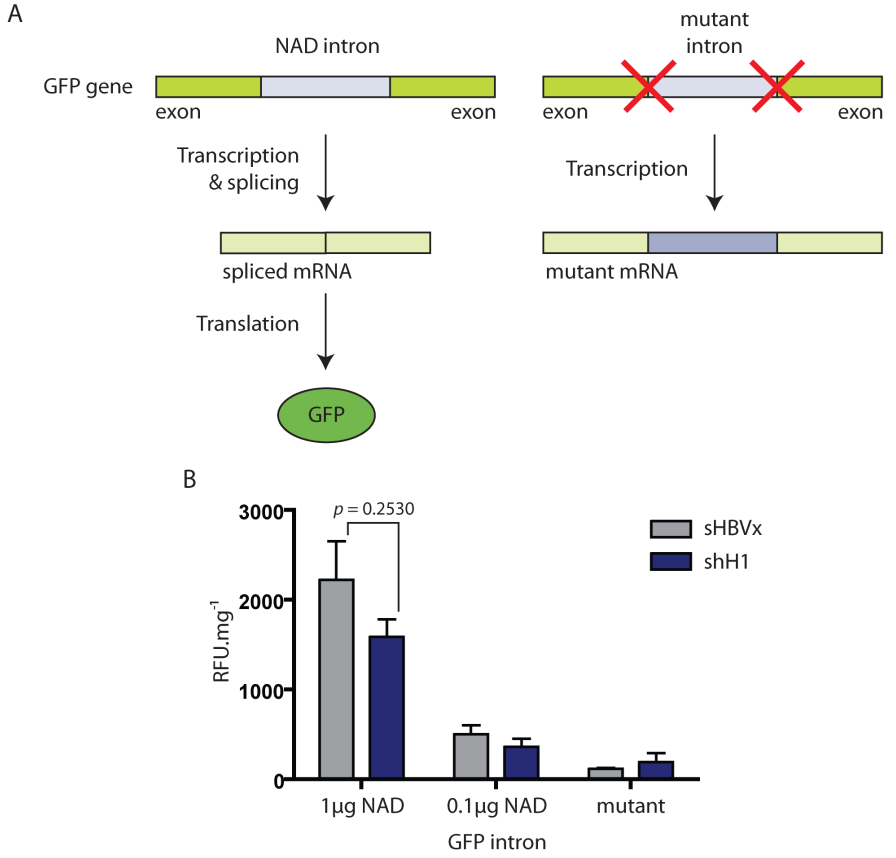


Figure 2.17: HTATSF1 suppression does not effect splicing of a GFP reporter. A. Schematic of GFP reporter constructs used in splicing assay. An intron from the NAD gene was inserted within a GFP expression cassette such that GFP was only expressed following normal transcription, splicing and translation. A construct with mutated splice sites, where no GFP was produced, was used as a control. **B.** TZM-bl cells were transfected with HTATSF1-targeting shH1 expression plasmid or negative control shHBVx. Forty-eight hours later, cells were transfected with either the NAD- or mutant-intron GFP construct in triplicate. Fluorescence was determined by fluorometry another 48 hours later and reported relative to total cellular protein. One-way ANOVA with Dunnett post-tests relative to pTZU6 and an unpaired, two-tailed t-test.

Significant differences in the production of GFP, relative to total cellular protein, were not observed between cells treated with either 1.0 or 0.1 μg of shHBVx or shH1 expression plasmids (p >0.25; Figure 2.17B). This demonstrates that HTATSF1 is not critical for splicing of such a reporter construct and may not be involved in cellular splicing.

2.3.8 U6-driven shRNA expression in TZM-bl cells is not cytotoxic and does not activate an immune response

To verify that the high levels of U6-driven shRNA expression were not cytotoxic, the colourimetric MTT assay was performed to indicate the number of viable cells. TSA was again used as a positive control for cytotoxicity. None of the shRNA expression plasmids altered the number of viable cells 48 hours post-transfection compared to the pTZU6 control (Figure 2.18A). In addition, levels of eGFP expression in cells cotransfected with shRNA expression cassettes and pCI-eGFP were unchanged, determined by fluorescence microscopy (Figure 2.18B). Together this provides good evidence that the levels of shRNAs being produced in transfected TZM-bl cells over the short time-scales involved in the HIV-1 reporter assays were not cytotoxic.

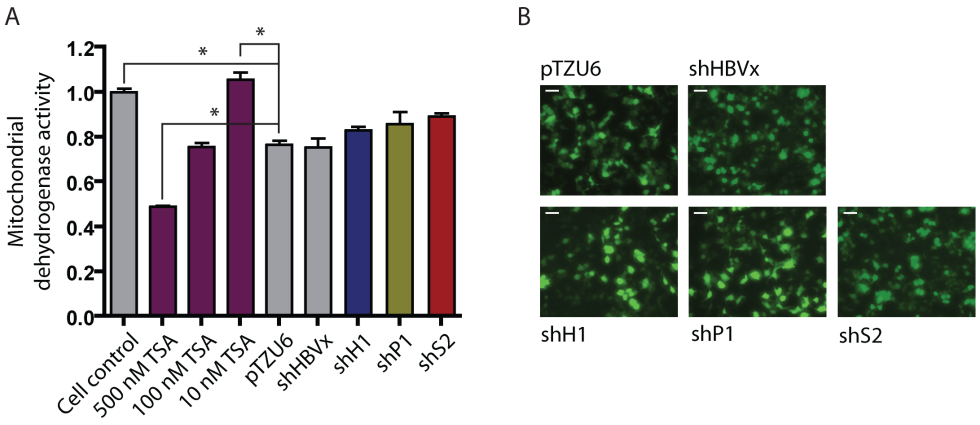


Figure 2.18: shRNA expression does not alter cell viability. **A.** TZM-bl cells were treated with MTT 48 hours post-shRNA expression plasmid transfection. TSA treatment was used as a positive control for reduced cell viability. **B.** Fluorescence microscopy images of overlayed GFP and transmitted light. Bars 100 μm. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

Double-stranded RNAs (dsRNAs) may induce a type I immune response, a component of which is increased interferon-β expression (Stark *et al.* 1998). However, none of the cellular factor-targeting shRNAs significantly increased interferon-β mRNA (*ifnb1*) levels relative to *actb* mRNA compared to the pTZU6 control, in contrast to the positive control poly(I:C) that caused a >10-fold increase (p <0.001; Figure 2.19). This suggests that the shRNAs are not inducing an immune response.

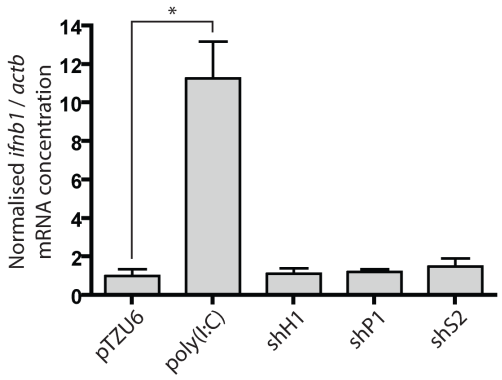


Figure 2.19: shRNAs do not induce an immune response. Levels of interferon-β mRNA (*ifnb1*) relative to β-actin mRNA (*actb*) were determined by qRT-PCR on total cellular RNA extracted 48 hours post-transfection in triplicate. One microgram of poly(I:C) RNA was used as a positive control for immune response induction. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to pTZU6.

The MTT assay, fluorescence microscopy and interferon-β expression data suggest that any effects mediated by the shRNAs on viral replication are not a result of cytotoxicity and immune response induction. Thus, the results in this chapter show that RNAi-mediated suppression of the cellular factor HTATSF1 can inhibit Tat-induced luciferase activity in HIV-1-infected TZM-bl cells without cytotoxic effects. However, long half-lives of targets PSIP1 and SMARCB1 hampers elucidation of their role in viral replication when using the TZM-bl system and transient shRNA expression.

2.4 Discussion

The approach used in this study was to validate the role of cellular factors HTATSF1, DDX3X and SMARCB1 in HIV-1 replication through shRNA-mediated PTGS. Whilst the results confirm that HTATSF1 is a HDF, they are inconclusive with respect to DDX3X and SMARCB1.

shRNAs were unable to suppress *ddx3x* mRNA levels (Figure 2.6), despite considerable suppression of cognate targets in a dual luciferase reporter assay (Figure 2.5). This demonstrates the drawback of using target reporters, as the isolated target sequence may be more accessible to RNAi than when in the context of the endogenous mRNA, such that the reporter system lacks physiological relevance. Several mechanisms can protect a mRNA from RNAi, including mRNA-bound proteins blocking the target site, the adoption of an obstructive structure and, possibly, localisation to a subcellular compartment that is not surveyed by the RNAi machinery (Overhoff *et al.* 2005; Patzel *et al.* 2005; Schubert *et al.* 2005). Such factors may also account for the variation in mRNA suppression observed with the shRNAs targeting HTATSF1, PSIP1 and SMARCB1 (Figures 2.6 and 2.8B).

The relationship between mRNA and protein levels is complex, however. Of the three genes whose mRNA levels were suppressed by shRNAs, *htatsf1* mRNA displayed the least inhibition (~60%; Figures 2.6 and 2.8B), although the greatest suppression of HTATSF1 protein (>95%; Figure 2.15). Conversely, suppression of *psip1* mRNA and p75 protein was high and low, respectively. This demonstrates that the kinetics of mRNA and protein turnover, transcription and translation vary between genes. This means that use of RNAi in TZM-bl cells to validate HDFs may lead to notable omissions, such as p75. The long half-life of p75 identified here has been reported previously (Llano *et al.* 2006a) and it is not surprising, therefore, that PSIP1 was notably absent from the lists of HDFs identified in large-scale RNAi screens (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b). Thus, despite its inclusion as a positive control for HDF suppression-mediated inhibition of HIV-1 replication, in this reporter system, it did not fulfil this role.

Of note is that only shP1 caused a slight inhibition of Tat-induced luciferase activity and not shP2 (Figures 2.10 and 2.11). shP2 targets a region common to both p52 and p75 isoforms, although only p75 is involved in HIV-1 integration. The relative levels of p52 and p75 in TZM-bl cells are unknown but it is possible that p52 acts as a sponge for shP2 such that less shRNA is available for targeting of p75. Furthermore, it has been reported that p75 levels need to be reduced by >95% to have an effect on HIV-1 replication, on the basis that only at this level of suppression are proteins depleted from chromatin (Llano *et al.* 2006a). Thus, it seems the slight difference in p75 mRNA suppression mediated by shP1 and shP2 (~93% and ~86%, respectively;

Figure 2.8B) in TZM-bl cells is sufficient to result in a significant difference in the ability of the shRNAs to inhibit HIV-1 integration. Unfortunately, Western blot analysis of shP2-treated cells was not obtained to test this hypothesis.

From the results reported here it remains unclear whether SMARCB1 is a HDF. The moderate inhibition of Tat-induced luciferase activity following primary infection (~20%; Figures 2.10 and 2.11) suggests that SMARCB1 does play a viral cofactor role during an early stage of infection. This fits with other studies that have reported a role for SMARCB1 in the early life cycle, including nuclear import of the PIC (Turelli *et al.* 2001) and remodelling of the +1 nucleosome to facilitate Tat-mediated activation of HIV-1 transcription (Ariumi *et al.* 2006; Mahmoudi *et al.* 2006; Treand *et al.* 2006). In addition, a more recent study has proposed that the Brm-type SWI/SNF complex is required for long transcript synthesis in the absence of Tat during the very earliest stages of Tat-independent viral transcription (Mizutani *et al.* 2009). Conversely, another study proposed an antiviral role for SMARCB1 during early stages of infection (Maroun *et al.* 2006). Central to this observation, however, was the use of an HIV-1 molecular clone containing an integrase (IN) mutant (K71R) that retained catalytic activity and the ability to bind p75, but a 2-fold decreased association with SMARCB1. The possibility that such a mutant was capable of enhanced association with other factors that resulted in increased viral replication was not considered. Further, overexpression of SMARCB1 suppressed replication (Maroun *et al.* 2006), which was seen as supporting an antiviral role. However, overexpression may have destabilised the higher-order structure of IN, as occurs with overexpression of the p75 integrase-binding domain (Llano *et al.* 2006b; Hombrouck *et al.* 2007). Subsequently, it has been demonstrated that multimerisation of SMARCB1 is necessary for its ability to inhibit integration and influence particle production (Das *et al.* 2009).

Several studies have demonstrated SMARCB1 inclusion in HIV-1 virions (Yung *et al.* 2001; Yung *et al.* 2004; Sorin *et al.* 2006). Furthermore, incorporation is restricted to HIV-1 and it is not found in other closely related primate lentiviruses, such as HIV-2 or SIV-1, or oncoretroviruses, such as Mo-MLV (Yung *et al.* 2004). The incorporation of SMARCB1 was correlated with the infectivity of HIV-1 particles and reverse transcriptase activity (Sorin *et al.* 2006). Mutants of IN have been shown to affect reverse transcription and this function has been ascribed to the direct binding of IN to RT (Wu *et al.* 1999; Hehl *et al.* 2004; Zhu *et al.* 2004). Perhaps SMARCB1 association with IN stabilises the RT complex, although it cannot be excluded that SMARCB1 is necessary for a step prior to reverse transcription or that reduced infectivity of virions could be due to a lack of other SMARCB1-associated proteins, such as the SAP18-HDAC1 complex (Sorin *et al.* 2009). Admittedly, no evidence for a role of SMARCB1 in infectious particle production was obtained here (Figures 2.13 and 2.14).

It is unclear why moderate PSIP1 and SMARCB1 suppression inhibited Tat-induced luciferase activity in primary-infected cells but not p24 levels and infectious particle production. One explanation predicts that these proteins have two independent roles in HIV-1 replication and that the levels of the protein required to fulfil these roles differ. For example, that 50% SMARCB1 suppression in TZM-bl cells is insufficient to maintain its putative function in HIV-1 transcription but sufficient for its putative incorporation into virions. Interestingly, several members of the SWI/SNF complex are up-regulated in actively-infected versus latently-infected cells and activated- versus resting-infected cells (Krishnan and Zeichner 2004; Mizutani *et al.* 2009). Dual roles in HIV-1 replication for PSIP1 have not been described, however.

Alternatively, the discord in the effects of PSIP1 and SMARCB1 suppression had on Tat-induced luciferase activity and HIV-1 infectious particle production may be a result of biased sensitivity of the system toward factors influencing early stages of the life cycle. This may be understood with metabolic control analysis. In simple terms, the luciferase reporter output is reached in fewer 'steps' than p24 and infectious particle outputs. Following integration and low-level Tat production, only transcription and translation are necessary for luciferase activity. In contrast, full-length viral transcripts, from which p24 is derived, are only produced during the later stages of viral transcription (Schwartz *et al.* 1990). Furthermore, p24 production requires not just transcription and translation, but also assembly of Gag and Gag-Pol precursor proteins with viral RNA to form immature viral particles, release of PR and cleavage of the polypeptides to form p24 and mature viral particles with a condensed core (Yoshinaka and Luftig 1977; Wiegers *et al.* 1998). Inhibition of pathway flux is dependent not just on the extent of suppression of an intermediate, but also on the relative rates of transition from one intermediate to the next. Thus, as particle assembly, maturation and budding are considerably slower than Tat-activated transcription, the same degree of PSIP1 and SMARCB1 suppression may inhibit the reporter output but not the p24 output. This is because although the production rates of the components necessary for infectious particle production are reduced, the slow rate of the assembly, maturation and budding processes ensures that there is sufficient accumulation of those components for infectious particle production to be unaffected.

Although this hypothesis may account for the lack of inhibition of p24 production in shH1-treated cells (Figure 2.13), it does not account for the discord in the effects of HTATSF1 suppression on p24 levels and infectious particle production (Figures 2.13 and 2.14); particularly as HTATSF1 levels were dramatically reduced (>80% 48 hours post-transfection; Figure 2.15) and inhibition of primary Tat-induced luciferase activity was similar to that mediated by the positive controls, shU5 and shTAT (Figures 2.10 and 2.11). The data rather suggests a role for HTATSF1 that impacts infectious particle production but not p24 levels. Besides a role in transcription

elongation, two other roles have been suggested for HTATSF1 in viral replication: (1) splicing of HIV-1 transcripts (Miller *et al.* 2009); and, (2) formation of influenza viral ribonucleoprotein complexes (vRNP) (Naito *et al.* 2007).

Splicing is known to occur cotranscriptionally (Moore and Proudfoot 2009; Perales and Bentley 2009) and it has been shown that, primarily through the first RNA recognition motif, HTATSF1 interacts with spliceosomal U small nuclear ribonucleoproteins (snRNPs) (Yan *et al.* 1998; Fong and Zhou 2001), which strongly stimulates HIV-1 transcription elongation (Fong and Zhou 2001). As the HTATSF1-U snRNP complex also stimulates splicing *in vitro* (Fong and Zhou 2001), HTATSF1 may serve as a dual-function factor to facilitate reciprocal activation of transcription and splicing. A more recent study suggests that HTATSF1 plays no role in HIV-1 transcription elongation at all, affecting HIV-1 splicing only (Miller *et al.* 2009), which is in contrast to all previous investigations into the role of HTATSF1 in HIV-1 replication (Zhou and Sharp 1996; Li and Green 1998; Kim *et al.* 1999; Parada and Roeder 1999; Fong and Zhou 2001; Zhou *et al.* 2004a). Although a unique role of HTATSF1 in splicing is consistent with the data, since p24 is not derived from spliced transcripts (Schwartz *et al.* 1990), an elegant study has confirmed HTATSF1 as a facilitator of efficient RNA Pol II elongation (Chen *et al.* 2009).

HTATSF1 was found to stimulate influenza viral RNA synthesis through chaperoning the viral nucleoprotein (NP) and enhancing its association with viral RNA to form vRNPs (Naito *et al.* 2007). These complexes are the basic unit for transcription and replication of the virus genome. Several splicing proteins are candidate host cofactors for influenza virus replication, and influenza gene transcription depends on functional RNA Pol II (Chan *et al.* 2006). Thus, HTATSF1 may enhance viral RNA synthesis through promoting the association of transcriptionally functional RNA Pol II, vRNPs and splicing factors. Interestingly, the yeast homolog of HTATSF1, CUS2, chaperones U2 snRNP prior to spliceosomal assembly in a manner dependent on the first RRM domain (Yan *et al.* 1998). Thus, together the data suggests a role of HTATSF1 in enhancing RNA synthesis through chaperoning snRNPs and coupling splicing to elongation, which may be common to influenza virus, HIV-1 and normal cellular function.

There was no evidence of cytotoxicity associated with shRNA-mediated HTATSF1 suppression (Figures 2.16, 2.17, 2.18 and 2.19). Perhaps the levels of HTATSF1 needed for efficient HIV-1 replication are higher than those required for its normal cellular functions. Also, there are several factors capable of enhancing splicing and elongation in cells, and this redundancy may protect cells from toxic side effects associated with HTATSF1 suppression and explain why no significant decrease in splicing of a GFP reporter was observed (Figure 2.17). In contrast, productive elongation of HIV-1 transcripts involves a highly coordinated series of events, including recruitment of Tat and cyclin T1 to the TAR element, phosphorylation of the CTD of RNA

Pol II and remodelling of the +1 nucleosome. The role of HTATSF1 in this process may be more critical than in general elongation.

In summary, HTATSF1 has been validated as an HDF, although how its suppression inhibits infectious particle production without altering p24 levels is unclear. In addition, expressed RNAi mimics can suppress HTATSF1 efficiently without cytotoxic effects. This implicates HTATSF1 as an anti-HIV therapeutic target. However, the TZM-bl reporter system for the identification of HDFs is hampered by the inherent bias for factors involved in early stages of the life cycle. Further, the kinetics of viral replication may significantly differ in this cell line compared to a physiological substrate of HIV eg CD4+ T cells, where host gene expression is different. Moreover, transient expression of RNAi mimics may be insufficient to suppress target proteins with long half-lives. Consequently, PSIP1 did not function as a positive control as hoped and whether SMARCB1 is an HDF or not remains unclear. Therefore, further validation of the anti-HIV therapeutic potential of targeting cellular factors HTATSF1 and SMARCB1 required a T cell-derived line, stably expressing RNAi mimics. This was the objective of the work described in Chapter 3.

CHAPTER 3

Persistent suppression of HIV-1 replication in a CD4+ SupT1 cell line using lentivirus-delivered shRNAs targeted to HTATSF1

3.1 Introduction

The work in Chapter 2 demonstrates that HTATSF1 functions as an HDF and can be suppressed to inhibit HIV-1 replication without cytotoxicity. However, there are a number of limitations associated with validating the anti-HIV therapeutic potential of targeting cellular factors with RNAi in reporter cell lines such as TZM-bl cells. These include a potential bias for factors influencing early stages of the viral life cycle. In addition, TZM-bl cells are not derived from cells that are the natural hosts of HIV, which may alter the kinetics of HIV-1 replication and the relative import of HDFs since the expression of genes varies between cell lines (see 1.3.5). This is especially important when the redundancy in cellular processes is considered, as it may be possible for HIV-1 to adapt to another cofactor in one cell line but not another. The large number of putative HDFs identified in HIV-1 host factor screens, which display little overlap with one another but enrichment for proteins involved in certain biological processes (Goff 2008; Bushman *et al.* 2009), may reflect host factor redundancy (see 1.3.4). In addition, putative HDFs identified in two genome-wide screens are enriched in immune cells (Brass *et al.* 2008; Konig *et al.* 2008), suggesting that these cells are especially proficient in the functions HIV needs for optimal replication. It is important, therefore, to determine the ability of HTATSF1 suppression to inhibit HIV-1 replication under conditions that resemble those in the host as closely as possible.

Since T cell-derived lines are difficult to transfect, a lentivirus may be employed to deliver shRNA expression cassettes into these cells. Lentiviruses are a subclass of retroviruses and so integrate their genomes into the host cell DNA allowing long-term expression of inserted therapeutic genes. This ensures maximal, sustained target suppression is obtained, even of proteins with long half-lives, such as SMARCB1 and PSIP1. Lentiviruses are capable of transducing both dividing and non-dividing cells, but since native viruses can cause disease, non-replicating viruses are used in transgene expression. The safest forms are self-inactivating (SIN) vectors where the U3 region of the LTR is deleted, and a heterologous promoter (eg CMV or EF1-α) ensures transcription of the entire vector mRNA, including the inserted expression cassette. This stable shRNA expression permits examination of the impact of cellular factor suppression on HIV replication kinetics over an extended time course.

In this chapter, the ability of HTATSF1 suppression to mediate prolonged suppression of HIV-1 replication in a T cell-derived line was determined. This was achieved by transducing cells with lentiviruses carrying shRNA expression cassettes. To exclude changes in proliferation rate as the source of variation in viral replication, shRNA-expressing cell lines were examined for their proliferative capacity. Whether the inhibition of viral replication on HTATSF1 suppression was a consequence of aberrant splicing of viral transcripts was also investigated.

3.2 Materials and methods

3.2.1 Mammalian cell culture

Cells of the HEK293T cell line, HEK293 cells transformed by the large T antigen from Simian vacuolating virus 40 (SV40) that inactivates retinoblastoma protein (pRb), were maintained in DMEM (BioWhittaker, MD, USA) supplemented with 10% heat-inactivated FCS (Delta Bioproducts, Johannesburg, SA) at 37 °C and 5% CO_2 in a Forma Series II 3110 water-jacketed incubator (Thermo Fisher Scientific, Inc., MA, USA).

SupT1 cells, a non-Hodgkin's T cell lymphoma cell line expressing high levels of surface CD4, were maintained in the same media as HEK293T cells. SupT1 cells grow in suspension and, therefore, serial passage required centrifugation of cells at 130 x *g*.

To determine the correct number of cells for seeding, 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).

3.2.2 Selection of lentivector for shRNA expression cassettes

Lentiviruses derived from two, second generation lentivectors, pLVTH (Addgene plasmid 12262; deposited by D. Trono; Addgene, Inc., MA, USA) (Wiznerowicz and Trono 2003) and pLVCTH (a gift from D. Trono), were analysed for their ability to transduce SupT1 cells and express their GFP reporter genes, driven by the EF1-α and CMV promoters, respectively. Five hundred microlitres of lentivirus generated by D. Ivacik were incubated for 2 hours with 5 x 10^4 SupT1 cells per well in a 12-well plate in triplicate. Cells were then collected by centrifugation at 130 x *g*, resuspended in 1 ml media and returned to the culture plate for a further 48 hours.

Cells were harvested by centrifugation at 130 x *g* and resuspended in 200 μl PBS containing 4% FCS. Samples were run on a FACSCalibur™ flow cytometer (BD Biosciences, NJ, USA) and 5 x 10^4 events per sample acquired using CellQuest Pro™ software (BD Biosciences, NJ, USA).

Analysis was performed using FlowJo 9.1 (Tree Star, Inc., OR, USA). SupT1 cells were gated based on forward and side scatter and GFP+ cells determined from that subset by comparing SupT1 cells transduced with no virus (GFP-) and those transudced with lentivirus (GFP+). All gates were consistent throughout samples.

3.2.3 Cloning lentivector-expressed shRNAs

To generate shRNA-expressing lentivectors, shRNA expression cassettes were removed from pTZ plasmids and cloned into pLVTH. Five micrograms of pTZU6+shRNA plasmids were digested with 8.3 U *Eco*RI and 21.7 U *Acc*I (Fermentas, WI, USA) for 3 hours at 37 °C in a reaction volume of 30 μl containing 2x Tango Buffer (Fermentas, WI, USA).

The pLVTH backbone was prepared by removal of the H1 promoter from within the LTR/SIN. Firstly, 3 μg of pLVTH was digested with 5 U *Cla*I (Fermentas, WI, USA) for one and a half hours at 37 °C in a reaction volume of 15 μl containing 1x Tango Buffer. This was followed by a second digestion with 5 U *Eco*RI in a volume of 20 μl containing 2x Tango buffer for one and a half hours at 37 °C. Finally, 5 U Antarctic phosphatase (New England Biolabs®, Inc., MA, USA) and Antarctic phosphatase buffer (New England Biolabs®, Inc., MA, USA) were added to a final volume of 25 μl. The dephosphorylation reactions were incubated for 1 hour at 37 °C, followed by enzyme inactivation at 65 °C for 5 min.

Digestion products from both U6+shRNA and pLVTH reactions were resolved on a 2% agarose gel (see A1.2.3) and the desired fragment excised and purified using the MinElute™ Gel Extraction Kit (QIAGEN, CA, USA) (see A1.2.4). Forty femtomoles of the dephosphorylated backbone was used in 1 hour, room temperature ligation reactions containing 90 fmol of U6+shRNA insert and 5 U T4 DNA ligase (Fermentas, WI, USA) in a volume of 15 μl. A mock pLVTH construct, containing the U6 promoter but no shRNA sequence, was also generated. Transformation, colony screening and bulk vector preparation were as described in 2.2.2, with the exception that colonies were screened by *Bam*H1 (Fermentas, WI, USA) digestion.

3.2.4 Dual luciferase reporter assay

To verify shRNA expression from pLVTH, a dual luciferase reporter assay was performed, as described in 2.2.6, using 1.2 x 10^5 HEK293T cells per well. psiCheck™-2 vectors with targets for the control shRNAs, shHBVx, and shU5 and shTAT, were provided by A. Ely and S. Barichievy, respectively.

3.2.5 Generation of lentiviruses

To generate lentiviruses from the shRNA-expressing lentivectors, 3.6 x 10^6 HEK293T cells were seeded in a 60 cm^2 Costar® dishes (Corning Inc, NY, USA) and each dish transfected 24 hours

later with 5 μg pLVTH-U6+shRNA, 3.8 μg psPAX2 (Addgene plasmid 12260; deposited by D. Trono; Addgene, Inc., MA, USA) and 2.5 μg pMD2.G (Addgene plasmid 12259; deposited by D. Trono; Addgene, Inc., MA, USA) using the calcium chloride method (see A1.2.5). psPAX2 provided Gag, Pol, Tat and Rev, which are needed for replication, and pMD2.G determined tropism of the resultant lentiviruses by providing Env, in this case VSV-G, which permits transduction of SupT1 cells (Figure 3.1). Both 24 and 48 hours post-transfection, the culture media (10 ml) was collected, pooled, 0.22 μm-filtered and stored at -80 °C in 0.5 ml aliquots.

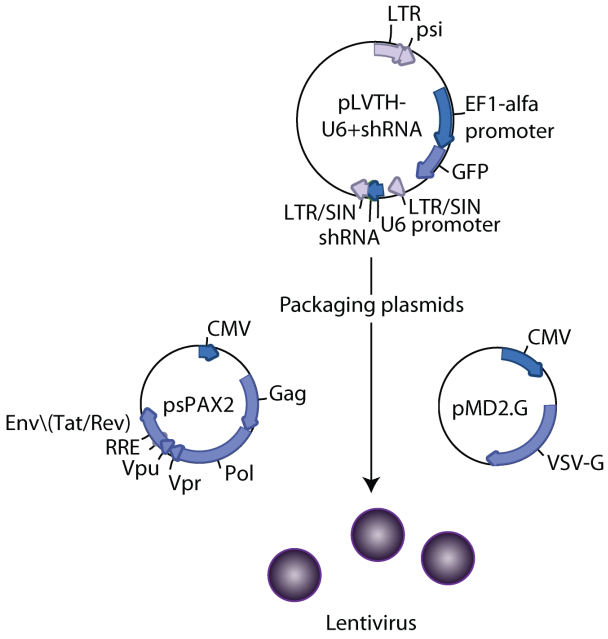


Figure 3.1: Schematic showing plasmids used to generate lentiviruses capable of delivering shRNA expression cassettes to SupT1 cells. HEK293T cells were transfected with pLVTH-U6+shRNA plasmids along with packaging plasmids psPAX2 and pMD2.G. psPAX2 provides proteins required for a single-round of replication and pMD2.G provides the envelope protein VSV-G, which enables the resultant lentivirus to transduce SupT1 cells.

3.2.6 Biological transduction titre of shRNA-expressing lentiviruses

Lentiviruses were titred based on the number of GFP+ SupT1 cells following transduction. SupT1 cells were seeded at 1 x 10^4 cells per well in a 24-well plate in a volume of 100 μl. Four hundred microlitres of 10^0, 10^{-1} and 10^{-2} lentivirus dilutions were added to the cells in duplicate and incubated for 24 hours. Cells were then harvested by centrifugation at 130 x *g*, resuspended in 1 ml media and returned to the culture plate for another 48 hours. Seventy-two hours post-infection, cells were harvested by centrifugation, run on a FACSCalibur™ (BD Biosciences, NJ, USA) and the proportion of GFP+ cells determined as described in 3.2.2, with the exception that 5 x 10^3 events were acquired per sample. Lentivirus titres were determined by applying sigmoidal dose-

response non-linear regression analyses using GraphPad Prism 4.0c (GraphPad Software, Inc., CA, USA).

3.2.7 Generation of shRNA-expressing SupT1 cell lines

SupT1 cells were seeded at 3 x 10^5 cells in a 75 cm^2 culture flasks (Nunc™, Thermo Fisher Scientific, Inc., MA, USA) and incubated with a volume of lentivirus calculated to transduce 20% of SupT1 cells, based on the biological transduction titre. This was to minimise the probability of there being more than one integrated shRNA expression cassette per cell. Cells were cultured for 5 days to allow expansion prior to harvesting and resuspension in PBS with 4% FCS to a concentration of ~1 x 10^6 cells/ml. Cells were sorted on a FACSCalibur™ with sorting gates set based on the GFP expression of both lentivirus-transduced (GFP+) and untransduced (GFP-) SupT1 cells. The sorted GFP+ cells were concentrated by centrifugation at 130 x *g* and cultured in DMEM with 20% FCS, 100 U/ml penicillin, 100 μg/ml streptomycin, 50 μg/ml tetracycline, 100 μg/ml ampicillin, 170 μg/ml chloramphenicol, 50 μg/ml kanamycin and 100 μg/ml ciprofloxacin for 1 week. Sorted cell lines were then cultured for a further week without antibiotics to expand the populations. Stocks were made and stored in liquid nitrogen so that low-passage cells could be used in all subsequent experiments. The proportion of GFP+ SupT1 cells was determined in each cell line by flow cytometry as described in 3.2.6, immediately prior to sorting (pre-sort) and freezing (post-sort).

3.2.8 Detection of processed shRNA guide strands using PAGE and Northern blot analysis

Total RNA was isolated from SupT1 cell lines using TriReagent™ (Sigma-Aldrich, MO, USA) (see A1.3.1). Thirty micrograms of RNA, as well as a ladder composed of 5 pmol of shHBVx (5' CCG TGT GCA CTT CGC TTC 3') and U6 (Table 2.5) probes 3'-end labelled with digoxigenin (DIG)-ddUTP, were combined with an equal volume of RNA loading dye (Ambion, TX, USA) and denatured at 80 °C for 5 min. Samples were resolved on a 15% polyacrylamide gel, visualised on a UV transilluminator and transferred and cross-linked to a nylon membrane as described in 2.2.9.

DNA oligonucleotide probes complementary to the processed shRNA guide strands (Table 2.5) were labelled using the DIG Oligonucleotide 3'-end Labelling Kit (Roche, Mannheim, Germany). Briefly, 100 pmol of oligonucleotide probe was incubated with 4 μl reaction buffer (1 M potassium cacodylate, 0.125 M Tris.HCl, 1.25 mg/ml bovine serum albumin, pH 6.6), 5 mM $CoCl_2$, 0.05 mM DIG-ddUTP and 20 U terminal transferase in a volume of 20 μl for 15 min at 37 °C. The labelling reaction was stopped with addition of 2 μl of 0.2 M EDTA, pH 8.0. Immediately prior

to hybridisation, 5 pmol of DIG-labelled probe was denatured at 95 °C for 5 min. Probes were then added to membranes pre-hybridised with 10 ml DIG Easy Hyb solution (Roche, Mannheim, Germany) in a rotating oven (Thermo Fisher Scientific, Inc., MA, USA). Hybridisation was permitted overnight at 42 °C.

Membranes were subject to a low stringency wash in 50 ml 5x SSC (750 mM NaCl, 65 mM disodium citrate, pH 7.0) and 0.1% sodium dodecylsulphate (SDS) (Sigma-Aldrich, MO, USA) for 20 min at room temperature. Two high stringency washes were then performed at 42 °C, each with 50 ml 1x SSC (150 mM NaCl, 13 mM disodium citrate, pH 7.0) and 0.1% SDS for 15 min. Membranes were then rinsed in 100 ml wash buffer (0.1 M maleic acid, 0.15 M NaCl, pH 7.5, 0.03% Tween 20) and blocked for 30 min at room temperature in 100 ml of 5% skimmed milk powder in PBS. DIG-labelled oligonucleotides were detected by incubation of the membranes with 1 μl alkaline phosphatase-conjugated anti-DIG antibody in 20 ml PBS with 1% skimmed milk powder for 30 min at room temperature. Membranes were then washed 4 times for 15 min each with wash buffer and incubated in detection buffer (0.1 M Tris.HCl, 0.1 M NaCl, pH 9.5) for 5 min, all at room temperature. Chemiluminescence detection of labelled probes was enabled by incubation of membranes with CDP-*Star*™ (Roche, Mannheim, Germany) for 5 min and image acquisition with a G-BOX (Syngene, MD, USA) using an exposure of up to 1 hour.

When necessary, membranes were stripped by rinsing in distilled water, washing twice with stripping buffer (0.2 M NaOH and 0.1% SDS) for 15 min each at 37 °C and finally rinsing the membrane in 2x SSC for 5 min at room temperature.

3.2.9 SupT1 cell proliferation rates by cell counting

To ascertain whether the different shRNA-expressing SupT1 cell lines differed in their proliferation rates, SupT1 cells were seeded at 5 x 10^4 cells per well in a 12-well plate in duplicate. Following seven days of culture, cells were harvested and counted as described in 3.2.1.

3.2.10 SupT1 cell proliferation rates by DNA quantification

To further verify the proliferation rates of shRNA-expressing SupT1 cell lines, cells were seeded and cultured as described in 3.2.9, in duplicate. Following harvesting, cells were incubated with 1 ml DNA extraction buffer (0.15 M NaCl, 20 mM Tris.HCl, 10 mM EDTA and 0.1% SDS) and 400 μg/ml proteinase K (Roche, Mannheim, Germany) at 37 °C overnight. DNA was extracted by addition of 1 ml phenol:chloroform:isoamyl alcohol (IAA) in a 25:24:1 ratio and centrifugation at 4 °C, with the resulting aqueous layer stored in 15μl 3 M NaOAc and 300 μl 100% ethanol overnight

at -80 °C. DNA precipitates were resuspended in 100 μl water and concentrations determined by NanoDrop (Thermo Fisher Scientific, Inc., MA, USA).

3.2.11 Assessment of p81A-4 replication in shRNA-expressing SupT1 cell lines

To assess the ability of SupT1 cell lines expressing shRNAs to inhibit HIV-1 replication, 2 x 10^4 SupT1 cells were seeded per well in a round-bottomed 96-well plate and some immediately infected with the subtype B molecular clone p81A-4 (NIH AIDS Research & Reference Reagent Program, MD, USA) (Chesebro *et al.* 1992; Walter *et al.* 2005) at a $TCID_{50}$ (see A1.4.5) of 50/ml in duplicate. Twenty-four hours post-infection, plates were centrifuged at 130 x *g* and media removed. Cells were washed three times with PBS to remove residual virus before resuspension in 350 μl media. Following another centrifugation to pellet cells, 150 μl of media was removed and stored as day 0 samples at -80 °C. 48 hours later (day 2), cells were similarly collected and 150 μl culture supernatant removed and stored. Cells were resuspended with the addition of 150 μl media to replace that removed. Samples were taken in the same fashion on days 4, 7, 10, 14 and 17 (ie 5, 8, 11, 15 and 18 days post-infection, respectively). HIV-1 p24 content in the frozen samples was subsequently determined by ELISA (Murex Biotech Ltd, Kent, UK) (see A1.4.4). Dilutions of the positive control within the kit were used to generate a standard curve of p24 levels from which absolute levels of p24 in the experimental samples were determined.

3.2.12 qRT-PCR measurement of mRNA concentrations

To examine the duration of host factor mRNA suppression mediated by shRNAs in the SupT1 cells, total RNA was extracted from cells cultured for different periods. Day 0, 10 and 20 refer to time points in the p81A-4 infection assay (see 3.2.11) and these samples were obtained from RNA extractions performed on SupT1 cells cultured for the equivalent periods. This approximates to 7, 17 and 27 days post-thaw of the shRNA-expressing SupT1 cell line stocks.

Isolated RNA was treated with 3 U RQ1 RNase-free DNase (Promega, WI, USA) in a reaction volume of 40 μl containing 4 mM Tris.Cl, pH 8.0, 10 mM $MgSO_4$ and 1 mM $CaCl_2$, at 37 °C for 30 min. The reaction was stopped with addition of EGTA, pH 8.0, to a final concentration of 2 mM and incubation at 65 °C for 10 min. Approximately 50 ng of DNase-treated RNA, 1.5 μg of random primers (Invitrogen™, Life Technologies, CA, USA) and 5 nmol of dNTPs (Fermentas, WI, USA) were incubated at 65 °C for 5 min and then placed on ice. Reverse transcription was carried out using 100 U SuperScript® III RT, 50 mM Tris.HCl, pH 8.3, 75 mM KCl, 30 mM $MgCl_2$, 5 mM DTT

and 20 U RNaseOUT™ (Invitrogen™, Life Technologies, CA, USA) in a volume of 10 μl. Reactions were incubated at 25 °C for 5 min, 55 °C for 45 min and finally 70 °C for 15 min.

qPCR and analysis was performed as described in 2.2.7, with the exception that the qPCR reactions utilised the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4).

3.2.13 Detection of HTATSF1 by SDS-PAGE and Western blot analysis

To examine the duration of shH1-mediated HTATSF1 suppression, protein extracted from SupT1 cells, cultured for periods equivalent to days 0 and 20 of the p81A-4 infection assay (see 3.2.11), were subject to SDS-PAGE, Western blot and analysis, as described in 2.2.11. A notable difference is that 100 μg of cellular protein was resolved on a 12% polyacrylamide gel. Day 20 samples were prepared in duplicate.

3.2.14 Quantification of GFP+ SupT1 cells in shRNA-expressing cell lines by flow cytometry

To assess whether the proportion of GFP+ cells within the shRNA-expressing populations was maintained over time, samples of the shRNA-expressing SupT1 cell lines were subject to flow cytometry as described in 3.2.6, at time points equivalent to days 0 and 20 of the p81A-4 infection assay (see 3.2.11).

3.2.15 Nuclear run-on assay to determine *htatsf1* transcription rates in SupT1 cell lines

Transcription rates of the *htatsf1* gene were examined in SupT1 cells expressing either shH1 or shU5 following periods of cell culture equivalent to days 0 and 20 of the p81A-4 infection assay (see 3.2.11). SupT1 cells (4×10^6) were harvested and washed with PBS. Cell pellets were resuspended in 1.5 ml NP-40 lysis buffer (10mM Tris.Cl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% NP-40) and kept on ice for 5 min. Samples were then centrifuged at 400 x *g* for 5 min at 4 °C. Following removal of the supernatant, the pellets were resuspended in 0.75 ml NP-40 lysis buffer and immediately centrifuged at 400 x *g* for 5 min at 4 °C. Following removal of the supernatant, the nuclei pellets were resuspended in 100 μl glycerol storage buffer (0.1mM EDTA, 40% glycerol, 5 mM $MgCl_2$, 50 mM Tris.Cl, pH 8.3) by gentle vortexing. Nuclei were flash-frozen by transfer to liquid nitrogen and stored for up to one month.

Thawed nuclei were subject to *in vitro* transcription reactions following addition of 100 μl transcription reaction buffer (10 mM Tris.HCl, pH 8.0, 5 mM $MgCl_2$, 0.3 M KCl, 2.5 mM of ATP, CTP

and GTP (Roche, Mannheim, Germany), 0.25 mM biotin-16-UTP (Roche, Mannheim, Germany) and 100 U RNaseOUT™ (Invitrogen™, Life Technologies, CA, USA). Reactions were carried out at 30 °C for 45 min. Reactions were stopped by the addition of 50 U RQ1 RNase-free DNase (Promega, WI, USA) and incubation at 37 °C for 10 min. Three hundred microlitres of HSB buffer (0.5 M NaCl, 50 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM Tris.HCl, pH 7.4) was used to resuspend nuclei prior to incubation at 30 °C for 5 min. Nuclei were lysed with addition of 100 μl of SDS/Tris buffer (5% SDS, 0.5 M Tris.HCl, pH7.4, 0.125 M EDTA). Five microlitres of proteinase K (Roche, Mannheim, Germany) was also added and samples incubated for 30 min at 42 °C. RNA was extracted by addition of 650 μl TriReagent™ (Sigma-Aldrich, MO, USA) and 100 μl chloroform. Samples were centrifuged at 16,000 x *g* at 4 °C for 10 min and the resulting aqueous layers transferred to 300 μl ice cold 100 % ethanol, 100 μl 3 M NaOAc and 3 μl glycogen (Ambion, TX, USA), which were stored overnight at -80 °C. The following day, RNA isolation was completed by centrifugation at 16,000 x *g* at 4 °C for 10 min. RNA pellets were washed with 75% ethanol. Final RNA preparations were resuspended in 50 μl DEPC-treated water (Ambion, TX, USA).

RNA samples were treated with 5 U RQ1 RNase-free DNase (Promega, WI, USA) in a reaction volume of 60 μl containing 4 mM Tric.Cl, pH 8.0, 10 mM $MgSO_4$ and 1 mM $CaCl_2$, at 37 °C for 30 min. The reaction was stopped with addition of EGTA, pH 8.0, to a final concentration of 2 mM and incubation at 65 °C for 10 min. RNA concentration and purity was assessed on a NanoDrop (Thermo Fisher Scientific, Inc., MA, USA).

Dynabeads® MyOne™ Streptavidin C1 beads (Invitrogen™, Life Technologies, CA, USA), which bind biotinylated RNA, were prepared for RNA binding by removal of RNase enzymes. Firstly, a volume of beads equivalent to 30 μl per sample were washed twice with a volume of solution A (DEPC-treated 0.1 M NaOH) equal to the volume of beads. Throughout, beads were separated from solution using a magnetic separator for 1-2 min. Beads were then washed twice with solution B (DEPC-treated 0.05 M NaCl) and three times with 2x bind & wash (B&W) buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA and 2 M NaCl). Beads were finally resuspended to a final concentration of 5 μg/ml (twice original volume) in 2x B&W buffer. Twelve micrograms of RNA at 200 ng/μl were added to an equal volume of beads (ie 60 μl sample + 60 μl beads) and incubated at room temperature on a gently rotating RM-2M Intelli-mixer (ELMI Ltd., Riga, Latvia) for 15 min. RNA-bound beads were washed three times with 1x B&W buffer before resuspension in 20 μl 1x B&W buffer.

Reverse transcription reactions using 1.5 μl of RNA-beads per reaction were carried out as described in 3.2.12. 1 U of RNase H (Invitrogen™, Life Technologies, CA, USA) was added to the samples to degrade bead-bound RNA and facilitate dissociation of cDNA. Reactions were

incubated at 37 °C for 20 min, followed by 95 °C for 5 min to inactivate the enzyme. cDNA was removed from the beads and used in qPCR, which was performed and analysed as described in 3.2.12.

3.2.16 Quantification of p81A-4 transcript classes by qRT-PCR

To examine whether HTATSF1 suppression affects levels of different classes of HIV-1 transcripts, SupT1 cells were seeded at 5 x 10^5 per well in a 6-well plate and infected with p81A-4 at a $TCID_{50}$ of 50/ml. Cells were harvested and washed with PBS prior to extraction of total RNA 48 hours later. RNA was DNase-treated, reverse transcribed and the resulting cDNA used in qPCR as described in 3.2.12 (see also A1.3.4).

Table 3.1: Primers used in qPCR to determine levels of HIV-1 transcript classes.

Primer	Sequence (5'-3')	Amplicon(s)
9 kb forward	TAGCAGTCCTCTATTGTGTGCA	Unspliced
9 kb reverse	GGAGGTTCTGCACTATAGGGTA	
9 + 4 kb forward	GCTGTTAAATGGCAGTCTAGCA	Unspliced & singly spliced
9 + 4 kb reverse	TAAAATGCTCTGCCTGGTCCTA	
2 kb forward	GAAGAGCTCATCAGAACAGTCA	Multiply spliced
2 kb reverse	AGATCGTCCCAGATAAGTGCTA	
Total forward	GGGTGGGAGCAGTATCTCGA	Unspliced, singly spliced & multiply spliced
Total reverse	TCTACAGCTGCCTTGTAAGTCA	

The qPCR primers were designed to amplify different classes of HIV-1 transcripts (Table 3.1). The 9 kb primers amplified a region only present in unspliced viral transcripts, the 4 + 9 kb primers a region present in both unspliced and singly spliced viral transcripts and the 2 kb primers span the Tat/Rev splice junction such that amplification only occurs from multiply spliced viral transcripts. In addition, primers were used to amplify a region present in all classes of viral transcripts. Levels of 4 kb, or singly spliced, transcripts were determined by subtracting the level of 9 kb transcripts from the expression level obtained with the 9 + 4 kb primers. Total viral transcripts were quantified by both summing the expression level of 9, 4 and 2 kb transcripts (sum total) and the value obtained by qPCR with the primers that amplify a region common to all HIV-1 transcripts (common total). Expression of viral transcripts is reported relative to β-actin mRNA (see Table 2.4 for primers). In addition, levels of *htatsf1* mRNA were verified (see Table 2.4 for primers).

3.2.17 Statistical analyses

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 $p < 0.05$. Data was analysed using non-linear regression, one-way ANOVAs, followed by a Dunnett's multiple comparison post-test, and two-way ANOVAs, followed by Bonferroni post-tests, where appropriate, calculated using GraphPad Prism 4.0c (GraphPad Software, Inc., CA, USA).

3.3 Results

3.3.1 Selection of the lentivector pLVTH for shRNA expression

To further examine the therapeutic potential of targeting HTATSF1 and SMARCB1, CD4+ and CCR5+ T cell-derived lines were generated that stably expressed U6-driven shRNAs. SupT1 cells, a non-Hodgkin's T cell lymphoma cell line expressing high levels of surface CD4, were used as these are readily infectable with HIV-1. Moreover, sustained inhibition of viral replication has been reported in SupT1 cells on stable expression of shRNAs targeting the integration cofactor PSIP1 (Llano *et al.* 2006a), which was included in this study as a positive control. Before cloning shRNA expression cassettes into a lentivector, lentiviruses LVTH and LVCTH, derived from lentivectors pLVTH and pLVCTH, respectively, were first tested for their ability to transduce and express their GFP reporter genes in SupT1 cells. The heterologous promoter driving expression of mRNA differed between the two vectors – EF1-α in pLVTH and CMV in pLVCTH (Figure 3.2A).

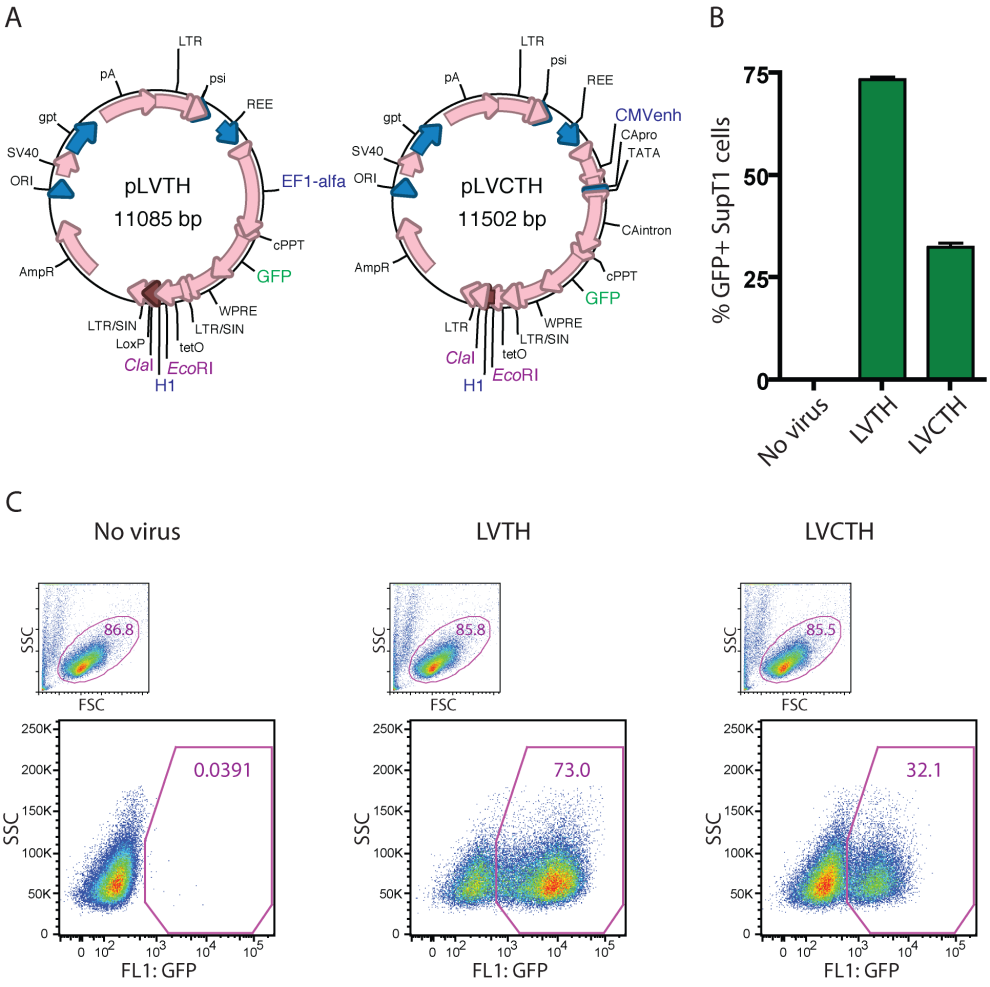


Figure 3.2: SupT1 GFP reporter gene expression is greatest when transduced with pLVTH-derived lentivirus. A. Two different lentivectors, pLVTH and pLVCTH, were packaged in HEK293T cells. **B.** SupT1 cells were transduced with the resulting lentiviruses, LVTH and LVCTH, in triplicate and the proportion of GFP+ SupT1 cells was determined by flow

cytometry 48 hours later. **C.** Representative flow cytometry plots showing the gating strategy. SupT1 cells were first selected (top-left plots) and then GFP+ cells determined from the SupT1 subset. Gates were consistent for all samples.

Following transduction of SupT1 cells, cell populations were examined by flow cytometry for GFP expression. A greater proportion of cells were GFP+ when transduced with LVTH compared with LVCTH (Figure 3.2B and C). As both lentiviruses were prepared under the same conditions, this suggests that either pLVTH-derived viruses generate better titres than those derived from pLVCTH, or expression of the GFP reporter is more efficient from pLVTH than pLVCTH. Regardless, pLVTH was selected as the lentivector into which shRNA expression cassettes were cloned.

3.3.2 shRNAs knock down cognate targets when expressed from a lentivector

shRNA expression cassettes were cloned into *Cla*I and *Eco*RI sites within the LTR/SIN of pLVTH, such that the H1 promoter was replaced (Figure 3.2A). Insertion within the LTR/SIN ensured duplication of the U6-shRNA expression cassette in the provirus following reverse transcription. shRNA expression from the lentivector was verified by quantitation of cognate target inhibition with a dual luciferase reporter assay in HEK293T cells. All shRNAs significantly inhibited their target sequences, as determined by the ratio of *Renilla* to firefly luciferase activity, compared with controls (>85%, p <0.001; Figure 3.3). This demonstrates that transfer of shRNA expression cassettes to a new plasmid backbone did not disrupt their expression and efficacy.

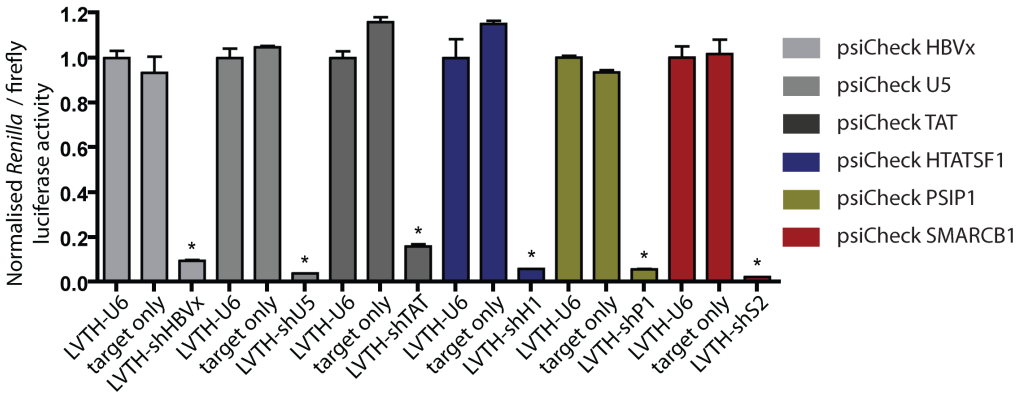
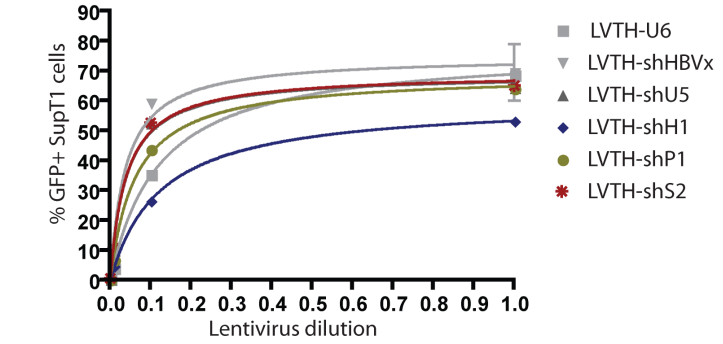


Figure 3.3: shRNAs knock down their cognate targets when expressed from a lentivector. HEK293T cells were transfected with pLVTH-U6+shRNA expression cassettes and cognate psiCheck reporter constructs resulting in a significant inhibition of *Renilla* luciferase activity relative to firefly luciferase activity 48 hours later. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

3.3.3 Generation of stable shRNA-expressing SupT1 cell populations

Lentivectors were packaged in HEK293T cells by cotransfection with two separate plasmids encoding the structural and regulatory genes *in trans*, and the *env* gene, which allowed pseudotyping with VSV-G and tropism for SupT1 cells. This approach excludes the risk of recreating replication competent wild type-like viruses by chance. The resulting lentiviruses, carrying the shRNA expression cassettes, were titred based on the proportion of GFP expression in transduced SupT1 cells (Figure 3.4A). Very low viral titres for LVTH-shTAT were obtained (data not shown), which is probably due to the silencing of *tat* mRNA in packaging cells. Given that another virus-targeting positive control, LVTH-shU5, was successfully packaged, shTAT was no longer used in the study.

A



B

$$Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{(1 + 10^{(\log(EC50 - X))})}$$

$$20 = 0 + \frac{(100 - 0)}{(1 + 10^{(\log(EC50 - X))})}$$

$$\log X = \log EC50 - \log 4$$

$$X = 10^{(\log EC50 - \log 4)}$$

Sample	EC50	Dilution for 20% GFP
LVTH-U6	0.2685	0.0671
LVTH-shHBVx	0.0898	0.0225
LVTH-shU5	0.1285	0.0321
LVTH-shH1	0.9052	0.2263
LVTH-shP1	0.2094	0.0525
LVTH-shS2	0.1267	0.0317

Figure 3.4: Biological transduction titre of shRNA-expressing lentiviruses. **A.** SupT1 cells were transduced with lentiviruses carrying shRNA expression cassettes and a GFP reporter gene. The proportion of GFP+ SupT1 cells was determined by flow cytometry 72 hours later. **B.** The virus dilution required to give 50% GFP+ SupT1 cells (EC50) was determined by sigmoidal dose-response non-linear regression analysis, where Top = 100 and Bottom = 0. Rearrangement of the non-linear equation permitted calculation of the dilution of virus needed to obtain 20% GFP+ SupT1 cells on transduction.

Titration curve equations, derived from non-linear regression, were used to calculate the amount of virus to give GFP expression in 50% (EC50) and 20% of cells (Figure 3.4B). This latter virus dilution was used to transduce SupT1 cells resulting in populations of SupT1 cells with 15-20% GFP expression (Figure 3.5A). This degree of transduction was selected so as to reduce the probability of more than one provirus being integrated per cell. Multiple shRNA expression cassette integration events may be associated with toxicities related to detrimental integration events or saturation of the RNAi pathway from high shRNA expression (Grimm *et al.* 2006; Grimm

et al. 2010). The transduced cell populations were sorted for GFP expression by flow cytometry, resulting in the generation of SupT1 cell lines with >90% GFP expression (Figure 3.5A and B).

A

Sample	pre-sort	post-sort
LVTH-U6	16.0	95.3
LVTH-shHBVx	15.2	93.6
LVTH-shU5	16.9	91.9
LVTH-shH1	17.7	92.9
LVTH-shP1	18.0	95.4

B

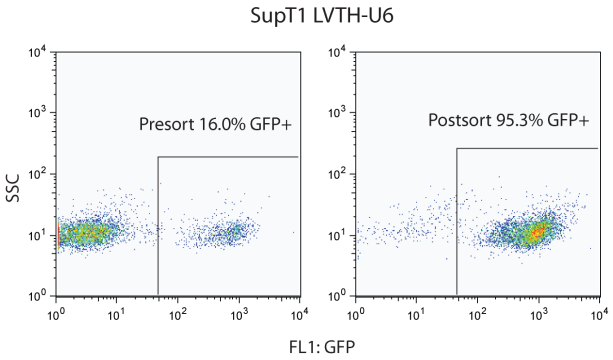


Figure 3.5: Sorting of lentivirus-transduced SupT1 cell lines. **A.** The proportion of GFP+ SupT1 cells was determined by flow cytometry following transduction with lentiviruses both pre- and post-sort. **B.** Representative flow cytometry plots demonstrating the cell sorting gating strategy. Gates were consistent across all samples.

LVTH-shS2 SupT1 cells proliferated at a noticeably reduced rate compared to the other cell lines and, following expansion post-sorting, the proportion of GFP+ cells was only 60% (data not shown). This suggested that the shS2-expressing cells were at a growth disadvantage to the shRNA-negative cells in the population. For this reason, this cell line was not used in the study and so whether SMARCB1 is an HDF in SupT1 cells was not established.

shRNA expression and guide strand processing was confirmed in SupT1 cells transduced with LVTH-shH1 and LVTH-shP1 by small RNA Northern blot. As expected, the shH1- and shP1-derived guide strands were detectable in SupT1 cells transduced with LVTH-shH1 and -shP1, respectively (Figure 3.6). This demonstrates that specific shRNAs are being expressed in the transduced and sorted SupT1 cell populations and that no cross-contamination occurred.

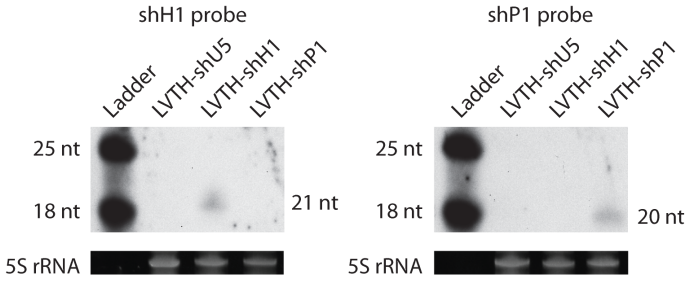


Figure 3.6: Processing of HDF-targeting shRNAs in SupT1 cells. Guide strands were detected by small RNA PAGE Northern blot. Equal loading of RNA was verified by levels of 5S rRNAs, which were detected by ethidium bromide-staining of the polyacrylamide gel.

3.3.4 Proliferative capacity is unaltered by shRNA expression in SupT1 cells

Different cell proliferation rates may affect HIV-1 replication kinetics because replication is influenced by the extent of cell activation and the amount of cells present. Having observed a reduction in proliferation rates on stable expression of the shRNA targeting SMARCB1, the other shRNA-expressing SupT1 cell lines were examined. The number of cells present following seven days culture did not significantly differ between the cell lines (p >0.05), whether quantified by counting of viable cells within a Trypan blue-stained population (Figure 3.7A), or spectrophotometry of isolated DNA (Figure 3.7B). This confirms that any differences in HIV-1 replication observed between the shRNA-expressing SupT1 cell populations will not be a result of a difference in the cells available for HIV-1 to infect.

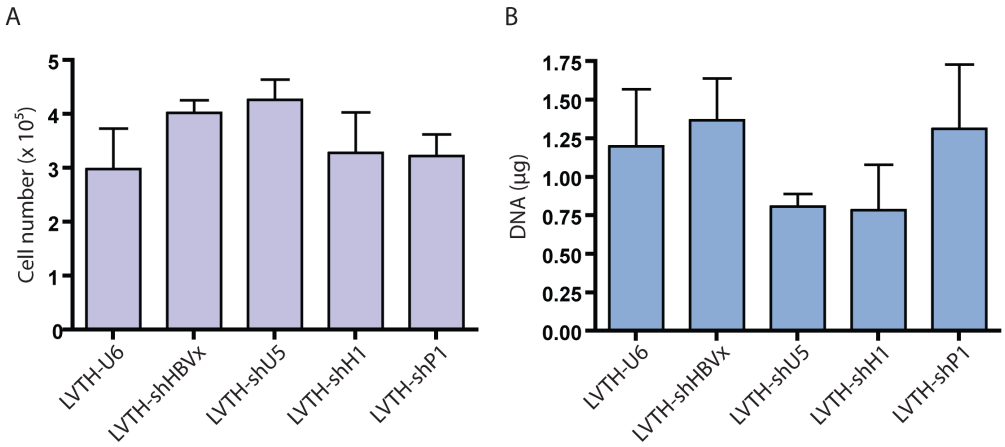


Figure 3.7: SupT1 cell line proliferation rates do not significantly differ. shRNA-expressing SupT1 cell lines were cultured for seven days prior to (A) counting following a Trypan blue stain, and (B) quantification of isolated DNA. One-way ANOVAs.

3.3.5 A shRNA targeting HTATSF1 mediates sustained inhibition of HIV-1 replication

The kinetics of subtype B molecular clone p81A-4 replication was examined in the SupT1 cells expressing shRNAs. LVTH-U6 cells were also cultured in the absence of virus as a negative control. 24 hours post-infection, cells were washed to remove any residual virus used in the infection (Day 0). Levels of the HIV-1 antigen p24 were quantified in culture supernatants at various time points up to 17 days post-infection.

Compared with SupT1 cells with no shRNA expression, SupT1 cells expressing shH1 significantly inhibited p24 production for more than 14 days post-infection (>95% on Day 10, p <0.001), and to a similar extent to that obtained with the positive control, shU5, which targets the virus directly (75% on Day 10, p <0.01; Figure 3.8A and B). This demonstrates that shRNA-mediated suppression of HTATSF1 expression in a T cell-derived line can inhibit HIV-1 replication. The extent of p24 inhibition with shH1 or shU5 expression was not as great as that observed in

the shP1-expressing cell line, which was negligible throughout the time course. This is in accordance with previous studies (Llano *et al.* 2006a) and may relate to the pre-integrative role of p75. p24 levels 17 days post-infection in the cell lines with no shRNA expression (LVTH-U6), or expression of the negative control, shHBVx, are not shown as these populations exhibited extensive cell death by this time point, which was observed by microscopy.

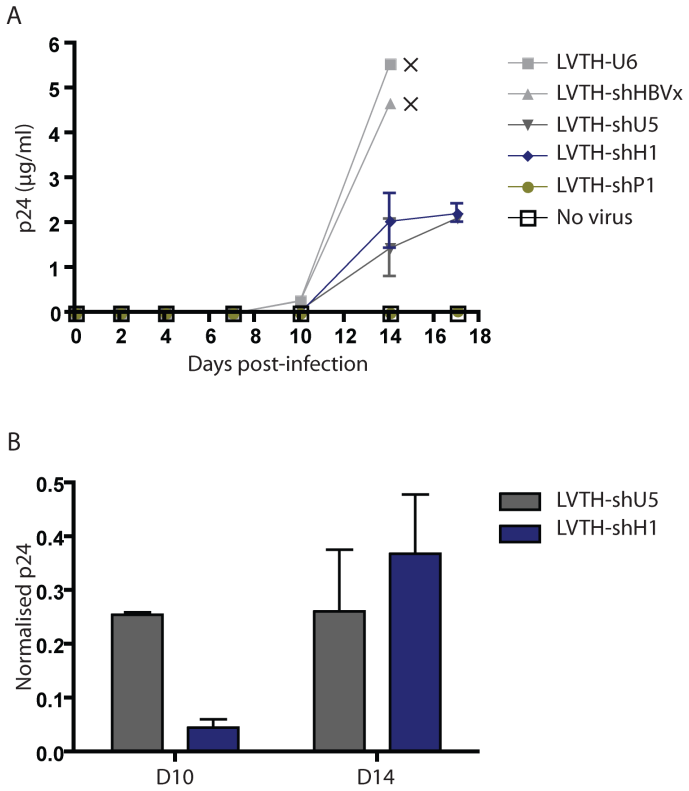


Figure 3.8: HTATSF1-targeting shRNA inhibits p81A-4 replication. **A.** Following infection of SupT1 cells with subtype B molecular clone p81A-4 at a $TCID_{50}$ of 50/ml, p24 levels were determined in culture supernatants by ELISA up to 17 days post-infection. Crosses indicate discontinued p24 measurement, which was a result of cell death. **B.** As (A) with p24 levels 10 and 14 days post-infection normalised to pLVTH-U6.

Interestingly, p24 levels in shH1-expressing SupT1 cells, relative to the pLVTH-U6 control, were increased 14 days post-infection compared with 10 days post-infection (Figure 3.8B). In contrast, the suppression of p24 levels in shU5-expressing cells was maintained. Several mechanisms may account for the slight rebound in p24 levels: adaptation of the virus to another cofactor, increased HTATSF1 as a result of increased HTATSF1 transcription and increased HTATSF1 as a result of decreased shRNA expression. The first possibility is unlikely given the time frame used in the assay. Also, escape from shU5 would be expected to occur more rapidly than cofactor adaptation, particularly as shH1 and shU5 mediate replication inhibition to a similar degree. The other two possibilities were investigated.

3.3.6 HTATSF1 levels increase through reduced shH1 expression

Firstly, whether HTATSF1 levels were increasing over time was investigated. Fresh stocks of SupT1 cell lines were raised and cultured for the same duration as used in the p81A-4 challenge assay. At time points equivalent to days 0, 10 and 20 post-infection, total RNA was isolated and subject to qRT-PCR. Levels of *htatsf1* mRNA were suppressed at each time point in the shH1-expressing cell line compared to the control LVTH-U6 (Figure 3.9). However, there was a significant increase in *htatsf1* mRNA levels at day 20 (~65%) from day 10 (~49%, p <0.01; Figure 3.9). In contrast, suppression of *psip1* mRNA was sustained in the shP1-expressing cell line throughout the time course, demonstrating that the reduction in RNAi over time was specific to the shH1-expressing cell line.

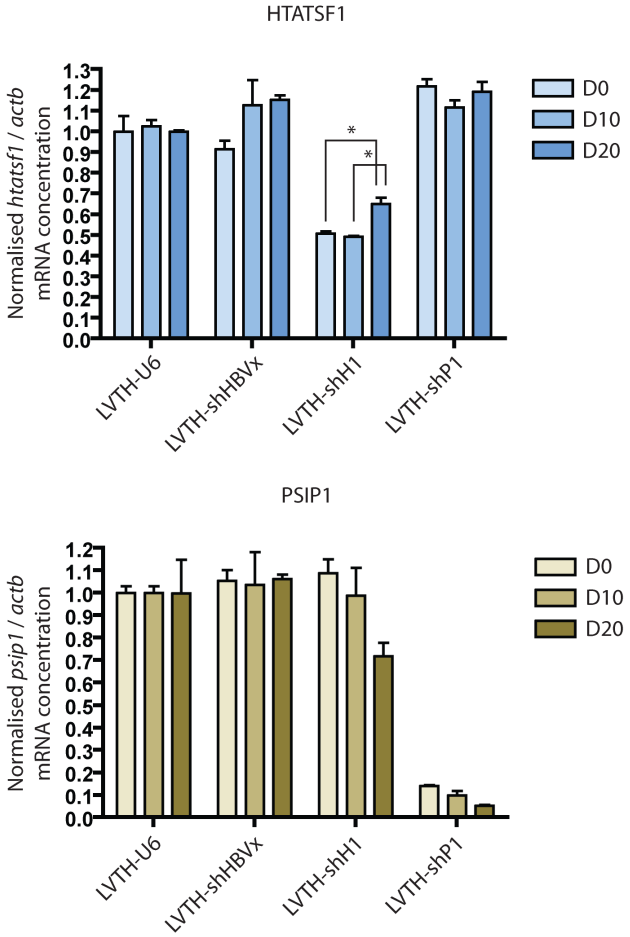


Figure 3.9: HTATSF1 mRNA levels increase in shH1-expressing SupT1 cells over time. HTATSF1 or PSIP1 mRNA levels relative to β-actin mRNA (*actb*) were determined by qRT-PCR from total RNA extracted from shH1- or shP1-expressing SupT1 cells, respectively. RNA was extracted at time points equivalent to days 0, 10 and 20 of the p81A-4 challenge assay. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

The increase in *htatsf1* mRNA on day 20 was accompanied by an increase in HTATSF1 protein in the shH1-expressing cell lines from ~31% of LVTH-U6 control on day 0, to ~60% on day

20 (Figure 3.10). This suggests that the slight attenuation of p24 production inhibition over time in cells expressing shH1 may have been a result of increased HTATSF1 expression.

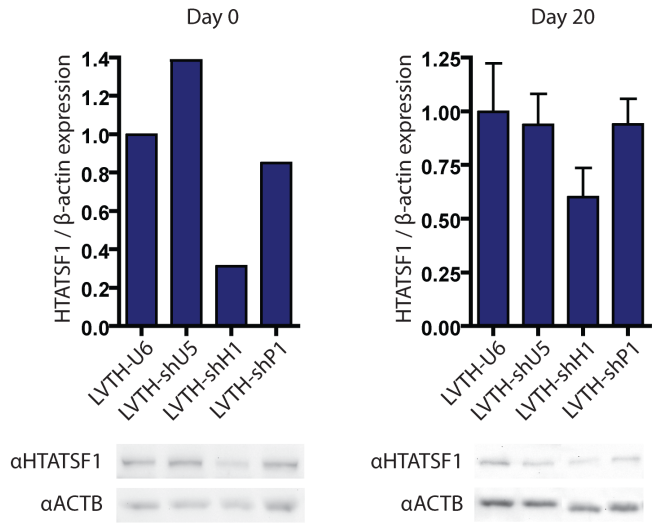


Figure 3.10: HTATSF1 increases in shH1-expressing SupT1 cell lines over time. HTATSF1 relative to β-actin was determined in SupT1 cell lines after different periods of passage. One hundred micrograms of total cellular protein was subject to PAGE and Western blot. Day 20 samples were prepared in duplicate and, as such, representative bands for this time point are shown.

Several experiments were performed to elucidate the mechanism of an increase in HTATSF1 expression over time. Flow cytometry of shRNA-expressing SupT1 cell populations revealed that the increase in HTATSF1 was accompanied by a drop of ~3% in the proportion of GFP+ cells in the shH1-expressing population (Figure 3.11), which was not observed in the shP1-expressing population. However, only a single population were cultured and analysed by flow cytometry for each cell line, so the statistical relevance of these results is unclear.

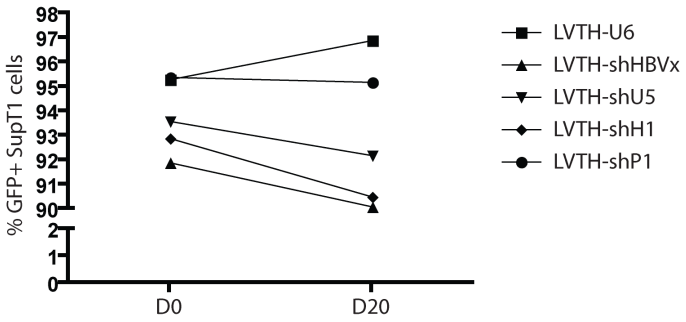


Figure 3.11: The proportion of GFP+ SupT1 cells in the shH1-expressing cell line decreases over time. SupT1 cell lines were analysed by flow cytometry for the percentage of GFP+ cells over time.

Transcription rates of *htatsf1* were determined by nuclear run-on assay performed on nuclei isolated from SupT1 cells expressing either shH1 or shU5 cultured for different time periods. As expected, no change in *htatsf1* transcription rates were observed in the cells

expressing shU5 between day 0 and day 20 (Figure 3.12A). Similarly, *htatsf1* transcription rates were unaltered in shH1-expressing cells over time, suggesting no regulatory mechanism in the cell for the preservation of HTATSF1 expression operating at the level of transcription. In contrast, small RNA Northern blot revealed a drop of 68% in shH1-derived guide strand expression relative to 5S RNA from day 0 to day 20 (Figure 3.12B). This was much greater than the observed reduction in shP1-derived guide strand expression (23%; Figure 3.12B). Together the data suggests that the attenuation of p24 inhibition in shH1-expressing SupT1 cells is a result of decreased shH1 expression and, consequently, increased HTATSF1 over time. The cause of decreased shH1 expression may be a result of epigenetic silencing of the integrated expression cassette or a growth advantage of the cells within the population without shH1 expression.

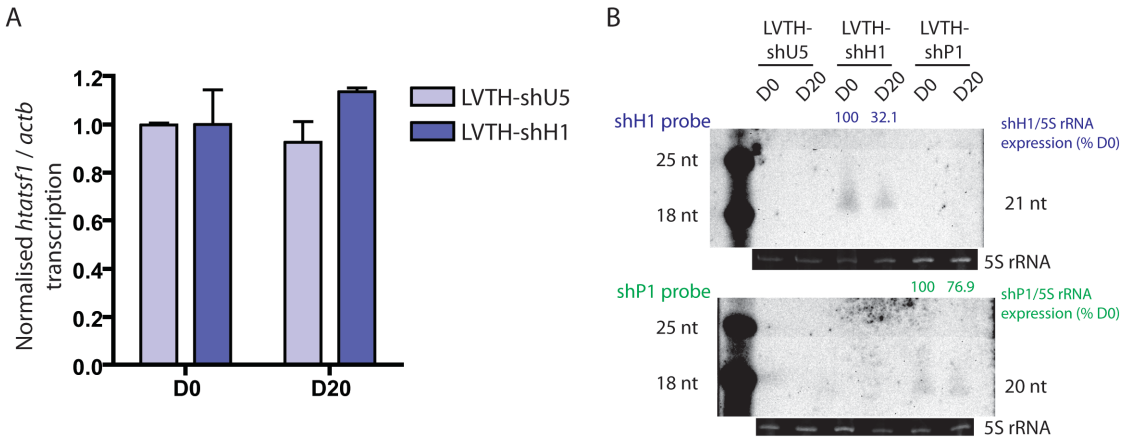


Figure 3.12: Suppression of HTATSF1 does not alter *htatsf1* transcription but shH1 expression decreases over time. **A.** SupT1 cell lines were analysed for transcription of the *htatsf1* gene by nuclear run-on assay. Two-way ANOVA. **B.** Guide strands were detected by small RNA PAGE Northern blot. Levels of guide strands are given relative to 5S rRNAs, which were detected by ethidium bromide-staining of the polyacrylamide gel.

3.3.7 HTATSF1 suppression does not alter the ratio of HIV-1 transcript classes in SupT1 cells

One group has solely attributed the HIV-1 cofactor role of HTATSF1 to enhancement of viral transcript splicing (Miller *et al.* 2009). To investigate whether the inhibition of viral replication observed in SupT1 cells expressing shH1 (Figure 3.8) was a result of defects in splicing, various primer sets were designed to amplify different classes of HIV-1 transcripts (Figure 3.13). One set uniquely amplified ~9 kb, or unspliced, transcripts. A second set amplified both ~4 kb, or singly spliced, and ~9 kb transcripts. A third set of primers spanned the Tat/Rev splice junction and so only amplified ~2 kb, or multiply spliced, transcripts. For the ~4 kb class, expression was calculated by subtracting the expression of unspliced transcripts from that obtained with the second primer set that quantified both unspliced and singly spliced transcripts. Total p81A-4

transcript expression was quantified by summing expression for each of the three classes, and also using a fourth primer set that amplified a region common to all three classes.

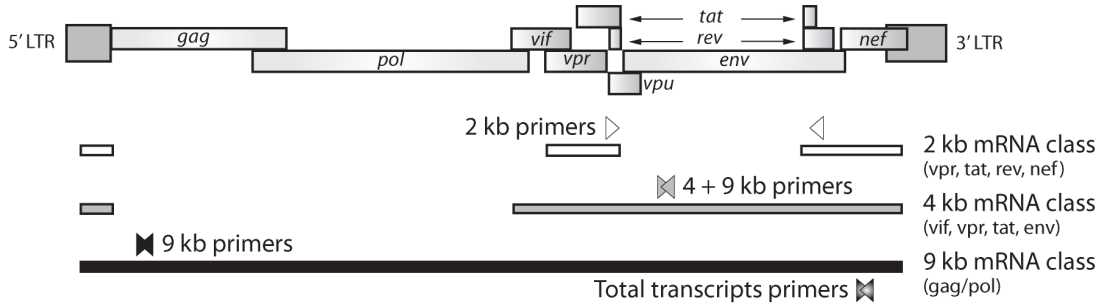


Figure 3.13: Schematic of HIV-1 transcript classes and primers used to quantify them. HIV-1 transcripts fall into three classes: ~9 kb class, which are unspliced, ~4 kb class, often referred to as singly spliced and ~2 kb class, which are multiply spliced. Primers were designed to amplify the 9 or 2 kb classes only, the 4 and 9 kb class and all HIV-1 transcripts.

Following infection of shRNA-expressing SupT1 cell lines with p81A-4, total RNA was extracted from washed cells 48 hours later and analysed by qRT-PCR. Firstly, suppression of *htatsf1* mRNA was confirmed in the shH1-expressing cell line (~60%, p <0.01; Figure 3.14A). The same cDNA samples were then analysed for expression of the different classes of HIV-1 transcripts, relative to *actb* mRNA. As expected, shP1-expressing SupT1 cells displayed reduced expression of all three classes, consistent with its role in viral integration (Figure 3.14B). Unspliced transcripts were increased in the shU5-expressing cell line, relative to the LVTH-U6 control, suggesting that the target site may be less accessible in unspliced transcripts compared with spliced. Surprisingly, an increase in all three classes of HIV-1 transcripts was observed in the shH1-expressing cell line compared to LVTH-U6 cells (Figure 3.14B).

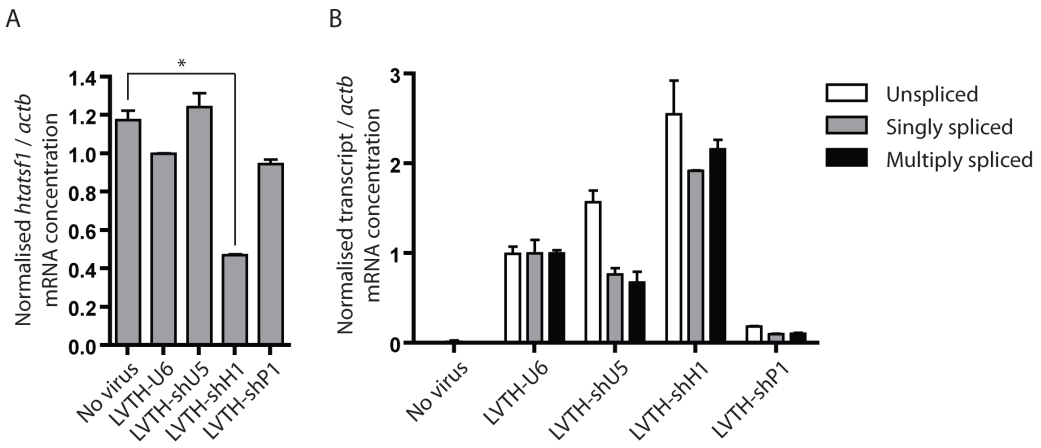


Figure 3.14: All classes of p81A-4 transcripts are increased in shH1-expressing SupT1 cells. SupT1 cell lines were infected with p81A-4 at a $TCID_{50}$ of 50/ml and total RNA extracted 48 hours later. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to uninfected SupT1 cells. **A.** Levels of HTATSF1 mRNA relative to β-actin mRNA (*actb*) were determined by qRT-PCR. **B.** The same RNA was also used to quantify p81A-4 transcripts.

The general increase in HIV-1 transcripts in shH1-expressing cells was not accompanied by a change in the relative ratios of the three transcript classes, however (p >0.05; Figure 3.15A), and this was the case for all SupT1 cell lines. In addition, the total number of p81A-4 transcripts was similar whether determined by summing the expression for the three classes or using the primers common to all classes (Figure 3.15B), which suggests that the observed results are unlikely to be a result of significant differences in the amplification efficiency of the various primer sets.

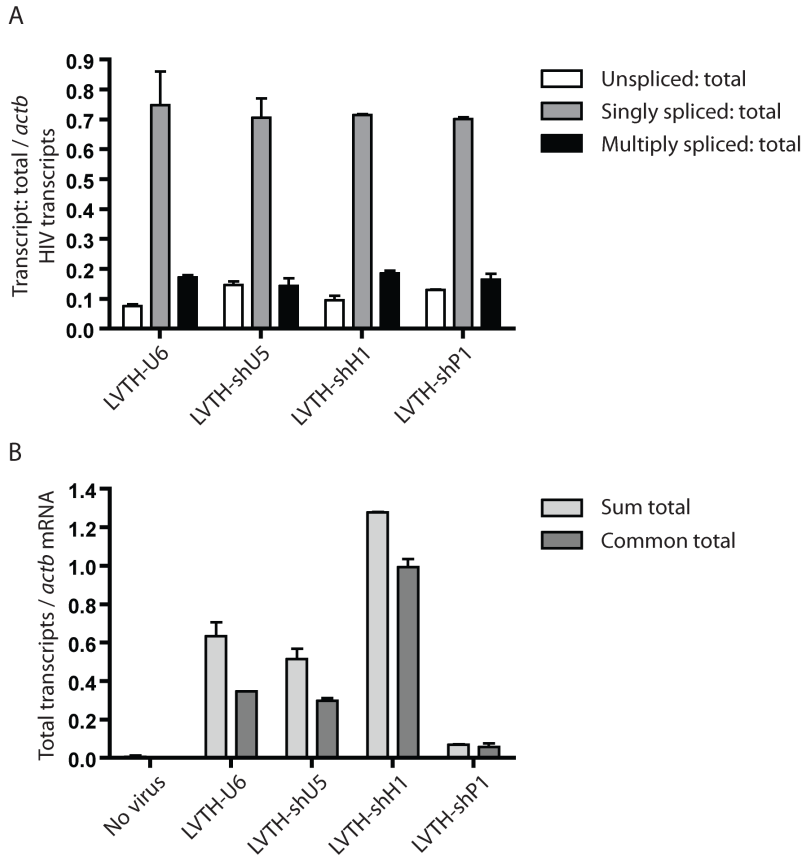


Figure 3.15: HTATSF1 suppression does not alter the ratio of p81A-4 transcript classes. SupT1 cell lines were infected with p81A-4 at a $TCID_{50}$ of 50/ml and total RNA extracted 48 hours later. **A.** The ratio of different classes of p81A-4 transcripts to total viral transcripts relative to β-actin mRNA (*actb*) were determined by qRT-PCR. Total viral transcripts are given by summing the different classes. **B.** Total viral transcripts relative to β-actin mRNA (*actb*), determined both by summing each class (sum total) and qRT-PCR using primers which amplify a region common to all classes of transcripts (common total).

This data provides no evidence of a role for HTATSF1 in splicing of viral transcripts in SupT1 cells. Rather, HTATSF1 suppression leads to an accumulation of all classes of viral transcripts two days post-infection. Nevertheless, the inhibition of viral replication over 14 days demonstrates that the increase in p81A-4 transcripts is not translated into enhanced infectious particle production.

3.4 Discussion

The approach used in this study was to validate the role of cellular factors HTATSF1 and SMARCB1 in HIV-1 replication through shRNA-mediated PTGS in a T cell line. The ability of SMARCB1 suppression to inhibit p81A-4 replication in SupT1 cells was not determined as cells expressing a shRNA targeting SMARCB1 displayed diminished proliferative capacity. This is seemingly in contrast to the role of SMARCB1 as a tumour suppressor in rhabdoid cells, which occurs by regulating the cell cycle through the pRb pathway [reviewed in (Stojanova and Penn 2009)]. Mutations in *smarcb1* have been found in malignant rhabdoid tumours, a rare yet aggressive paediatric cancer (Versteege *et al.* 1998; Biegel *et al.* 1999). Indeed, a survey of 100 primary rhabdoid and central nervous system tumours found that 75% of these carried genetic alteration of the *smarcb1* locus (DeCristofaro *et al.* 1999). However, SMARCB1 regulates approximately 5-10% of genes (Gresh *et al.* 2005) and, recently, it was demonstrated that deficiency in SWI/SNF chromatin remodelling complexes, of which SMARCB1 is a core component, results in aberrant chromatin organisation and mitotic failure (Bourgo *et al.* 2009). In addition, inactivation of SMARCB1 in murine embryonic fibroblasts led to growth suppression triggered by p53 and p21 (Isakoff *et al.* 2005) and SMARCB1 loss is associated with polyploidy and chromosome instability (Vries *et al.* 2005), but not disruptions in the DNA damage response (McKenna *et al.* 2008). Knudson's 2-hit hypothesis for oncogenesis, which requires the occurrence of at least two random somatic alterations (Knudson 1971), may account for the discrepancy in enhanced cell proliferation between malignant rhabdoid tumours and shS2-expressing SupT1 cells. Of course, whether SupT1 cells, which are derived from a leukaemia, have functioning p53, p21 or DNA damage response is unknown and so it is unclear by what mechanism SMARCB1 suppression inhibits proliferation or induces apoptosis in these cells. Admittedly, the decrease in proliferation may have been a result of an off-target effect of the shRNA, not mediated by SMARCB1 suppression. Nevertheless, the association of SMARCB1 loss with cancer suggests that, regardless of whether this protein is a HDF or not, it is not a good anti-HIV therapeutic target.

Suppression of HTATSF1 inhibited HIV-1 replication for over 14 days in SupT1 cells (Figure 3.8). However, the major cofactor role HTATSF1 plays in p81A-4 replication in SupT1 cells is unclear. Previously, HTATSF1 was shown to enhance splicing of HIV-1 transcripts, with shRNA-mediated suppression of HTATSF1 leading to an increase in unspliced, relative to total, HIV-1 transcripts (Miller *et al.* 2009). However, the work presented in this thesis suggests that HTATSF1 suppression has no effect on splicing of viral transcripts (Figure 3.15A) and rather leads to an accumulation of all classes of HIV-1 transcripts (Figure 3.14B).

Miller and colleagues used a very different approach to this study. Firstly, levels of the three different HIV-1 transcript classes were assessed by Northern blot on RNA samples extracted from anti-HTATSF1 shRNA-expressing T-Rex-293 cells 48 hours post-transfection with the subtype B HIV-1 plasmid pSG3ΔEnv. Similar to the findings here (Figure 3.15A), HTATSF1 depletion lead to an increase in unspliced and singly spliced transcripts. However, this was accompanied by a decrease in fully spliced transcripts, although this was not obvious from the blot as abundance was quantified relative to fuzzy control bands, particularly for the fully spliced class. In addition, transcript class abundance was not normalised to a housekeeping gene, despite GAPDH expression clearly being greater in the samples expressing a HTATSF1-targeting shRNA. They went on to analyse the same RNA samples by qRT-PCR, showing an increase in the ratio of unspliced to total HIV-1 transcripts. However, again expression of neither were normalised to a housekeeping gene. Also, it is unclear how the qPCR data was processed because the ratio of unspliced to total HIV-1 transcripts was greater than one and the increase in unspliced transcripts was greater (>1.5-fold) than when quantified from Northern blot bands (~1.2-fold), despite the two analyses being performed on the same samples. In summary, critical analysis of the data presented in Miller *et al*. raises doubt over their conclusion that HTATSF1 regulates the relative levels of unspliced and spliced HIV-1 transcripts. Most importantly, Miller and colleagues offer no explanation as to why singly spliced transcripts are unaffected by HTATSF1 suppression.

In contrast to Miller *et al*., this study used SupT1 cells and cells were infected, not transfected, with HIV-1. This alone may account for the differences between the two studies. For example, HTATSF1 may be a splicing cofactor but this may not be its rate-limiting role in SupT1 cells. Alternatively, transfecting plasmid versus infecting with virus will lead to differences in the kinetics of replication and it is possible that the time point at which the effect on splicing is most discernible was not used in this study. Indeed, even Miller and colleagues comment that the effect on splicing was only observed 48 hours post-transfection. Moreover, a recent study demonstrated a role for HTATSF1 in promoting a transition from fully spliced to singly spliced transcripts in the presence of Tat (Jablonski *et al.* 2010), indicating that any role of HTATSF1 in splicing is complex and may change during the course of the viral life cycle.

Of interest is the accumulation of all classes of viral transcripts 48 hours post-infection on HTATSF1 suppression (Figure 3.14B). This was surprising given proposed roles for HTATSF1 in transcription elongation and splicing. Conversely, the increase in transcripts did not reflect increased HIV-1 replication, as p24 production was reduced for over 14 days compared with LVTH-U6 and LVTH-shHBVx controls (Figure 3.8). This suggests a post-transcription role for HTATSF1, necessary for the formation of infectious particles, not yet described.

HTATSF1 has been implicated in the stabilisation of a number of nuclear RNA-protein complexes, including spliceosomal U snRNPs (Yan *et al.* 1998; Fong and Zhou 2001), the RNA Pol II elongation complex (Zhou and Sharp 1996; Li and Green 1998; Parada and Roeder 1999; Fong and Zhou 2001; Zhou *et al.* 2004a) and influenza vRNPs (Naito *et al.* 2007). Perhaps, therefore, HTATSF1 functions as a more general enhancer of the formation of nuclear RNA-protein complexes. This is certainly compatible with the structure of HTATSF1, which is composed of a highly acidic carboxyl-terminal half and an amino-terminal half containing two RNA recognition motifs (RRMs) (Zhou and Sharp 1996). Nucleic acid-binding proteins contain basic domains, although these tend to aggregate in the absence of nucleic acid or molecular chaperones. The acidic domain of HTATSF1 may bind such basic regions, serving as a chaperone. HTATSF1 may also simultaneously bind RNA, through the amino-terminal RRMs, and basic proteins, through the carboxyl-terminal acidic domain, stabilising RNA-protein complexes. Thus, a hypothesis is that HTATSF1 is involved in export of viral transcripts in SupT1 cells, either through chaperoning of the RNA-binding proteins required (eg Rev), or stabilisation of the exported RNA-protein complexes themselves. If this were the case, suppression of HTATSF1 would result in accumulation of viral transcripts without increasing replication, as observed here. Further experiments, such as RNA fluorescence *in situ* hybridisation (FISH), are required to test this hypothesis.

The magnitude of the inhibition of p24 production relative to LVTH-U6 control was not maintained in the shH1-expressing SupT1 cell lines (Figure 3.8), which is probably a result of increased HTATSF1 expression over time (Figures 3.9 and 3.10). The increase in HTATSF1 suppression was not a result of increased *htatsf1* transcription (Figure 3.12A) but likely a result of decreased shH1 guide strand expression (Figure 3.12B). Although an effect on shRNA processing cannot be excluded with the data reported here, the lack of a role for HTATSF1 in RNAi renders it more likely that the decrease in shH1-derived guide strands is a result of a decrease in the expression of the shRNA. This is unlikely to be through an HTATSF1 suppression-mediated effect on transcription elongation, as HTATSF1 has only been identified in RNA Pol II, not RNA Pol III, elongation complexes (Kim *et al.* 1999; Parada and Roeder 1999; Chen *et al.* 2009).

Two mechanisms may account for the decreased shRNA expression: (1) epigenetic silencing of the shRNA expression cassette; and, (2) decreased growth of the shRNA-positive cells within the population relative to the shRNA-negative cells. Although epigenetic silencing of integrated viral sequences is acknowledged as a limitation of gene therapy, the former is less likely to account for decreased shRNA expression because RNA Pol III U6 promoters are more resistant to epigenetic silencing than RNA Pol II promoters [reviewed in (Zhang and Rossi 2010)], although the possibility cannot be excluded with the data here.

For normal cells to have a growth advantage over HTATSF1-suppressed cells seems plausible, particularly when the role of HTATSF1 in transcription elongation is considered. Indeed, stable HTATSF1 knockdown has been suggested to cause growth arrest in HeLa cells (Chen *et al.* 2009). However, under these circumstances, decreased GFP+ cells in the population should accompany decreased shRNA expression. The discord in the extent of change in GFP expression (Figure 3.11) and shH1 expression (Figure 3.12B) over 20 days may relate to the methods used to quantify the respective parameters – although flow cytometry is extremely sensitive, GFP expression was assessed in only one replicate. In addition, shH1 expression was quantified on a Northern blot relative to 5S RNA on an ethidium bromide-stained polyacrylamide gel, although transfer efficiency probably differed between the ~120 nt 5S RNA and ~21 nt guide strands. Furthermore, despite all experiments being performed on cell lines prepared at the same time, the experiments were conducted up to several months apart. Thus, differences in culture conditions between sampling for the qRT-PCR, Western blot, Northern blot, nuclear run-on, flow cytometry, cell counting and DNA isolation, cannot be excluded.

Of course, growth arrest is not a desirable side effect, particularly in immune cells. This brings into question the therapeutic potential of HTATSF1 targeting as an anti-HIV strategy. However, significant differences in the proliferation rates between the various SupT1 cell lines were not observed (Figure 3.7). Admittedly, the methods used (ie cell counting and quantification of isolated DNA) may have been insufficiently sensitive over the seven days examined and an assay based on the extent of BrdU incorporation would have been preferable. Nevertheless, the effect HTATSF1 suppression has on proliferation appears modest. Moreover, it has been demonstrated that cells with greater resistance to HIV-1 replication undergo preferential expansion *in vivo* (Swan *et al.* 2006). Therefore, the selective disadvantage accrued on HTATSF1 suppression may be outweighed *in vivo* by a selective advantage in the context of a HIV infection. Further experiments are needed to verify this is the case, but it suggests that targeting of HTATSF1 may be more suited to post-infection treatment rather than prophylaxis.

Even if HTATSF1 suppression does not confer a selective disadvantage *in vivo*, there is another concern with respect to using HTATSF1 as an anti-HIV therapeutic target: the inhibition of HIV-1 replication on HTATSF1 suppression was moderate, in contrast to the pronounced inhibition observed in LVTH-shP1 cells (Figure 3.8). This suggests that whilst HTATSF1 may increase the efficiency of viral replication, it is not a critical HDF in SupT1 cells.

In summary, HTATSF1 may have complex roles in HIV-1 replication. Its suppression can inhibit viral replication in T cell-derived lines but the effect is moderate and may be accompanied by a growth disadvantage. HTATSF1 is, therefore, a HDF but not necessarily a promising anti-HIV therapeutic target. This study demonstrates that a thorough interrogation of HIV-1 replication

kinetics and cytotoxicity in relevant cells is necessary to validate HDFs as therapeutic targets. As such, a similar approach to that employed here should be applied to other novel HDFs identified in genome-wide screens.

Given that this thesis validates HTATSF1 as an HDF, and a role for DDX3X in HIV-1 replication was demonstrated elsewhere (Ishaq *et al.* 2008), whether TGS of these genes would inhibit HIV-1 replication was examined. This was the objective of the work described in Chapter 4.

CHAPTER 4

Inhibition of HIV-1 replication through transcriptional gene silencing of HTATSF1

4.1 Introduction

TGS provides an alternative approach to PTGS for reducing gene expression (see 1.4). Like PTGS, Ago-bound small RNAs elicit gene silencing although, rather than targeting a mRNA, the guide strands are complementary to promoter-associated transcripts (Han *et al.* 2007) (see 1.4.2). A number of mechanisms have been proposed for the modulation of gene expression induced by promoter-targeting small RNAs (see 1.4.1-4). Classical TGS is accompanied by epigenetic changes at the targeted region, including histone modifications and/or DNA methylation, resulting in a more compact chromatin structure that is less conducive to transcription (see 1.4.1). It has been demonstrated that reduced gene expression may be sustained for over a month in the absence of the promoter-targeted small RNA, providing silent-state chromatin was achieved during the TGS induction (Hawkins *et al.* 2009). This highlights one of the principal advantages of classical TGS over PTGS – there is the potential for long-term silencing, and so the frequency of guide strand administration/expression may be reduced compared with PTGS, reducing toxicities associated with saturation of the RNAi pathway.

Host proteins involved in HIV-1 replication are attractive therapeutic targets given the high mutability of the virus (see 1.2.5). Chapters 2 and 3 of this thesis have validated HTATSF1 as a HDF and shown that its suppression inhibits HIV-1 replication in both reporter TZM-bl cells and T cell-derived lines. Although validation of the HDF DDX3X was not achieved in this study, concurrent work has confirmed its potential as a therapeutic target since its suppression inhibits HIV-1 replication without inducing apoptosis in HEK-based reporter cells (Ishaq *et al.* 2008). Given the advantages of TGS over PTGS, a proof of principle study was conducted to investigate the potential of TGS induction at HDFs to inhibit viral replication.

Obtaining TGS is sometimes hampered by the identification of target sites that correspond to promoter-associated transcripts but not mRNAs. Initial attempts to induce TGS at *htatsf1*, *ddx3x* and *psip1* in HEK293T cells failed (see A2). In this pilot study, several siRNAs reduced target expression but nuclear run-on analysis showed that the observed suppression was not at the transcriptional level. Subsequent interrogation of the siRNA target sites revealed that the database used to retrieve TSSs was not always accurate and a number of the siRNAs were in

fact targeting coding transcripts. Thus, difficulties identifying the most active TSS led to design of siRNAs that were not targeting the appropriate region in gene promoters.

Unlike *Saccharomyces cerevisiae* (David *et al*. 2006), human genes rarely have a defined TSS, with transcription initiation more commonly occurring over a region (Juven-Gershon *et al.* 2008). In addition, many genes have alternative promoters (Carninci *et al.* 2006; Jacox *et al.* 2010). Identification of the active promoter is critical, however, to the design of TGS-inducing small RNAs. The preferred method for identification of the most active TSS in a cell line for a specific gene is 5' rapid amplification of cDNA ends (RACE). However, in the absence of such data, promoters may be inferred from data obtained by the Functional Annotation of the Mammalian Genome (FANTOM) project (Kawaji *et al.* 2006b; Kawaji *et al.* 2009; Kawaji *et al.* 2010).

The FANTOM research project is an international collaboration to characterise the mammalian transcriptome. The project led to the development of cap analysis of gene expression (CAGE) technology (Maeda *et al.* 2008): through a combination of cap trapping, cleavage by restriction enzymes and deep sequencing, the 5' 27 nt sequence of capped mRNAs are obtained (Carninci *et al.* 2005; Carninci *et al.* 2006). These can be aligned to the genome providing a genome-wide map of TSSs.

During the latest (fourth) stage of the FANTOM project, the transcription regulatory network was examined by mapping TSSs in a differentiating macrophage-like cancer cell line, THP-1 (Kawaji *et al.* 2009). Interestingly, TSSs may be only a few base pairs apart with different expression patterns (Balwierz *et al.* 2009). To accommodate this when constructing a "promoterome", only nearby, coexpressed TSSs were grouped into transcription start clusters (TSCs). TSCs with overlapping proximal promoters were further clustered into transcription start regions (TSRs). TSCs and TSRs were annotated based on data from 56 human CAGE libraries, 18 of which were obtained from THP-1 samples. TSR data, including the most active TSS within a TSR, combined with enrichments of transcription factor binding sites, were compiled to provide a map of promoter regions (Kawaji *et al.* 2009).

This chapter examines the feasibility of transcriptional silencing of HDFs *htatsf1* and *ddx3x* as an anti-HIV strategy. The FANTOM project-derived TSRs and promoter regions were used as a proxy for target gene promoters in the HEK293T and TZM-bl cell lines. Algorithms designed to predict PTGS target sites were used in TGS target site selection, on the assumption that similar thermodynamic principles apply, as Ago proteins are involved in both small RNA-mediated gene-silencing pathways (Meister *et al.* 2004; Janowski *et al.* 2006; Kim *et al.* 2006). The mechanism of siRNA-mediated silencing was interrogated by quantification of promoter-associated transcripts and nuclear run-on analysis for the assessment of target gene transcription rates.

4.2 Materials and methods

4.2.1 Mapping the *htatsf1* promoter

A rational approach to the design of *htatsf1* (Entrez gene ID 27336) and *ddx3x* (Entrez gene ID 1654) promoter-targeting siRNAs was employed, which first involved mapping features of the promoters to the NCBI36/hg18 (Mar 2006) genome (Lander *et al.* 2001). The RefSeq transcription start sites of the *htatsf1* transcript NM_014500 (chrX:135,407,337 (+)) and *ddx3x* transcript NM_001356 (chrX:41,077,595 (+)) were obtained from the UCSC Genome Browser (http://genome.ucsc.edu/) (Kent *et al.* 2002).

A region 1000 bp either side of the RefSeq TSS was examined for the presence of various core promoter elements (Table 4.1 and Figure 4.1) (Juven-Gershon *et al.* 2008; Anish *et al.* 2009). Their presence was considered evidence that these sequences function in transcriptional regulation.

Table 4.1: Core promoter elements.

Motif	Consensus sequence	Position relative to TSS	Associated factor
Initiator (Inr)	YYANWYY	-2 to +5	TFIID
TATA box	TATAWAWR	-30/31 to -24/23	TBP (subunit of TFIID)
TFIIB recognition element, upstream (BRE^u)	SSRCGCC	Upstream of TATA box	TFIIB
TFIIB recognition element, downstream (BRE^d)	RTDKKKK	Immediately downstream of TATA box	TFIIB
Downstream core element (DCE) S_I	CTTC	~+8	Proximal to TAF1
Downstream core element (DCE) S_{II}	CTGT	~+18	
Downstream core element (DCE) S_{III}	AGC	~+31	
X core promoter element 1 (XCPE1)	DSGYGGRASM	-8 to +2	Unknown
X core promoter element 2 (XCPE2)	VCYCRTTRCMY	-9 to +2	

The region 1000 bp either side of the RefSeq TSS was examined for the presence of these core promoter elements. Degenerate nucleotides within consensus sequences are indicated according to the IUPAC nucleotide code.

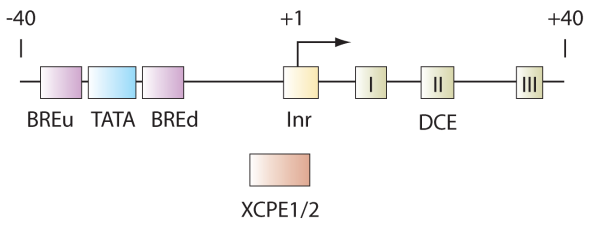


Figure 4.1: Schematic of the distribution of various core promoter elements. The region 1000 bp either side of the RefSeq TSS was examined for the presence of these core promoter elements.

Other promoter features were determined using data from the FANTOM project (Kawaji *et al.* 2006b; Kawaji *et al.* 2009; Kawaji *et al.* 2010). The locations of the *htatsf1* (TSR ID TSR_hg18_v1_chrX_+_135407060, located chrX:135407036-135407424 (+)) and *ddx3x* (TSR ID TSR_hg18_v1_chrX_+_41078359, located chrX:41078354-41078375 (+)) TSRs were obtained from FANTOM data available on the SwissRegulon Portal (http://www.swissregulon.unibas.ch/cgi-bin/regulon) (Pachkov *et al.* 2007). Also from this database were the *htatsf1* (cluster ID L2_chrX_+_135407060, located chrX:135407036-135407064 (+)) and *ddx3x* (cluster ID L2_chrX_+_41078359, located chrX:41078357-41078365 (+)) promoter regions and Motevo algorithm-predicted transcription factor binding sites (TFBS) (van Nimwegen 2007) retrieved. The *htatsf1* and *ddx3x* mapped FANTOM4 TSSs (chrX:135407060 (+) and chrX:41078359 (+)) may be defined as the most active or representative TSS, governed by the highest number of sequence tags within the TSR and promoter region.

4.2.2 Design of promoter-targeting siRNAs

A number of different tools were used in design of promoter-targeting siRNAs, although developed for the design of siRNAs targeting mRNAs. These online tools were Dharmacon's siDesign® Center (http://www.dharmacon.com/designcenter/designcenterpage.aspx), siRNA target finder (http://www.ambion.com/techlib/misc/siRNA_finder.html) and siDirect v2.0 (http://sidirect2.rnai.jp/doc/) (Naito et al. 2004; Naito et al. 2009). The algorithms used by all three tools include elements to favour functional siRNAs, based on efficacy predictions incorporated in other algorithms (Amarzguioui and Prydz 2004; Reynolds *et al.* 2004; Ui-Tei *et al.* 2004), and to reduce off-target effects, through elimination of siRNAs with near-perfect matches to mRNAs or high thermodynamic stability in the seed region (Ui-Tei *et al.* 2008).

To further reduce off-target effects, guide and passenger strand sequences were subject to nucleotide BLAST (blastn) against all human genomic and transcript sequences. The current standard reference assembly of the human genome, hg19/Genome Reference Consortium human genome build 37 (GRCh37) (May 2010), was used in alignments. Those siRNAs with ≥1 mismatches in the seed sequence or ≥3 mismatches overall to non-targeted regions were eliminated. In addition, G nucleotides were avoided in the 2-nucleotide overhang to prevent RNase activity on the siRNAs. Finally, siRNAs were selected that did not fully overlap TFBS and that spanned the *htatsf1* TSR and promoter region (Table 4.2).

Table 4.2: Oligonucleotide templates used to generate *htatsf1* and *ddx3x* promoter-targeting siRNAs.

siRNA	Antisense/guide oligonucleotide template	Sense/passenger oligonucleotide template	Design tool	Target		
				chrX: (+)	GC content (%)	Length (bp)
siH13	*ATGCCGGCGTCGGCGCTTTTTGA*CCTGTCTC	ATTCAAAAAGCGCCGACGCCGGCCCTGTCTC	siDirect	135407008-135407030	72	23
siH14	*AATGGCGGCTGCGTCGGCCTG*CCTGTCTC	AACAGGCCGACGCAGCCGCCACCTGTCTC	Ambion	135407080-135407100	71	21
siH15	*CCTGAGCAGGGCTTAGTTTAG*CCTGTCTC	CCCTAAACTAAGCCCTGCTCACCTGTCTC	Dharmacon*	135407097-135407117	52	21
siH16	*CAGGGCTTAGTTTAGAAGTAA*CCTGTCTC	CATTACTTCTAAACTAAGCCCCCTGTCTC	Dharmacon*	135407103-135407123	38	21
siH17	*AAGTAATTTCCTGACGTTGCC*CCTGTCTC	AAGGCAACGTCAGGAAATTACCCTGTCTC	Ambion	135407118-135407138	43	21
siH18	*CAGCGCCCTGGACCTAGCGGC*CCTGTCTC	CAGCCGCTAGGTCCAGGGCGCCCTGTCTC	None	135407216-135407236	76	21
siH19	*AAGCCAGGTGGGCGGCGGCGC*CCTGTCTC	AAGCGCCGCCGCCCACCTGGCCCTGTCTC	Ambion	135407258-135407278	81	21
siH20	*AAGAGCTCACAGCATCAGCGCG*CCTGTCTC	AACGCGCTGATGCTGTGAGCTCCCTGTCTC	Ambion	135407281-135407302	59	22
siH21	*AAGGGGCGGTGCAGCCGCCGAG*CCTGTCTC	AACTCGGCGGCTGCACCGCCCCCCTGTCTC	Ambion	135407326-135407347	77	22
siD13	*AACTCCGAGGCTGAGACTAGG*CCTGTCTC	AACCTAGTCTCAGCCTCGGAGCCTGTCTC	Ambion	41078078-41078098	57	21
siD14	*TAGCTCGAGAGAACTGGGACA*CCTGTCTC	TATGTCCCAGTTCTCTCGAGCCCTGTCTC	Dharmacon	41078124-41078144	52	21
siD15	*CTGCTAGGGGCGACAGGACTA*CCTGTCTC	CTTAGTCCTGTCGCCCCTAGCCCTGTCTC	Dharmacon	41078175-41078195	62	21
siD16	*CCGCGACAGGGAATTGCGGTG*CCTGTCTC	CCCACCGCAATTCCCTGTCGCCCTGTCTC	Dharmacon	41078254-41078274	67	21
siD17	*ACGTGCTGACGTAGCCGGCTT*CCTGTCTC	ACAAGCCGGCTACGTCAGCACCCTGTCTC	Dharmacon	41078297-41078307	62	21

All siRNA oligonucleotide templates had the CCTGTCTC sequence at their 3' end to allow hybridisation with the T7 promoter primer within the *Silencer*® siRNA Construction Kit (Ambion, TX, USA). Target sequences are in italics. Two nt overhangs are underlined. * denotes target sequences slightly modified from that suggested by the design tool to allow for two nt overhangs in the siRNAs without G nucleotides.

4.2.3 Generation of siRNAs by *in vitro* transcription

siRNAs were constructed using the *Silencer*® siRNA Construction Kit (Ambion, TX, USA), which involves *in vitro* transcription using T7 RNA polymerase (Figure 4.2). Oligonucleotide templates were designed that carried an 8-nt leader sequence at the 3' end to allow hybridisation with the T7 promoter primer (Table 4.2). Two microlitres of 100 μM template oligonucleotide were incubated with 2 μl T7 promoter primer (concentration unknown) and 6 μl DNA hyb buffer (proprietary information) at 70 °C for 5 min and then room temperature for a further 5 min. To this reaction, 2 μl of 10x Klenow reaction buffer (proprietary information), 10x dNTP mix and exo- Klenow (concentrations unknown) were added and the reaction made up to 20 μl with nuclease-free water. These reactions were incubated at 37 °C for 30 min.

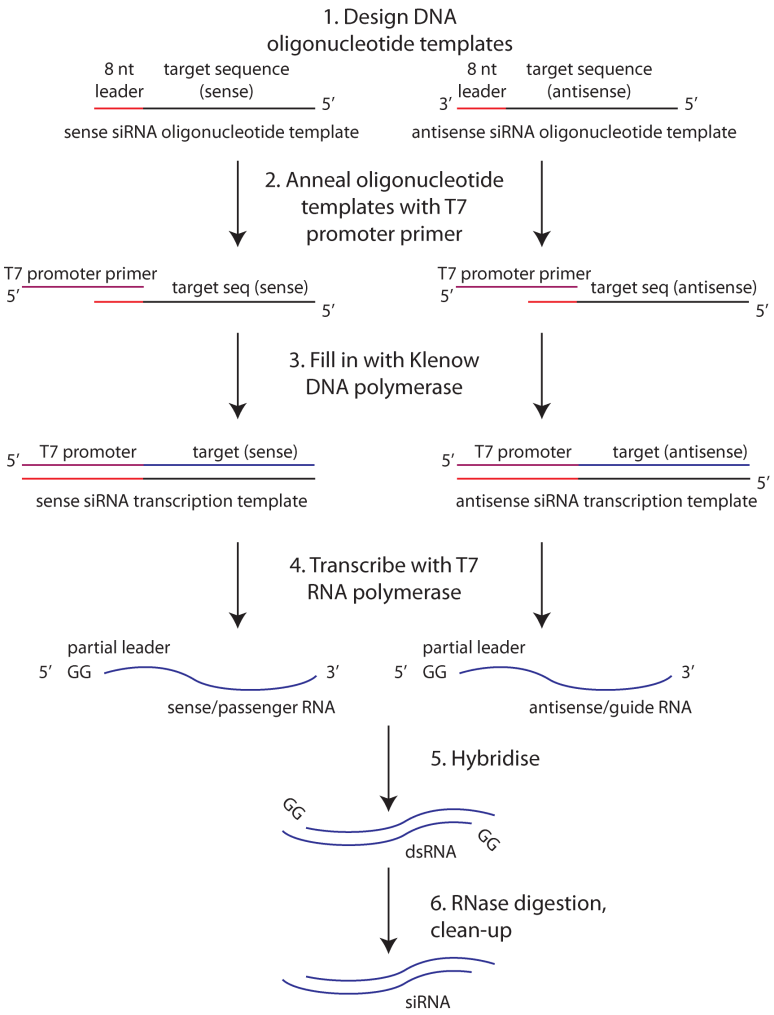


Figure 4.2: Procedure for construction of siRNAs by an *in vitro* transcription method. siRNAs were constructed using the *Silencer*® siRNA Construction Kit (Ambion, TX, USA), which employs T7 RNA polymerase to synthesise sense and antisense strands that are subsequently hybridised to form siRNAs with two nt overhangs. Figure adapted from *Silencer*® siRNA Construction Kit (Ambion, TX, USA).

Transcription reactions were assembled using 2 μl of siRNA template from the previous Klenow DNA polymerase reaction, 2 μl 10x T7 reaction buffer (proprietary information), 10 μl 2x dNTP mix and 2 μl T7 enzyme mix (concentrations unknown) in a final volume of 20 μl. Transcription reactions were incubated at 37 °C for 2 hours before combining the cognate sense and antisense reactions. Continued incubation at 37 °C overnight maximised RNA yield and facilitated hybridisation of sense and antisense strands.

The single-stranded RNA leader sequence of two G nt and DNA template were simultaneously digested. To the transcription reactions, 6 μl of digestion buffer (proprietary information), 3 μl RNase and 2.5 μl DNase (concentrations unknown) were added along with nuclease-free water to a volume of 100 μl. Digestion reactions were incubated at 37 °C for 2 hours.

siRNAs were purified on supplied filter cartridges (Ambion, TX, USA). Firstly, 400 μl binding buffer (proprietary information) was added to the digestion reactions and incubated at room temperature for 2-5 min. The siRNA-binding buffer mix was applied to a filter cartridge pre-wet with 100 μl wash buffer (proprietary information). Cartridges were centrifuged at 16,000 x *g* for 1 min and the flow-through discarded. The filters were then washed twice by addition of 500 μl wash buffer followed by centrifugation at 16,000 x *g* for 1 min and flow-through discard. Filters were transferred to a clean tube and wet with 100 μl nuclease-free water, preheated to 75 °C. Following a 2 min incubation at room temperature, siRNAs were eluted by centrifugation at 16,000 x *g* for 2 min.

The concentration of siRNAs was determined by spectrophotometry on a NanoDrop (Themo Fisher Scientific, Inc., MA, USA) where 1 absorbance unit at a wavelength of 260 nm (A_{260}) equates to 40 μg/ml siRNA. Purity of samples was also determined by spectrophotometry ie proximity of the A_{260}/A_{280} ratio to 2.0. The molar concentration (μM) of siRNAs was determined by dividing the concentration in μg/ml by 14 (as there are 14 μg of RNA in 1 nmol of an average 21-mer dsRNA). Aliquots of siRNAs at 33 μM were prepared and stored at -60 °C.

In addition to synthesis of siRNAs designed to target the *htatsf1* and *ddx3x* promoters (Table 4.2), a control siRNA targeting the *gapdh* (Entrez gene ID 2597) mRNA, siGAPDH, was prepared using the control oligonucleotide templates (sequences not provided) supplied in the *Silencer*® siRNA Construction Kit (Ambion, TX, USA).

4.2.4 Transfection of siRNAs

Transfections were carried out in a similar manner as for DNA plasmid transfections (see 2.2.5). However, the transfection reagent used was Lipofectamine™ RNAiMAX (Invitrogen, CA, USA) at a volume of 3 μl for every 100 pmol siRNA (ie 1:1 volumetric ratio with 33 μM siRNA stock). In all experiments, siRNAs were transfected at a concentration of 100 nM in the culture media.

4.2.5 Mammalian cell culture

HEK293T and TZM-bl cells were cultured as described in 3.2.1 and 2.2.4, respectively.

4.2.6 Screening siRNAs for inhibition of target expression

HEK293T or TZM-bl cells were seeded at 1.5 x 10^5 or 3 x 10^4 cells per well, respectively, in a 24-well plate. Twenty-four hours later, cells were transfected with either siRNAs at 100 nM (see 4.2.4) or 500 ng of DNA plasmids pTZU6+1 and pTZU6+shH1 (see 2.2.5) per well in triplicate – the latter to provide a positive control for *htatsf1* mRNA suppression. In some experiments, 24 hours post-transfection, cells were treated with the histone deacetylase inhibitor TSA (Sigma, MO, USA) at either 50 or 100 nM, depending on whether HEK293T or TZM-bl cells were used, respectively. Seventy-two hours post-transfection, RNA was extracted (see A1.3.1), DNase-treated and reverse transcribed (see 3.2.12). qPCR on the resulting cDNA was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4) and primers listed in Table 2.4.

4.2.7 Detection of promoter-associated transcripts

To examine whether transcripts were detectable in various regions of the *htatsf1* promoter, RNA was extracted from both HEK293T and TZM-bl cells. One microgram of RNA was DNase-treated and reverse transcribed (see 3.2.12) in quadruplicate. Another 1 μg of RNA per cell line was DNase-treated but not reverse transcribed to allow quantification of undigested genomic DNA (gDNA). An additional control lacking RNA was included per cell line.

qPCR was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4). Primers were designed to amplify a region upstream of the FANTOM promoter region (cluster ID L3_chrX_+_135407060) and a region within the TSR (TSR_hg18_v1_chrX_+_135407060) (Table 4.3). A coding region was also amplified using primers described previously (Table 2.4), to which promoter transcript abundance could be compared.

Expression of transcripts was reported as a function of the C_p median, which was calculated by the Lightcycler® V.2 software (Roche, Mannheim, Germany):

$$Expression = 2^{(-C_p(+RTsample))} - 2^{(-C_p(-RTcontrol))} \quad (4.1)$$

Following qPCR, samples were extracted from capillaries by centrifugation at 400 x *g* for 15 sec prior to resolution on a 2% agarose gel (see A1.2.3).

Table 4.3: Primers used to determine levels of *htatsf1* promoter-associated transcripts.

Primer	Sequence (5'-3')	Amplicon length (bp)
upstream forward	AGCAGCGGGTCCCCCCAA	92
upstream reverse	ACACGCCCACTATGGCGTCA	
TSR forward	TGCGTCGGCCTGAGCA	151
TSR reverse	ACCGCCGCTAGGTCCA	

Orientation of promoter-associated transcripts spanning the TSR and coding regions were determined in a similar manner but the same DNase-treated RNA extracted from HEK293T or TZM-bl cells was subject to strand-specific reverse transcription. The procedure described in 2.2.7 was followed with the exception that reverse transcription was primed with one of the primers listed in Table 4.3 only, not random hexamers. It should be noted that the 'forward' primers (same sequence as the + strand of DNA and, therefore, the sense transcript) prime reverse transcription of antisense transcripts and 'reverse' primers prime reverse transcription of sense transcripts.

4.2.8 Examination of the role of Ago proteins

HEK293T or TZM-bl cells were seeded at 2 x 10^5 or 5 x 10^4 cells per well, respectively, in a 24-well plate. Immediately following administration of cells to culture plates, cells were transfected with 25 nM of Dharmacon SMARTpool (Thermo Fisher Scientific, Inc., MA, USA) in triplicate. Each SMARTpool contains four siRNAs designed to induce a potent reduction in target gene expression without off-target effects. Three SMARTpools were used, two targeting Ago1 (EIF2C1; Entrez gene ID 26523) or Ago2 (EIF2C2; Entrez gene ID 27161), and one containing non-specific siRNAs as a control. Transfections were carried out as described in 2.2.5 but using 1 μl Dharmacon® DharmaFECT® per well. To examine the ability of the SMARTpools to suppress their targets, RNA was extracted 48 hours post-seeding and transfection (see A1.3.1).

To examine the role of Ago1 and Ago2 in siRNA-mediated silencing of *htatsf1*, seeding and SMARTpool transfection was repeated, as above. Twenty-four hours later, cells were again

transfected, this time with siRNAs at 100 nM (see 4.2.4) in triplicate and 72 hours later, RNA was extracted (see A1.3.1).

All extracted RNA was DNase-treated and reverse transcribed (see 3.2.12). qPCR on the resulting cDNA was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4) and primers listed in Table 2.4. Additional primers specific to Ago1 and Ago2 were also used: Ago1 forward 5' GCA CTG CCC ATT GGC AAC GAA 3'; Ago1 reverse 5' CAT TCG CCA GCT CAC AAT GGC T 3'; Ago2 forward 5' GGC CCA GTA TTT CAA GGA CA 3'; and, Ago2 reverse 5' TTT CTG CTC CTG TCC GAC TT 3' (Meister et al. 2004; Turner et al. 2009). Controls, cycling parameters and reporting of data were as described in 2.2.7.

4.2.9 Nuclear run-on assay to determine *htatsf1* transcription rates

The ability of promoter-targeted siRNAs to alter the transcription rate of *htatsf1* was assessed by nuclear run-on assay. HEK293T cells were seeded at 5 x 10^5 cells per well in a 6-well plate. Twenty-four hours later, cells were transfected with siRNAs at 100 nM (see 4.2.4) or 2 μg of DNA plasmids pTZU6+1 and pTZU6+shH1 per well (see 2.2.5) in triplicate – the latter to provide a negative control for *htatsf1* transcription rate change. Seventy-two hours post-transfection, media was removed and cells washed with PBS. Cells were harvested by treatment with trypsin and centrifugation at 400 x *g* for 5 min at 4 °C. Cell pellets were then subject to the same procedure described for SupT1 cell pellets (see 3.2.15). All remaining stages of the nuclear run-on assay were performed as described in 3.2.15.

4.2.10 Quantification of p81A-4 infectious particle production

TZM-bl cells were seeded at 3 x 10^4 cells per well in a 24-well plate. Cells were transfected 24 hours later with either siRNAs at 100 nM (see 4.2.4) or 500 ng of DNA plasmids pTZU6+1 and pTZU6+shH1 per well (see 2.2.5) in triplicate. Twenty-four hours later, cells were infected with various R5-tropic HIV-1 subtype B molecular clone p81A-4 (NIH AIDS Research & Reference Reagent Program, MD, USA) (Chesebro et al. 1992; Walter et al. 2005) (see A1.4.2) at a $TCID_{50}$ (see A1.4.3) of 1000/ml in the presence of 15 μg/ml DEAE-D (primary infection).

Cells were washed three times with PBS (Gibco, BRL, UK) 24 hours post-infection. Forty-eight hours post-infection, 100 μl of culture supernatant was used to infect additional TZM-bl cells seeded at 3 x 10^4 cells per well in a 24-well plate the preceding day, in the presence of 15 μg/ml DEAE-D (secondary infection). Tat-induced luciferase activity was assessed 48 hours post-

secondary infection in cell lysates using the Bright-Glo™ Luciferase Assay System (Promega,WI, USA) (see 2.2.10).

4.2.11 Statistical analyses

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. Data was analysed using one-way ANOVAs, followed by a Dunnett's multiple comparison post-test, and two-way ANOVAs, followed by Bonferroni post-tests, where appropriate, calculated using GraphPad Prism 4.0c (GraphPad Software, Inc., CA, USA).

4.3 Results

4.3.1 Promoter-targeted siRNAs inhibit target HDF expression

Initial attempts to induce TGS at HDFs failed (see A2), related in part to inaccurate promoter annotation. Three siRNAs were already screened for TGS at *htatsf1* and *ddx3x*: neither siH10, siH11 nor siH12 reduced *htatsf1* mRNA expression (Figure A2.3); siD10, siD11 and siD12 all suppressed *ddx3x* mRNA levels, but transcription rates were unaltered in siD10-treated cells (Figure A2.9). To increase the likelihood of promoter targeting, 5' RACE was performed to identify the most active TSS for these genes in both HEK293T and TZM-bl cells using the GeneRacer™ Kit (Invitrogen, CA, USA). However, despite no RNA degradation and the use of a nested PCR approach, no amplification products from decapped, extended mRNA were detected (data not shown). Thus, as a proxy for TSS and promoter identification, various features were mapped, including TSRs, TFBS and core promoter elements (Figure 4.3).

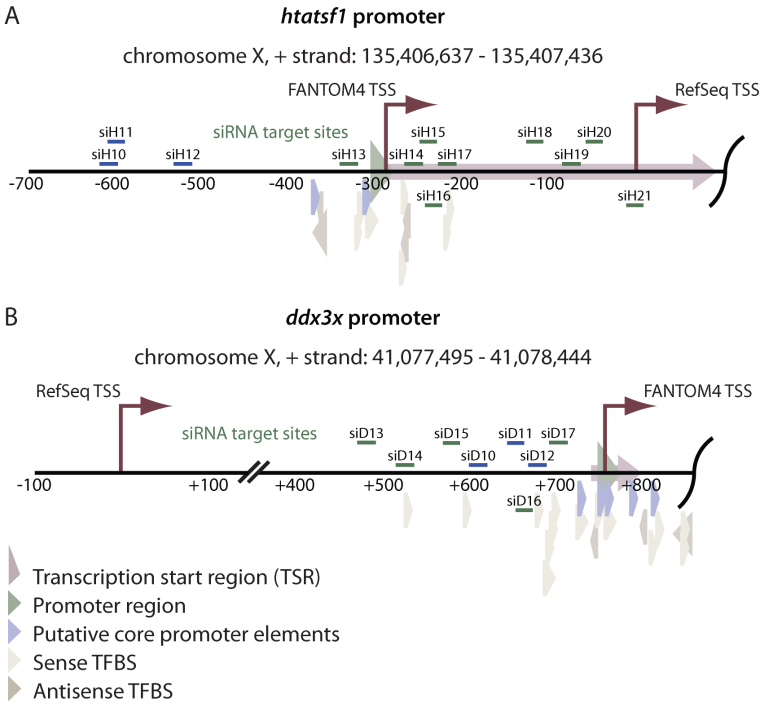


Figure 4.3: Rational design of promoter-targeting siRNAs. Features of the (A) *htatsf1* and (B) *ddx3x* promoters were mapped to increase the likelihood of promoter targeting and, therefore, obtaining transcriptional silencing. Features considered included the RefSeq TSS, TSR, promoter region, core promoter elements and TFBS. Most of the data was retrieved from FANTOM project portals. siRNA target sites in blue were screened previously (see A2).

For *htatsf1*, the upstream and downstream TFIIB recognition elements, BRE^u and BRE^d, respectively, were present. For *ddx3x*, the region was found to contain the TATA box, two Inr elements and DCE S_I and S_{III}. The presence of such core promoter elements proximal to the TSRs

of both genes supported that these regions are important in transcription initiation. siRNA target sites were selected from regions within and proximal to the TSRs of HDFs *htatsf1* and *ddx3x* (Figure 4.3). TFBS were avoided where possible to minimise the possibility of decreasing expression through blocking binding of a transcription factor, which may not be accompanied by the epigenetic changes associated with long-term silencing.

siRNAs were screened for their ability to suppress target gene expression by qRT-PCR in HEK239T cells. Levels of *ddx3x* mRNA were not significantly reduced in cells treated with siD13-17 (Figure 4.4B), with only the positive control siD10 decreasing target gene expression (p <0.01), which functions via PTGS (Figure A2.9). siD14-treatment led to a moderate increase in *ddx3x* mRNA compared to a GAPDH mRNA-targeting control, siGAPDH (30%; p <0.05), suggesting a degree of TGA (see 1.4.4). This was not investigated further as the aim of this study was to silence HDF expression with the ultimate goal of inhibiting HIV-1 replication.

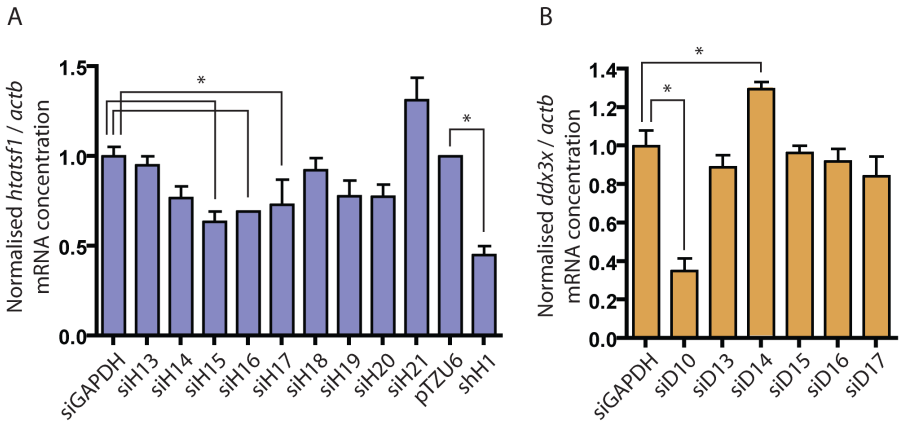


Figure 4.4: Promoter-targeting siRNAs can suppress *htatsf1* mRNA levels. A. siRNAs targeting sequences in the *htatsf1* promoter were screened in HEK293T cells for their ability to decrease levels of target mRNA in triplicate. Expression of *htatsf1* mRNA is reported relative to *actb* mRNA and normalised to the negative control siGAPDH. shH1 was included as a positive control for *htatsf1* suppression and normalised to the mock plasmid pTZU6+1. qRT-PCR was performed on total RNA extracted 72 hours post-transfection. **B.** As (A) but siRNAs targeted sequences in the *ddx3x* promoter and *ddx3x* expression is reported. siD10 was included as a positive control for *ddx3x* suppression *, p <0.05, one-way ANOVA with Dunnett post-tests relative to siGAPDH/pTZU6.

siH15, siH16 and siH17 significantly decreased *htatsf1* mRNA levels 72 hours post-transfection compared to the control siRNA by up to 35% (p <0.05; Figure 4.4A). However, as demonstrated by the positive controls used for target gene silencing (ie shH1 and siD10), qRT-PCR to discern mRNA levels in total RNA extracts is not discriminatory between PTGS and TGS.

4.3.2 siH15-mediated *htatsf1* silencing is attenuated by TSA in HEK293T cells

Classical TGS is accompanied by epigenetic changes at the target site associated with silent-state chromatin, including methylation of histone residues such as H3K27 and H3K9 (Ting *et*

al. 2005). During TGS induction, acetyl groups, which on histone residues such as H3K27 and H3K9 are associated with active transcription, require removal by histone deacetylases to allow for subsequent methylation. Consequently, classical TGS induction may be sensitive to the histone deacetylase inhibitor TSA.

TSR-targeting siRNA-mediated suppression of *htatsf1* mRNA levels was attenuated by treatment with 50 nM TSA in HEK293T cells (Figure 4.5A). Silencing mediated by shH1 was not affected by TSA treatment in both HEK293T and TZM-bl cells (Figure 4.5A and B), consistent with a mRNA-targeted PTGS mechanism. Despite using twice the concentration of TSA, attenuation of siH15-mediated silencing was not observed in TZM-bl cells (Figure 4.5B). Similarly, the siH21-mediated increase in *htatsf1* levels was enhanced by TSA treatment in HEK293T (Figure 4.5A), but not in TZM-bl (Figure 4.5B), cells. Although it is possible that an even higher concentration of TSA was needed to see an effect in TZM-bl cells, this data suggests that both promoter-targeted siRNA-induced suppression and activation of gene expression in HEK293T, and not TZM-bl, cells involves chromatin remodelling.

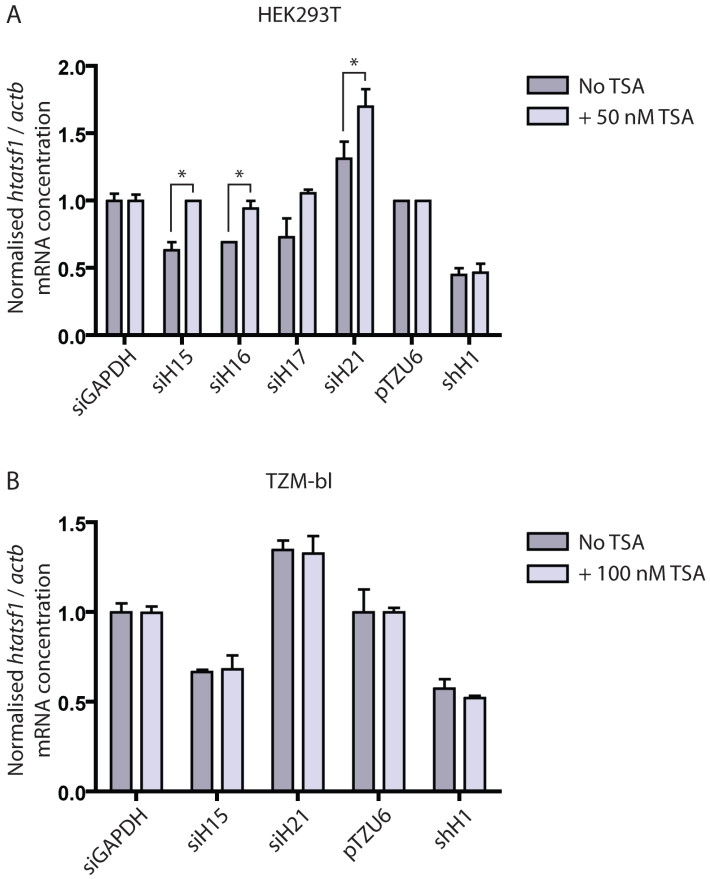


Figure 4.5: Promoter-targeting siRNA-mediated silencing is sensitive to TSA in HEK293T cells. **A.** HEK293T cells were transfected with siRNAs and 24 hours later media was replaced with either fresh media or media containing 50 nM TSA in triplicate. qRT-PCR was performed on total RNA extracted 72 hours post-transfection. **B.** As (A) but using TZM-bl cells that were treated with 100 nM TSA. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

4.3.3 Bidirectional promoter-associated transcripts are present at the siH15 target site

Since promoter-associated transcripts are required for TGS (Han *et al.* 2007; Gonzalez *et al.* 2008; Schwartz *et al.* 2008; Hawkins *et al.* 2009; Tan *et al.* 2009b; Yue *et al.* 2010), whether there was a difference in target region-associated transcript levels between the two cell lines was investigated. As the most active TSS was unknown, transcript levels associated with two different regions, within and upstream of the *htatsf1* TSR, were examined (Figure 4.6A).

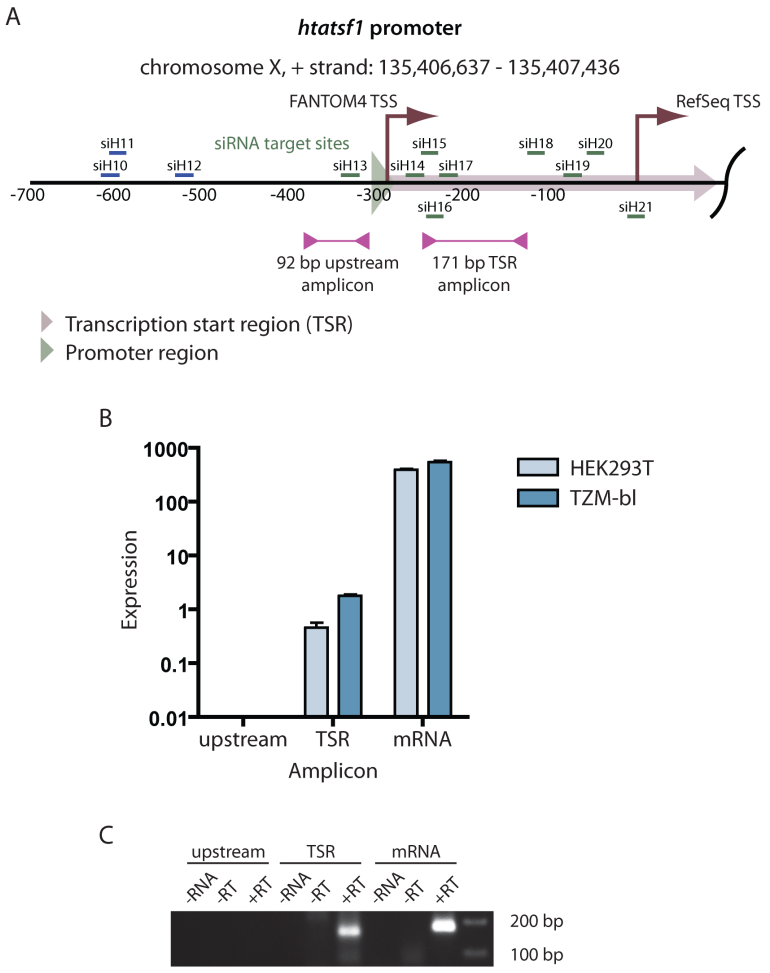


Figure 4.6: Transcription is detectable across the TSR but not upstream of the FANTOM TSS. **A.** Schematic of the *htatsf1* promoter showing the regions analysed for transcription. Primers were designed to amplify transcripts present upstream of, and within, the TSR. **B.** qRT-PCR was performed on total RNA extracted from both HEK293T and TZM-bl cells in quadruplicate. Primers designed to amplify different regions of the *htatsf1* promoter were used. Primers specific to a region in the *htatsf1* mRNA coding region were also included (amplicon 171 bp). **C.** Following qPCR, samples were resolved on an agarose gel. Representative samples from HEK293T cells are shown.

No transcripts were detectable by qPCR upstream of the TSR in either HEK239T or TZM-bl cells (Figure 4.6B & C). Transcripts associated with a region within the TSR were detected in both cell lines (Figure 4.6 B & C), albeit 100-fold less abundant than mRNAs (Figure 4.6B). There was no

significant difference in the levels of mRNA and TSR-associated transcripts between the HEK293T and TZM-bl cell lines. The 100-fold difference between mRNA and TSR-associated transcript suggests that the latter may be a low-copy, promoter-associated, non-coding transcript.

The orientation of the TSR-associated transcripts was investigated; directional RT cDNA products were subject to qPCR using both mRNA and TSR primers. Again, similar results were obtained for both cell lines (Figure 4.7A).

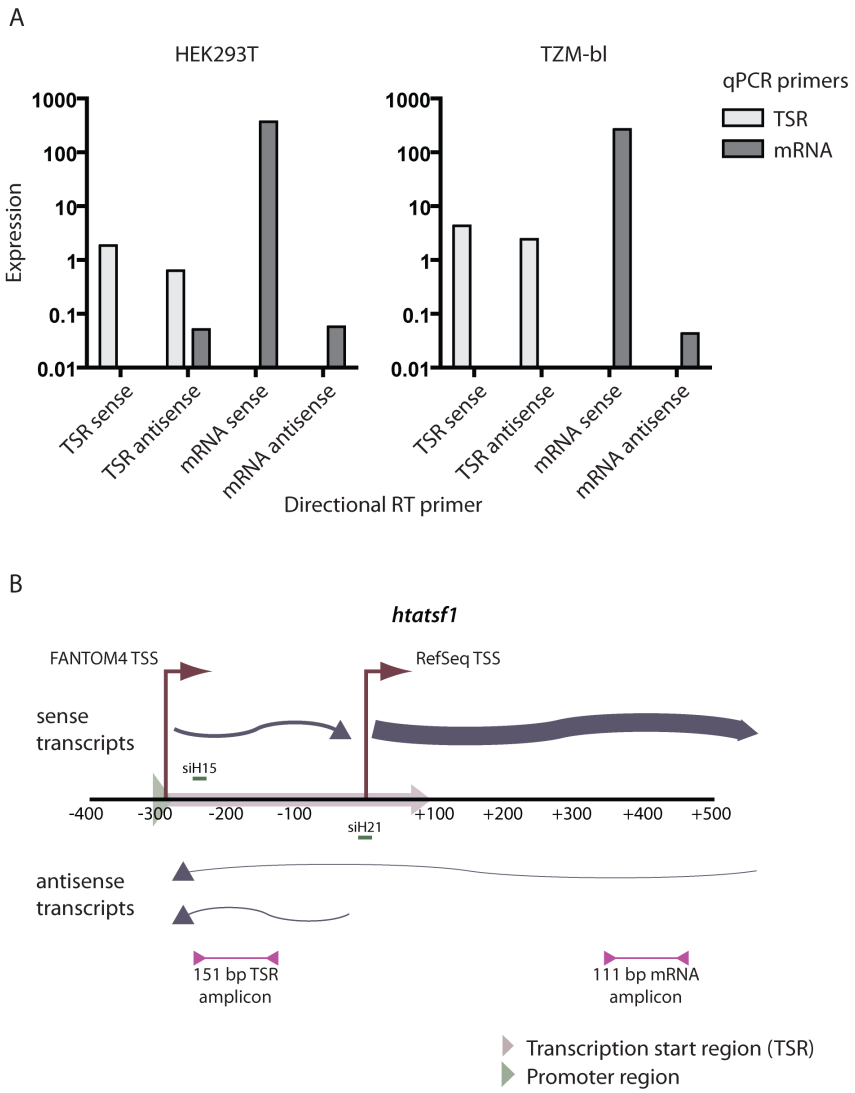


Figure 4.7: Sense and antisense transcription is detectable within the TSR. **A.** Total RNA was extracted from both HEK293T and TZM-bl cells and subject to directional reverse transcription using a strand-specific primer complementary to sequences in either the TSR or coding region. The resulting cDNA was quantified by real-time PCR. **B.** Schematic illustrating the data in (A).

Not surprisingly, transcripts associated with a coding region were predominantly sense, with antisense transcripts barely detectable. Both sense and antisense TSR-associated transcripts were detectable and at similar levels (Figure 4.7A), which may be a product of divergent transcription initiation (Core *et al.* 2008). This would also account for the observation that

antisense transcripts at the TSR were higher than that observed at the mRNA-associated region. As expected, no TSR sense transcripts were detected with mRNA qPCR primers as the directional RT employed the TSR reverse primer. Similarly, no mRNA antisense transcripts were detectable with the TSR qPCR primers as the directional RT used the mRNA forward primer. Of note is that no mRNA sense transcripts were detectable with the TSR qPCR primers, which suggests that the bulk of *htatsf1* mRNAs are initiated downstream of the TSR amplicon in both cell lines. Conversely, low-level antisense TSR-associated transcripts were detectable with the mRNA qPCR primers in HEK293T cells, suggesting that ~1/10th of these antisense transcripts are initiated downstream of the mRNA amplicon and, therefore, individual antisense transcripts may span the entire coding and promoter regions of the *htatsf1* gene (Figure 4.7B).

Together, this data suggests that both sense and antisense TSR-associated transcripts are present in HEK293T and TZM-bl cells at similar levels. Based on a comparison with *htatsf1* mRNA levels and directional RT-PCR, it appears the sense TSR-associated transcripts are not coding transcripts. This supports that siH15 is targeting a promoter-associated transcript.

4.3.4 Ago1 and Ago2 suppression does not attenuate siH15-mediated *htatsf1* silencing

Whilst Ago2 is unequivocally involved in siRNA-mediated PTGS (Meister *et al.* 2004), both Ago1 and Ago2 have been implicated in TGS (Janowski *et al.* 2006; Kim *et al.* 2006; Gonzalez *et al.* 2008; Hawkins *et al.* 2009; Tan *et al.* 2009b; Turner *et al.* 2009; Chu *et al.* 2010). Therefore, the relative roles of Ago1 and Ago2 in siRNA-mediated modulation of *htatsf1* mRNA expression were investigated.

Firstly, the ability of the SMARTpools to suppress target Ago expression was verified. In both HEK293T and TZM-bl cells, Ago1 and Ago2 mRNA levels were decreased by at least 60% 48 hours following transfection with the respective SMARTpool compared with a non-specific SMARTpool control (p <0.001; Figure 4.8). Having confirmed that the SMARTpools were effective in decreasing Ago mRNA expression, cells were first transfected with a SMARTpool and 24 hours later, transfected again with either the control, siGAPDH, or siH15 or siH21, which suppress and increase *htatsf1* mRNA levels, respectively (Figure 4.4A). Seventy-two hours post-siRNA transfection (and 96 hours post-SMARTpool transfection), total RNA was extracted and subject to qRT-PCR. No attenuation of siH15-mediated effects was seen on suppression of either Ago1 or Ago2 in both cell lines (Figure 4.9A). There was attenuation of the siH21-mediated increase in *htatsf1* mRNA on suppression of Ago2, although this was not statistically significant.

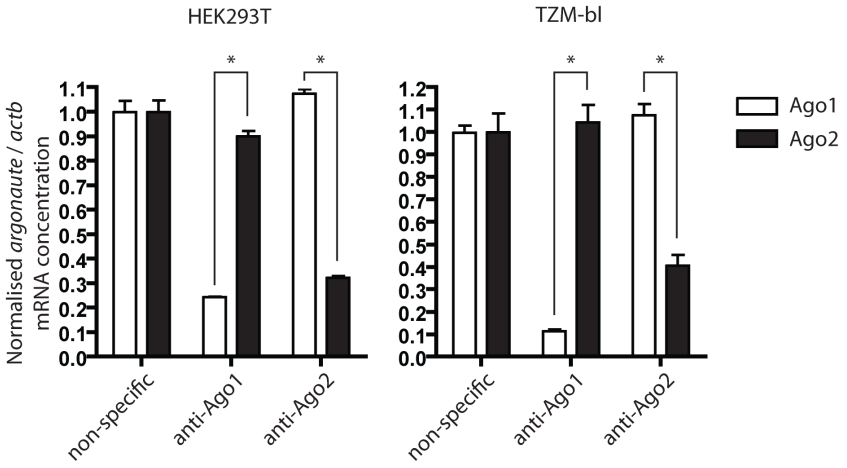


Figure 4.8: SMARTpools effectively suppress Ago expression. Both HEK293T and TZM-bl cells were transfected with SMARTpools targeting either Ago1 or Ago2, or a SMARTpool containing non-specific siRNAs as a control, in triplicate. Argonaute mRNA levels were determined by qRT-PCR 48 hours post-transfection. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

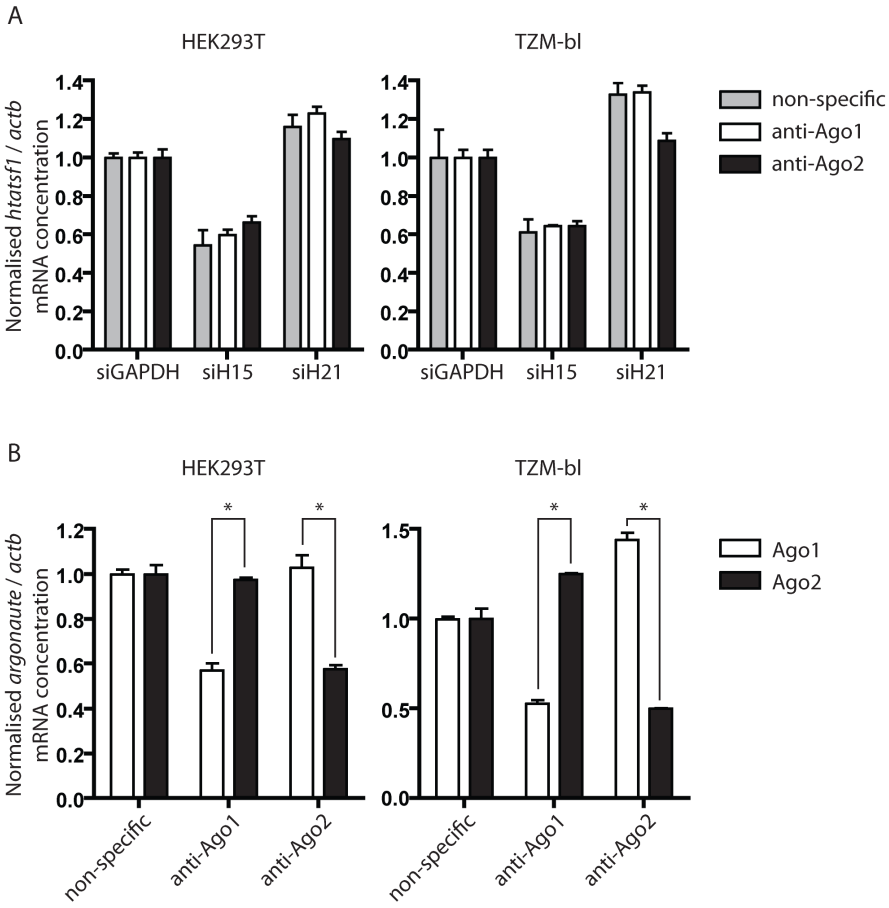


Figure 4.9: Ago suppression has no significant effect on siRNA-mediated silencing. **A.** Both HEK293T and TZM-bl cells were transfected with SMARTpools targeting either Ago1 or Ago2 or a SMARTpool containing non-specific siRNAs as a control. 24 hours later, cells were again transfected with promoter-targeting or control siRNAs in triplicate. *htatf1* mRNA levels were assessed relative to *actb* by qRT-PCR on RNA extracted 96 hours post-SMARTpool transfection. **B.** As (A), but showing levels of Argonaute mRNA levels in cells transfected with siGAPDH. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

The lack of significant attenuation of siH15 or siH21 effects may be due to insufficient suppression of Ago expression, which is difficult to determine without Western blot data. Nonetheless, SMARTpool efficacy in the samples was confirmed and, although the suppression of Ago1 and Ago2 mRNA was not as pronounced as that observed 48 hours post-transfection (Figure 4.8), there was still considerable reduction in levels at 96 hours post-transfection (p <0.001; Figure 4.9B). Moreover, the observed suppression of Ago mRNA levels is greater than that attained in previous experiments using anti-Ago siRNAs (A2.2.6; Figure A2.5B), which was sufficient to significantly attenuate siRNA-mediated PTGS of *psip1* (Figure A2.5A). Therefore, together with this data, it seems Ago1 and Ago2 have redundant roles in siH15-mediated silencing or the silencing is Ago1- and Ago2-independent, perhaps with another Ago protein involved (ie Ago3 or Ago4). The lack of attenuation of siH15-mediated silencing on Ago2 suppression does not support a PTGS mechanism. To clarify whether this is the case, transcription rates of *htatsf1* in siRNA-treated cells were quantified.

4.3.5 Promoter-targeted siRNAs modulate the *htatsf1* transcription rate

A nuclear run-on experiment may discriminate between silencing by PTGS and TGS mechanisms as it quantifies the transcription rate of a specific transcript. With PTGS, the levels of target mRNA are decreased but the rate of transcription is unaffected. With TGS, both target mRNA and transcription rate is suppressed.

Transcription rates were examined in HEK239T cells treated with shH1, known to reduce *htatsf1* mRNA by PTGS, and a shRNA control, shHBVx, as well as the *htatsf1* promoter-targeting siRNAs and the siRNA control, siGAPDH. As expected, the *htatsf1* transcription rate in cells treated with shH1 was unaltered (Figure 4.10). In contrast, a significant decrease (57%; p <0.01) and increase (41%; p < 0.01) in transcription rates from siH15- and siH21-treated cells, respectively, were observed. This suggests that siH15 and siH21 are modulating *htatsf1* expression at the transcriptional level in HEK293T cells. Unfortunately, inefficient *in vitro* transcription from TZM-bl nuclei did not enable a similar analysis in this cell line.

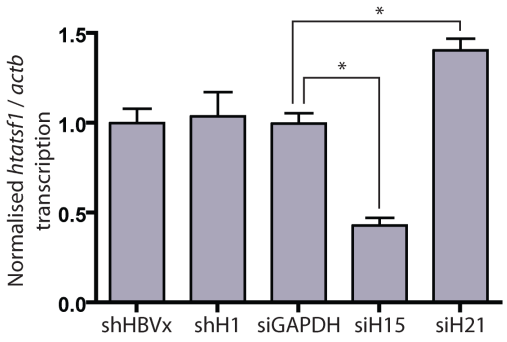


Figure 4.10: Promoter-targeting siRNAs may alter *htatsf1* transcription rates. A nuclear run-on assay was performed on samples 72 hours post-transfection with promoter-targeting siRNAs. Transfection with shH1, which silences HTATSF1 expression via a PTGS mechanism, was included as a control. Transcription rates relative to that of *actb* were normalised to the appropriate negative control. Results from two independent experiments are shown. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to shHBVx/siGAPDH.

4.3.6 siH15-mediated HTATSF1 suppression inhibits HIV-1 infectious particle production

Having demonstrated that siRNA-induced TGS of *htatsf1* is possible in HEK293T cells, whether this mode of silencing is sufficiently potent to inhibit subtype B molecular clone p81A-4 infectious particle production from TZM-bl cells was assessed. Tat-induced luciferase activity was significantly reduced in TZM-bl cells infected with culture media collected from cells treated with siH15 (84%, p < 0.01; Figure 4.10). As observed previously (Figure 2.14), shH1 also inhibits infectious particle production (59%, p < 0.01; Figure 4.11). No significant increase in infectious particle production was observed in cells treated with siH21, suggesting that physiological levels of HTATSF1 are not limiting for virus production. siRNAs targeted to the *htatsf1* promoter are, therefore, as effective at inhibiting HIV-1 infectious particle production from TZM-bl cells as a mRNA-targeted shRNA.

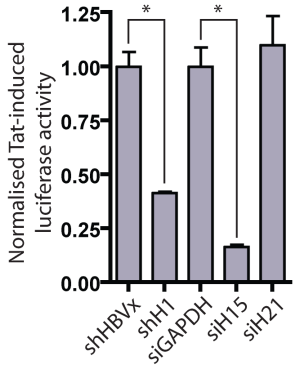


Figure 4.11: Promoter-targeting siRNAs can decrease p81A-4 infectious particle production. TZM-bl cells were transfected with either promoter-targeting siRNAs or a mRNA-targeting shRNA and appropriate controls in triplicate. 24 hours later, cells were infected with p81A-4 at a $TCID_{50}$ of 1000/ml. 48 hours later, virus-containing culture supernatant was used to infect additional untransfected TZM-bl cells and Tat-induced luciferase activity quantified 48 hours post-infection. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to shHBVx/siGAPDH.

4.4 Discussion

The approach used here was to screen siRNAs targeting regions within and proximal to the TSRs of *htatsf1* and *ddx3x* for TGS and examine whether silencing by this mechanism would be sufficient to inhibit HIV-1 replication. This study shows that, whilst not all promoter-targeting siRNAs mediated silencing of their targets, there is some evidence, including promoter-transcript levels and orientation, sensitivity of silencing to TSA and decreased transcription rates, that siH15 induces TGS of *htatsf1* in HEK293T cells.

It is unclear why some of the promoter-targeting siRNAs led to decreased target mRNA levels whilst others did not. siD10 mediated PTGS of *ddx3x* (Figures 4.4 & A2.9) whilst siRNAs targeting nearby sites had no effect. For those targeting *htatsf1*, siH15-17 which suppressed mRNA levels to the greatest extent were all clustered <100 bp downstream of the FANTOM4 TSS and >200 bp upstream of the RefSeq TSS. Whether this region is good to target because of its location with respect to other promoter features and the TSS is unknown, in part because 5' RACE data to confirm the TSS was not obtained. The FANTOM4 TSS may not be the most active TSS in HEK293T or TZM-bl cells, particularly as neither of these cell lines (or tissues from which they were derived) were among the 56 human CAGE libraries used in TSR determination (Balwierz *et al.* 2009).

The reason siH15-17 mediated the greatest *htatsf1* inhibition may have been unrelated to target location with respect to promoter features. The GC content of these three target sites were the only ones within the *htatsf1* promoter <55% (Table 4.2). This may be significant as it has been suggested that siRNAs with low GC content (30-50%) are more active in inducing PTGS than those with high GC content (Ambion's siRNA target finder, TX, USA). If this holds true for TGS, it may prove difficult to design many promoter-targeting siRNAs per gene because most promoters are GC-rich (Kim *et al.* 2005; Balwierz *et al.* 2009).

At present, only one study has examined where in the promoter to target in any depth. A series of overlapping siRNAs were designed to target the open complex region revealing that the resulting TI was extremely sensitive to single base shifts (Janowski *et al.* 2005). The sensitivity was attributed to the accessibility of the target to the guide strand. Less is known about where to target in a promoter to obtain classical TGS. Some studies have used pools of siRNAs, and so provide little information as to what constitutes a good target site (Pulukuri and Rao 2007; Gonzalez *et al.* 2008). Most have focused on sequences up to 200 bp of the predicted TSS. The rationale being that this is most likely to correspond to the nucleosome-depleted region (NDR) and, therefore, the best place to focus chromatin remodelling.

In contrast, Gonzalez *et al* demonstrated TGS of four genes by miR-17-5p and miR-20a with the majority of target sites residing >1 kb upstream of the RefSeq TSS (Gonzalez *et al.* 2008). The same article also reported that exogenous siRNAs could induce TGS of the INK4b-ARF-INK4a locus by targeting the well-known regulatory domain >1 kb upstream. Thus, similarly to TGA (Matilainen *et al.* 2010), TGS may be induced by targeting distant regions. Furthermore, regulatory elements may be important in choosing target sites and, thus, the optimal distance from the TSS could vary substantially between genes.

In this study, the only regulatory elements mapped were TFBS. Different outcomes have been reported following TFBS targeting: Janowski *et al* observed that siRNAs spanning Sp1 binding sites were better inducers of TI (Janowski *et al.* 2006). In this case, blocking the TF binding site did not result in chromatin remodelling, which is desirable to attain long-term silencing. However, other studies have designed guide strands to span transcription factor binding sites that did result in heterochromatin formation (Murayama *et al.* 2006; Suzuki *et al.* 2008; Turner *et al.* 2009; Yamagishi *et al.* 2009; Palanichamy *et al.* 2010). In this study, TFBS were avoided where possible to minimise the possibility of inducing a temporary TGS without epigenetic changes. siH13 and siH17 did partially overlap TFBS, although these siRNAs did not significantly inhibit *htatsf1* mRNA levels. This could be due to a number of reasons, including the generic, such as the thermodynamics of the siRNAs and the accessibility of the target transcript. In addition, it is unknown whether these regulatory elements are active in the cell lines used and whether blocking of a single TFBS would have an effect given that such transcription factors may act in a redundant or coordinate fashion.

To date, insufficient data on negative target sites is available to enable a bioinformatic analysis on what makes a good TGS target site. Although algorithms designed to predict PTGS target sites are often used, these algorithms do not take into account the context of the sequence within the promoter architecture.

TGA of *ddx3x* and *htatsf1* was observed with siD10 and siH21, respectively (Figure 4.4). Further investigation of the effects of siH21 revealed increased transcription rates (Figure 4.10) and an attenuation of *htatsf1* activation on suppression of Ago2 expression (Figure 4.9). It is possible that an Ago2-containing RITS complex directly modulates chromatin architecture to enhance transcription. Details of how this would occur are unclear, although it has been suggested that such a mechanism involves targeting of antisense transcripts (Schwartz *et al.* 2008). In contrast, there have been reports where promoter-targeted siRNAs have induced TGA through unintended PTGS of an antisense transcript with regulatory function, mediated by the passenger strand of the siRNA (Morris *et al.* 2008). The feasibility of such a mechanism is supported by reports that both strands of a siRNA duplex are equally eligible for incorporation

into Ago2-containing RISC in HEK293 cells (Wei *et al.* 2009). In addition, long, non-coding antisense transcripts have been described as endogenous TGS effectors (Hawkins and Morris 2010). Despite the bulk of reported evidence describing an indirect TGA mechanism, both mechanisms are feasible with the data here, as antisense transcripts were identified above and below the siH21 target site and a proportion were likely spanning the coding and promoter regions (Figure 4.7). Alternatively, TGA may result if the siRNA simply blocked binding of a transcriptional repressor or enhanced the association of RNA Pol II through stabilisation of the open complex. Discrimination between potential TGA mechanisms would be facilitated by examination of whether antisense cleavage products were generated and whether the endogenous antisense transcripts subdued *htatsf1* transcription.

The presence of transcripts at the TSR and antisense transcripts at the coding region may be a result of pervasive transcription (Kapranov *et al.* 2002; Bertone *et al.* 2004; Kampa *et al.* 2004; Schadt *et al.* 2004; Carninci *et al.* 2005; Cheng *et al.* 2005; Carninci *et al.* 2006; Birney *et al.* 2007; Kapranov *et al.* 2007a; Kapranov *et al.* 2007b; van Bakel *et al.* 2010). However, the greater abundance of antisense transcripts at the TSR over the coding region is not consistent with simply 'background transcription'. What may account for this observation is that CpG islands are associated with bidirectional promoter activity (Trinklein *et al.* 2004; Core *et al.* 2008), and widespread divergent initiation, though not accompanied by productive elongation upstream, has been described at human promoters (Core *et al.* 2008).

Several lines of evidence support siH15-induced TGS at the *htatsf1* promoter in HEK293T cells: mRNA levels and transcription rates were decreased, target region-associated sense transcripts were less abundant than coding region transcripts and did not span the TSR and coding regions and silencing was sensitive to TSA (Figure 4.5A). No attenuation of siH15-mediated *htatsf1* silencing was observed on suppression of Ago1 or Ago2. This may reflect redundancy in the role of Ago proteins in TGS at the *htatsf1* promoter, which has been reported at other genes, including *pgr* (Janowski *et al.* 2006) and *hoxd4* (Tan *et al.* 2009b). Admittedly, TSA sensitivity is not definitive of TGS as the target gene may be more susceptible to a TSA-induced increase in basal transcription than the reference gene, in this case *actb*. Nonetheless, a mechanism of TGS for the observed silencing of *htatsf1* is consistent with the data presented here. It cannot be ruled out, however, that silencing is occurring by negative feedback of the siRNA on another gene that affects transcription of *htatsf1*, or by interfering in *cis* with another RNA associated with the transcription of *htatsf1*. Further experiments are required to clarify the mechanism, perhaps looking at chromatin signatures using ChIP analysis, including those associated with silent-state chromatin eg H3K27me3, and ncRNAs eg H3K39me3.

siH15-mediated *htatsf1* silencing inhibited p81A-4 infectious particle production in TZM-bl cells (Figure 4.11), although whether by a TGS mechanism is less clear. No sensitivity to TSA was observed in this cell line (Figure 4.5B). On the other hand, levels of transcripts at the target site were similar to those seen in HEK293T cells (Figures 4.6 and 4.7), supporting the targeting of a non-coding promoter-associated transcript. Unfortunately, nuclear run-on data was not obtained to clarify how transcription rates were affected. Regardless, the aim of this study was to show that a HDF could be silenced at the transcriptional level with siRNAs, and this objective was achieved in HEK293T cells. As already discussed, TSS and promoter usage may vary between tissues and so confirming TGS in TZM-bl cells would not further the ultimate goal of inducing TGS in a physiological substrate of HIV, such as T cells. To this end, it would be advantageous to screen siRNAs in T cells. However, not only is this complicated by the technical difficulties associated with electroporation of siRNAs or retroviral-transduction of shRNAs into cells, but also by the potential for TSS and promoter usage variation with T cell activation. At present, therefore, the best approach for the induction of TGS at a HDF is still to identify candidate target sites by studies in easy-to-transfect cells, such as HEK293Ts. Appropriate inducible shRNA expression cassettes can then be stably introduced into HIV-1-infectable cell lines and studied for their ability to induce TGS at a number of conditions. Time course ChIP experiments could then be used to quantify the duration of the induced TGS and how this impacts HIV-1 replication kinetics.

In summary, induction of TGS at a HDF as an anti-HIV strategy is complicated by the difficulty in screening siRNAs in T cells and the variation in TSS and promoter usage between tissues. Experiments can be performed, however, to confirm the susceptibility of a gene to TGS in an easy-to-transfect cell line, such as HEK293T cells. Here, the HDF HTATSF1 was transcriptionally silenced by a siRNA targeted to the TSR, which was described by the FANOM4 project. Treatment with the TSR-targeted siRNA mediated sufficient silencing of HTATSF1 to inhibit HIV-1 replication. Therefore, TGS at HDFs may provide a promising anti-HIV-1 strategy.

The approach used here may be used to confirm that certain genes are capable of undergoing TGS before proceeding to a more time-consuming screen in T cells. TGS was not induced at *ddx3x* and, whilst this may be a result of failing to identify the appropriate target site, it is also conceivable that the promoters of some genes may not undergo RNA-directed chromatin remodelling. Since screening for target suppression with promoter-targeted siRNAs is time consuming, whether certain promoter features are associated with TGS susceptibility was investigated in Chapter 5.

CHAPTER 5

Promoter features associated with susceptibility to small RNA-induced transcriptional gene silencing

5.1 Introduction

TGS offers great therapeutic potential (see 1.4), particularly as there are a number of advantages to TGS over PTGS: TGS may be sustained in the absence of the small RNA that initialises the TGS (Hawkins *et al.* 2009); and, TGS targets promoters that are limited in number (ie 2 copies per cell), whereas there may be many hundred mRNAs that are targeted by PTGS. Thus, TGS may prove a more potent silencing strategy than PTGS, reducing toxicities associated with RNAi saturation and off-target effects. The work in Chapter 4 provided proof of concept that TGS at HDFs may inhibit HIV-1 replication, through targeting *htatsf1*. However, TGS may be applied not just to the silencing of HDFs to combat HIV-1, but in other diseases such as cancer and those characterised by dysregulated chromatin. Despite the enormous potential of TGS-based therapy, few genes have been reported as regulated by small RNAs at the transcriptional level. It is, therefore, currently unknown whether all human gene promoters are susceptible to therapeutic transcriptional regulation in a small RNA-inducible manner.

Promoter architecture is important for selection of TGS target genes. Gonzalez *et al* searched for genes significantly down-regulated in two different colon cancer cell lines on miR-17-5p and miR-20a expression and with a canonical seed-match hexamer in the forward strand of the promoter regions (Gonzalez *et al.* 2008). Of the 104 genes fulfilling this criteria, they focused on five genes at random, only four of which were susceptible to TGS upon introduction of pre-miRNAs or miRNA mimics (Gonzalez *et al.* 2008; Iida *et al.* 2008). Furthermore, in yeast, siRNA-initiated silencing was observed at only a subset of target loci, which correlated with antisense transcription and over-expression of Swi6, the yeast homologue of HP1 (Iida *et al.* 2008). Consequently, the target sequence alone is not sufficient for TGS and other promoter characteristics are required for a TGS-permissive state.

Studying genes with contrasting promoter architectures may inform which promoter characteristics confer enhanced susceptibility to TGS. Cairns has suggested that some yeast promoters show extreme contrasts in architecture, classed as ‘open’ or ‘covered’ (Cairns 2009) (Figure 5.1). Open promoters are typically TATA-less and associated with constitutively expressed genes. They are characterised by large (~150 bp) nucleosome-depleted regions (NDRs) containing poly(dA:T) tracts that deter nucleosome formation (Struhl 1985), maintaining exposure of *cis*-

regulatory elements. In contrast, covered promoter *cis*-regulatory elements are fully or partially hidden by nucleosomes when the genes are in their repressed state (Cairns 2009). Covered promoters are, therefore, more reliant on chromatin remodelling for activity and are associated with highly regulated genes (Kawaji *et al.* 2006a; Cairns 2009). Of course, the majority of genes possess promoters which are somewhere between these two extremes.

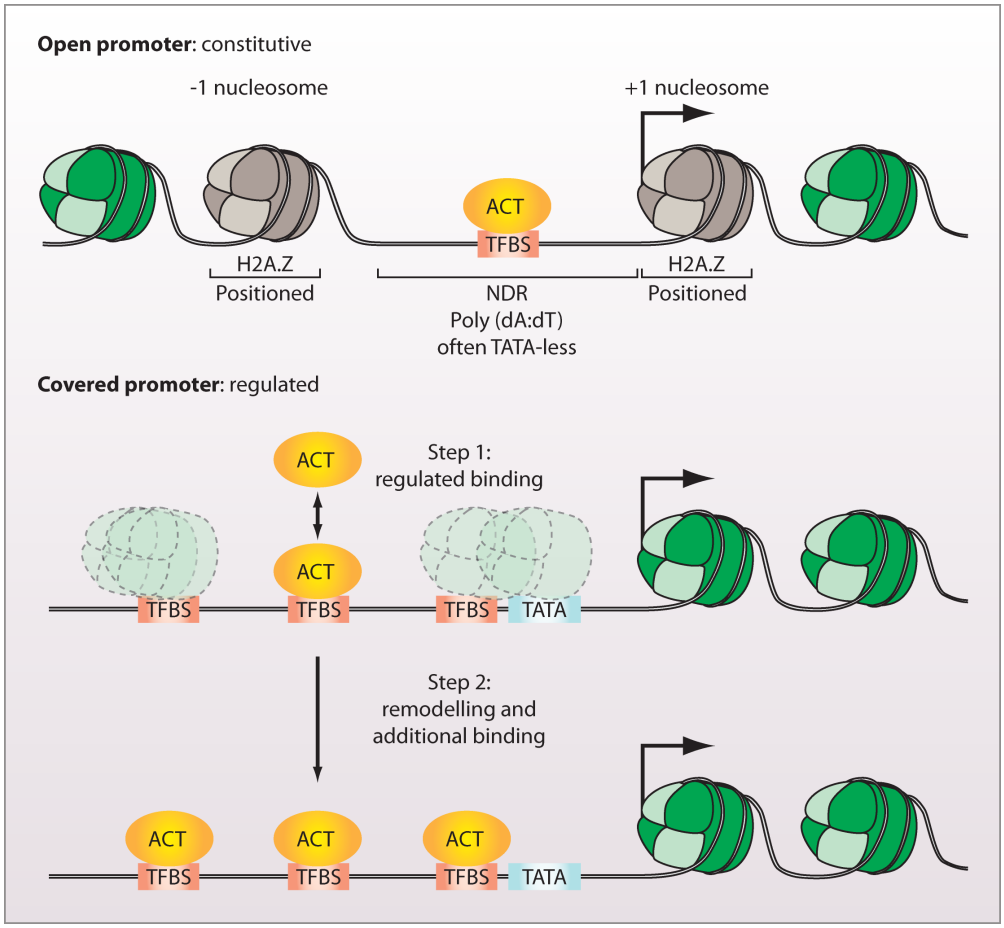


Figure 5.1: Extreme contrasts in promoter architecture. 'Open' promoters are associated with constitutively expressed genes and fixed position-nucleosomes composed of the H2A variant H2A.Z, which is associated with transcription initiation. Such promoters typically contain broad NDRs but no TATA boxes. In contrast, 'covered' promoters are more reliant on chromatin remodelling for exposure of *cis*-regulatory elements and activity.

Mammalian promoters can also be classified into two types and, in this case, the two groups include the majority. Using data from the FANTOM project (Kawaji *et al.* 2006a; Cairns 2009; Kawaji *et al.* 2009), promoters were classed based on their C or G (CG) and CpG dinucleotide content (Balwierz *et al.* 2009). Significantly, promoters of low CG and CpG content were associated with narrow TSRs (<10 bp) and high CG and CpG content with wide TSRs (25-250 bp) (Balwierz *et al.* 2009). Most importantly, more than 91% of promoters had a high probability of belonging to either the high- or low-CpG class. Thus, the method is suitable for unambiguous

classification of human promoters into two classes and different mechanisms may be involved in their regulation.

The two promoter classes are enriched in different core promoter elements, which may contribute to their differential regulation and focus the TSS. TATA boxes were strongly over-represented in narrow TSR promoters, whereas consensus Sp1 sites were over-represented in broad promoters (Carninci *et al.* 2006). This is consistent with the ability of Sp1 to recruit TATA-binding protein (TBP) in the absence of TATA boxes (Butler and Kadonaga 2002). Clearly, promoters can be classed based on their architectures, which may reflect differences in regulation and transcriptional plasticity.

Since, chromatin modifications and nucleosome remodelling are required for classical TGS, one could speculate that the two classes of human promoters are differentially susceptible to TGS. One hypothesis is that genes with broad TSRs are less conducive to TGS because of features within the promoter that deter nucleosome formation. Broad TSRs and CpG islands are associated with ubiquitous transcripts (Carninci *et al.* 2006) and it has been suggested that promoters have an intrinsic tendency to exclude nucleosomes (Sandelin *et al.* 2007).

A contrary hypothesis is that broad TSRs are more susceptible to TGS as it implies a large NDR, which could favour the production of promoter-associated transcripts (see 1.4.2). CpG islands are associated with bidirectional promoter activity (Trinklein *et al.* 2004). Further, divergent transcription and gene activity correlate with each other and promoters containing a CpG island (Core *et al.* 2008). Promoter-associated transcript levels have also been shown to correlate generally, though not exclusively, with gene expression (Seila *et al.* 2008; Taft *et al.* 2009). These include tiRNAs, which are ~18 nt long and preferentially associate with GC-rich promoters (Taft *et al.* 2009). Recently, it was suggested that these tiRNAs play a role in epigenetic regulation of gene expression (Taft *et al.* 2010).

However, the complexities of transcriptional regulation and promoter architecture may disprove both these simple hypotheses, particularly as broad TSRs do not always indicate single, dispersed promoters with large NDRs, but also overlapping proximal promoters with tissue- and temporal-specific expression (ie multiple TSCs within a TSR). Indeed, distinct TSS usage is observed in different tissues within TSRs (Kawaji *et al.* 2006a). Broad TSRs can, therefore, be regions with "fuzzy" nucleosome architecture, with nucleosome position defining TSS usage (Jiang and Pugh 2009). Whether a gene is susceptible to TGS may depend, therefore, not just on the length of the TSR, but also on the number of TSCs present within that TSR.

In short, it is unknown whether small RNAs are capable of directing *de novo* heterochromatin formation at every gene promoter. This chapter examines whether particular promoter features are enriched in genes at which TGS has been reported with the objective of

facilitating prediction of TGS-susceptible genes. Following a literature review to identify genes at which TGS was induced, 14 gene promoters were analysed based on features within and surrounding the associated TSRs, comprising TSR length and CpG class, presence of TBP- and Sp1-binding sites, RNA Pol II, nucleosome and histone variant H2A.Z occupancy, ratio of RNA Pol II occupancy in resting versus activated cells, the number of TSCs per TSR, the combined epigenetic score (CES) and tiRNA and ncRNA presence. In addition, levels of 19 histone methylation and 18 histone acetylation modifications were examined. In total, data on 49 promoter features were retrieved, all from previously published, peer-reviewed studies. Initially, the TGS group was compared to a random set of 2,000 TSRs and then all other TSRs in the genome. To control for gene activation, two further analyses were performed comparing the TGS group features to those active in the THP-1 cell line, in which the FANTOM4 data was obtained, and housekeeping genes.

5.2 Materials and methods

5.2.1 Selection of small RNA-induced transcriptionally-silenced genes and TSR assignment

An extensive literature review was conducted to identify genes reportedly transcriptionally silenced by promoter-targeted small RNAs. Inclusion in the study as a transcriptionally-silenced gene required compliance with a number of criteria:

(a) Induced silencing in human cells (determined by qRT-PCR and/or Western blot);

(b) RNA guide strand <30 nt;

(c) Specific silencing (ie use of controls to show no off-target silencing);

(d) Accompanied by epigenetic changes (DNA methylation and/or enrichment for silent-state histone modifications); and,

(e) 5' end of the target site within 1,500 bp upstream and 1,000 bp downstream of the RefSeq TSS;

(f) 5' end of the target site within 1,500 bp upstream and 200 bp downstream of the most active TSS of the assigned TSR (see below).

Application of this selection criteria resulted in a list of 14 genes (Table 5.2) that composed the transcriptional gene-silenced (TGS) group.

Each gene within the TGS group was assigned a TSR: the TSR nearest the RefSeq TSS in the same orientation as the gene was designated the TSR for that transcriptionally-silenced gene.

5.2.2 Promoter feature data sources and quantification

To assess whether particular promoter architectures are associated with susceptibility to small RNA-induced transcriptional gene silencing, different features of promoters were examined. All analysis used the NCBI36/hg18 (Mar 2006) human genome (Lander *et al.* 2001).

TSR data, including TSR length, CpG class and the number of TSCs per TSR (Balwierz *et al.* 2009), were obtained from FANTOM data (Kawaji *et al.* 2006b; Kawaji *et al.* 2009; Kawaji *et al.* 2010) downloaded from the SwissRegulon Portal (http://www.swissregulon.unibas.ch/cgi-bin/regulon?page=downloads) (Pachkov *et al.* 2007). A Bayesian procedure was used to classify each TSR as either high- or low-CpG class, based on both G or C and CpG dinucleotide content (Balwierz *et al.* 2009). In this study, those TSRs of high-CpG class were scored one and those of low-CpG class scored zero. The number of TSRs analysed following subtraction of those belonging to the TGS group were 43,159.

The most active TSS within a TSR, or representative TSS, was identified from the genome location given in the TSR ID. The number of Sp1 binding sites and TATA boxes (Table 4.1) were determined in a 2,000 bp window centred on the representative TSS of each TSR. Consensus sequences were determined from the regulatory motif weight matrices available on the SwissRegulon Portal (http://www.swissregulon.unibas.ch/cgi-bin/regulon?page=downloads) (Pachkov *et al.* 2007).

Nucleosome and RNA Pol II occupancy data was obtained from human CD4+ T cells (http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcellnucleosomes.aspx) (Schones *et al.* 2008). The average nucleosome and Ser5P-Pol II occupancy within the representative TSS-centred 2,000 bp window was calculated based on the data obtained from cells activated with anti-CD3 and -CD28 antibodies. The log of the activated:resting CD4+ T cells Ser5P-Pol II ratio within the same window is also reported.

Table 5.1: Histone modifications quantified in a 2000 bp window centred on the representative TSS.

Histone methylations	Histone acetylations
H2BK5me1	H2AK5ac
H3K27me1	H2AK9ac
H3K27me2	H2BK120ac
H3K27me3	H2BK12ac
H3K36me1	H2BK20ac
H3K36me3	H2BK5ac
H3K4me1	H3K14ac
H3K4me2	H3K18ac
H3K4me3	H3K23ac
H3K79me1	H3K27ac
H3K79me2	H3K36ac
H3K79me3	H3K4ac
H3K9me1	H3K9ac
H3K9me2	H4K12ac
H3K9me3	H4K16ac
H3R2me1	H4K4ac
H3R2me2	H4K8ac
H4K20me1	H4K91ac
H4K20me3	

Histone (H) number is followed by either a lysine (K) or arginine (R) residue and, finally, the histone modification: methylation (me), dimethylation (me2), trimethylation (me3) or acetylation (ac).

By the same ChIP-Seq method, data on histone variant H2A.Z occupancy (Barski *et al.* 2007) and the frequency of a number of histone modifications were obtained for human CD4+ T cells. The genome-wide distribution of 19 arginine and lysine histone methylations (Table 5.1) (http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcell.aspx) (Barski et al. 2007) as well as 18 histone acetylations (Table 5.1) (http://dir.nhlbi.nih.gov/papers/lmi/epigenomes/hgtcellacetylation.aspx) (Wang *et al.* 2008), were retrieved. The average number of reads in a representative TSS-centred 2000 bp window was calculated for each TSR.

CpG islands, genomic regions characterised by exceptionally high CpG dinucleotide frequency, are known to associate with active promoters (Kim *et al.* 2005). According to the Gardiner-Garden sequence criteria (Gardiner-Garden and Frommer 1987), a genomic region has to fulfil three conditions to classify as a CpG island: (1) GC content >50%; (2) ratio of observed-to-expected number of CpG dinucleotides above 0.6; and, (3) length greater than 200 bp. However, this sequence only-based definition does not distinguish between CpG islands that function as transcriptional regulators from those lacking characteristics of promoter activity. The combined epigenetic score (CES), which predicts the likelihood of a CpG island being bona fide, summarises the inherent tendency of a CpG island, across different cell types and tissues, to exhibit an unmethylated, open and transcriptionally competent chromatin structure (Bock *et al.* 2007). Thus, this score facilitates discrimination between regions that are just CpG-rich and those that exhibit promoter activity. The CESs, which may be visualised as a CGI scoring track on the UCSC Genome Browser (http://genome.ucsc.edu/) (Kent et al. 2002), were downloaded (http://epigraph.mpi-inf.mpg.de/download/CpG_islands_revisited/) and included in the analysis where a CpG island overlapped the TSR. A CES value of zero corresponds to a completely silenced, inactive and buried CpG island, while a value of one corresponds to an unmethylated, highly accessible CpG island with strong promoter activity.

Promoter-associated transcripts have been implicated in small-RNA transcriptional gene silencing. Therefore, the occurrence of ncRNAs was examined at TSRs. Human ncRNA data was retrieved from the Functional RNA Database (http://www.ncrna.org/frnadb/download) and the frequency of ncRNAs within a TSR recorded. A ncRNA was scored as present even if the ncRNA did not fall entirely within the 2,000 bp window centred on the representative TSS of a TSR, and vice versa. In addition, the location of a ncRNA was not strand-specific, such that ncRNAs antisense to the TSR were also included.

A subset of ncRNAs, tiRNAs, was recently shown to colocalise to TSSs and splice sites in metazoans (Taft *et al.* 2009; Taft *et al.* 2010). These ~18 nucleotide RNAs where identified using data obtained by the FANTOM4 project (Kawaji *et al.* 2009; Kawaji *et al.* 2010), focusing on small

RNAs (~15 to ~40 nucleotides) from undifferentiated human THP-1 cells and at four time points (<96 hours) after stimulation with phorbol 12-myristate 13-acetate (PMA) (http://fantom.gsc.riken.jp/4/download/Supplemental_Materials/Taft_et_al_2009/) (Taft *et al.* 2009). Given the association of tiRNAs to TSSs, their frequency at 2,000 bp windows encompassing TSRs was examined.

5.2.3 Assignment of TSRs to the THP-1 dataset

The FANTOM4 project included measurement of the expression level of ~2,400 genes using a Illumina Human-6 v2 chip in THP-1 cells at various time points following stimulation with PMA (Taft *et al.* 2009). This expression data was deposited in the Center for Information Biology Gene Expression Database (CIBEX), Accession CBX46 (http://cibex.nig.ac.jp/cibex2/ExperimentMiame.do?queryExperimentalDesignAccession=CBX46). All genes were sorted by their measured expression levels at 12 hours post-PMA stimulation. The top 25% were considered 'active' and thus included in the THP-1 dataset.

Gene expression data was listed by gene symbol as an identifier. To map the set of genes to a TSR, the gene symbols were first translated into RefSeq mRNA IDs. Coordinates of the RefSeq mRNAs were then obtained for the human genome build NCBI36/hg18 (Mar 2006) from the UCSC genome browser (http://genome.ucsc.edu/) (Kent *et al.* 2002). Genes whose RefSeq TSS overlapped a TSR were assigned that TSR. The result was a list of TSRs (the THP-1 dataset) with an associated expression level (n = 351).

5.2.4 Assignment of TSRs to the housekeeping gene dataset

Housekeeping (HK) genes are defined as genes constitutively expressed in all tissues to maintain essential cellular functions. Experimental determination of HK genes suffers from a number of limitations and so HK genes have recently been identified based on physical characteristics of gene length and chromatin compactness. The HK genes used in this study were mined based on combining such characteristics to create a powerful Naïve Bayes classifier (De Ferrari and Aitken 2006). Each Ensemble transcript ID was assigned with a probability of being a HK gene. These transcript IDs could be readily mapped to a TSS and thus, to a TSR where overlap occurred. Only where the Naïve Bayes classifier was >90% confident was a gene included in the HK dataset (n = 737).

5.2.5 Classification of *htatsf1* and *ddx3x* promoter features

To examine which group the promoter features of *htatsf1* and *ddx3x* more closely resemble, representative feature vectors (RFV) were created for each group, namely TGS, all other TSRs, THP-1 and HK genes. Briefly, the values for each feature within a group were averaged to give a single feature vector for each group. For example, if a group was composed of three TSRs/vectors and defined by four features, a RFV would be calculated as follows:

Vector/TSR	Feature1	Feature2	Feature3	Feature4
A	5.0	22.3	4.4	5.3
B	6.6	18.7	3.3	4.3
C	7.4	14.1	0.8	7.3
D	2.3	25.5	0.9	3.4
RFV	5.3	20.2	2.4	5.1

The Euclidean distance of a given vector can then be calculated from the RFV. This involves summing the squared difference between the given vector and the RFV and taking the square root:

$$d(\mathrm{RFV},\mathrm{X}) = \sqrt{\sum_{f}^{n}\left(\mathrm{RFV}_f - \mathrm{X}_f\right)^2} \qquad (5.1)$$

where X is an unknown vector, *d* is the Euclidean distance, *n* is the number of features and *f* is the series of average feature values within the vector. Thus, the smaller the Euclidean distance, the closer the given feature set is to the average feature set of a group.

5.2.6 Statistical analyses

For each promoter feature, a set of values was obtained for each group. The median for each distribution was calculated and plotted, along with standard deviations. TSR lengths (and standard deviations) were scaled by 0.1 to enable the inclusion of all features on a single plot. A Mann-Whitney U test was performed to give the probability that the medians for the distributions are the same. When that probability was small (p <0.05), the difference in the medians was considered to be significant (*). The Benjamini and Hochberg correction was used to correct for multiple testing. This correction converts *p*-values to false discovery rates (fdr). A conservative fdr threshold of 0.05 was used, thus features with a calculated fdr <0.05 were considered to have

significantly different medians (*). A chi-square test was used to compare fractions between groups.

5.2.7 Statement on contributions

The work in this chapter is the result of collaboration between Loren Hansen, a PhD student supervised by Dr David Landsman, Computational Biology Research Branch, NCBI, NIH, Bethesda, USA, and myself. The concept behind the project is entirely my own and I sourced all the data used in the analyses. I conducted the literature review, compiled the TGS group and assigned them TSRs. Loren contributed by writing the programs (using Perl script) needed to retrieve the information I requested from the data sources I had identified. Loren also assigned TSRs to both the THP-1 and HK gene datasets, calculated medians, RFVs and Euclidean distances and performed the statistical tests. Loren also engaged in helpful discussions regarding the interpretation of the data.

5.3 Results

5.3.1 Identification of small-RNA induced transcriptionally-silenced TSRs

To address whether TGS-susceptible genes have particular promoter characteristics, small RNA-induced transcriptionally-silenced genes had first to be identified. A literature review was conducted to identify all reports of mammalian TGS. Criteria were applied resulting in a list of 14 genes that were used in the study as the transcriptional gene-silenced (TGS) group (Table 5.2). The criteria, detailed in 5.2.1, ensured that only human genes targeted by a guide strand less than 30 nt long (eg siRNA, shRNA, miRNA), at the 5' end near a TSR, resulting in specific silencing accompanied by epigenetic changes, were included in the study.

Epigenetic changes at the target site were included in the selection criteria as the focus was to identify promoter characteristics associated with classical TGS. TGS by this mechanism has the potential for long-term silencing providing silent-state chromatin is achieved; an advantage over PTGS as continual administration of the silencing small RNA is, therefore, unlikely to be necessary. The long-term silencing potential of TI has not yet been examined, and so notable exceptions from the TGS group include c-Myc (Entrez gene ID 4609) (Napoli *et al.* 2009). Progesterone receptor (PGR) was included despite initial reports of silencing by a TI mechanism unaccompanied by epigenetic changes (Janowski *et al.* 2005; Janowski *et al.* 2006; Janowski *et al.* 2007; Schwartz *et al.* 2008; Chu *et al.* 2010). This is due to the most recent report of PGR transcriptional-modulation showing enrichment for H3K27 trimethylation at the target site (Yue *et al.* 2010).

This study uses the TSR as the basic identifier, because of its importance in classification of different promoter classes (Balwierz *et al.* 2009) and the large amount of data already analysed for this feature. Therefore, two genes, CCR5 (Entrez gene ID 1234) (Kim *et al.* 2006; Han *et al.* 2007) and interleukin-2 (IL2; Entrez gene ID 3558) (Murayama *et al.* 2006), were excluded from the TGS group because no TSRs were present at the 5' end of these genes.

Only genes transcriptionally-silenced by small RNAs in human cells were included. Therefore, despite convincing evidence for TGS, vascular endothelial growth factor (VEGF-A) (Turunen *et al.* 2009) and transforming growth factor-beta receptor II (TGFβII) (Kim *et al.* 2007) were excluded as silencing was only reported in mouse and rat cell lines, respectively.

The final TGS group is composed of genes with a range of functions found on different chromosomes and silenced by different types of small RNAs in different cell lines (Table 5.2). These genes were assigned TSRs based on proximity to the 5' end of the RefSeq gene.

Table 5.2: Transcriptionally-silenced genes included in the study and the assigned TSRs.

Identifiers				Target			TSR assignment		Reported context			
Name	Gene	Transcript	Locus	nt	Start	End	TSR ID	Target-TSS	Effector	Cell line	Evidence	References
Eukaryotic translation elongation factor 1 alpha	1915	NM_001402	6q14.1(-)	21	74287582	74287562	TSR_hg18_v1_chr6_-_74287478	-104	si	HEK293T	DNA me; Histone me	(Morris *et al.* 2004; Weinberg *et al.* 2006; Han *et al.* 2007)
E-cadherin	999	NM_004360	16q22.1(+)	21	67328515 67328633	67328535 67328653	TSR_hg18_v1_chr16_+_67328654	-139 -21	si	HCT116 & MCF7	Histone me; Run-on	(Ting *et al.* 2005)
Ras association domain family 1	11186	NM_170714 NM_007182	3p21.3(-)	22	50353327	50353306	TSR_hg18_v1_chr3_-_50353286	-41	sh	HeLa	DNA me; Histone me	(Castanotto *et al.* 2005; Kim *et al.* 2006)
Plasminogen activator, urokinase	5328	NM_002658	10q24(+)	21	75340680 75340762 75340782 75340843	75340700 75340782 75340802 75340863	TSR_hg18_v1_chr10_+_75340716	-36 +46 +66 +127	Si[1]	PC3	DNA me; Run-on	(Pulukuri and Rao 2007)
Cyclin-dependen kinase inhibitor 2A isoform 3/p16	1029	NM_058197 NM_000077	9p21(-)	19	21984852	21984834	TSR_hg18_v1_chr9_-_21964857	+5	si	HEK293T	Histone me	(Wang *et al.* 2007)
Polymerase (RNA) III polypeptide D	661	NM_001722	8q21(+)	23	22158432	22158454	TSR_hg18_v1_chr8_+_22158588	-156	mi	HEK293T	Histone me; Run-on	(Kim *et al.* 2008)
Cyclin-dependent kinase inhibitor 2B/p15	1030	NM_078487 NM_004936	9p21(-)	19	22000477 22000351 22000092	22000459 22000333 22000074	TSR_hg18_v1_chr9_-_21999305	-1172 -1046 -787	Si[2]	HEK293T	Histone me	(Gonzalez *et al.* 2008)
Tubulin folding cofactor E-like	219899	NM_152715	11q23.3(+)	23	120399096	120399118	TSR_hg18_v1_chr11_+_120400020	-924	mi	HCT116	Histone me	(Gonzalez *et al.* 2008)
Ras p21 protein activator 2	5922	NM_006506	3q22-3q23(+)	23	142687289	142687311	TSR_hg18_v1_chr3_+_142688572	-1283	mi	HCT116	Histone me	(Gonzalez *et al.* 2008)
Rhophilin, Rho GTPase binding protein 2	85415	NM_033103	19q13.11(-)	23	38249138 38248785	38249116 38548763	TSR_hg18_v1_chr19_-_38247635	-1503 -1150	mi[3]	HCT116	Histone me	(Gonzalez *et al.* 2008)
Wolf-Hirschhorn syndrome candidate 1	7468	NM_001042424	4p16.3(+)	23	1841771 1842021 1842821	1841793 1842043 1842843	TSR_hg18_v1_chr4_+_1842766	-995 -755 -55	mi[3]	HCT116	Histone me	(Gonzalez *et al.* 2008)
Ubiquitin C	7316	NM_021009	12q24.3(-)	21	123965316	123965296	TSR_hg18_v1_chr12_-_123965150	-166	sh	HEK293GT	DNA me; Histone me; Run-on	(Hawkins *et al.* 2009)
Homeobox D4	3233	NM_014621	2q31.1(+)	23	176723303	176723325	TSR_hg18_v1_chr2_+_176724543	-1240	si & mi	MCF7; & MDA-MB-231	DNA me; Histone me; Run-on	(Tan *et al.* 2009a)
Progesterone receptor	5241	NM_000926	11q22-11q23(-)	19	100505763	100505745	TSR_hg18_v1_chr11_-_100505741	-22	si	T47D	Histone me	(Yue *et al.* 2010)

Fourteen transcriptionally-silenced genes met the criteria for inclusion in the study and were, therefore, assigned TSRs which composed the transcriptional gene-silenced (TGS) group. Target-TSS gives the nt distance of the 5' end of the target site(s) from the representative TSS of the assigned TSR. [1]Four siRNAs were pooled in all experiments. [2]Transcriptional silencing of this gene also reported with an endogenous asRNA (Yu et al. 2008). [3]Multiple target sites for a single guide strand within the target promoter ie multiple miRNA seed sequence matches.

5.3.2 The TGS group are enriched for promoter features associated with active transcription compared with 2000 random TSRs

To examine whether promoter features conferred susceptibility to TGS, the promoter features of the TGS group required comparison to a control group. Unfortunately, reports of failed TGS induction do not exist. There is no incentive to publish 'negative' data and it is difficult to exclude the possibility that the appropriate target site within the promoter was not found. Without a TGS-negative control group, the TGS group (*n* = 14) promoter features were examined for enrichment compared to 2,000 random TSRs (*n* = 2000).

TSRs within the TGS group were significantly longer and of a higher CpG class than the 2,000 random TSRs (Figure 5.2). There was also a higher CES in the TGS group and enrichment for Sp1 binding sites. Together this provides confirmation that the data retrieval and analysis was functioning because broad TSRs are associated with high GC and CpG content (Balwierz *et al.* 2009) and so would be expected to also be associated with high CpG class, high CES scores (which predict the probability of a bona fide CpG island) and increased Sp1 binding sites, which have previously been shown to be over-represented in broad promoters (Carninci *et al.* 2006). There were significant enrichments for H3K4me3 and H4K20me1 in the TGS group (Figure 5.2). In contrast, levels of H3K27me1 were significantly diminished. Increased H3K4me3 and H4K20me1 and decreased H3K27me1 is associated with active gene expression at human promoters (Wang *et al.* 2008). Concordant with this observation is that nucleosome occupancy was significantly decreased in the TGS group compared with the 2,000 random TSRs (Figure 5.2). The enrichment for promoter features associated with active transcription in the TGS group is unsurprising given that only expressed genes are targeted for TGS. Whether the TGS genes are active in the cell lines and conditions that the promoter feature data was extracted from (eg THP-1 and CD4+ T cells) is unknown. However, given that the TGS genes are active in at least one cell line and seen as potential therapeutic targets, it is a reasonable assumption that a higher fraction of TGS genes are likely to be active in the cell lines and conditions under which the data is extracted compared to 2,000 random TSRs. Other features associated with active transcription were increased in the TGS group, although not significantly. These included H2A.Z occupancy, a histone variant associated with the +1 nucleosome at actively transcribed genes (Verdel *et al.* 2004; Sarma and Reinberg 2005; Wang *et al.* 2008), and the elongating form of RNA Pol II, Ser5P-Pol II.

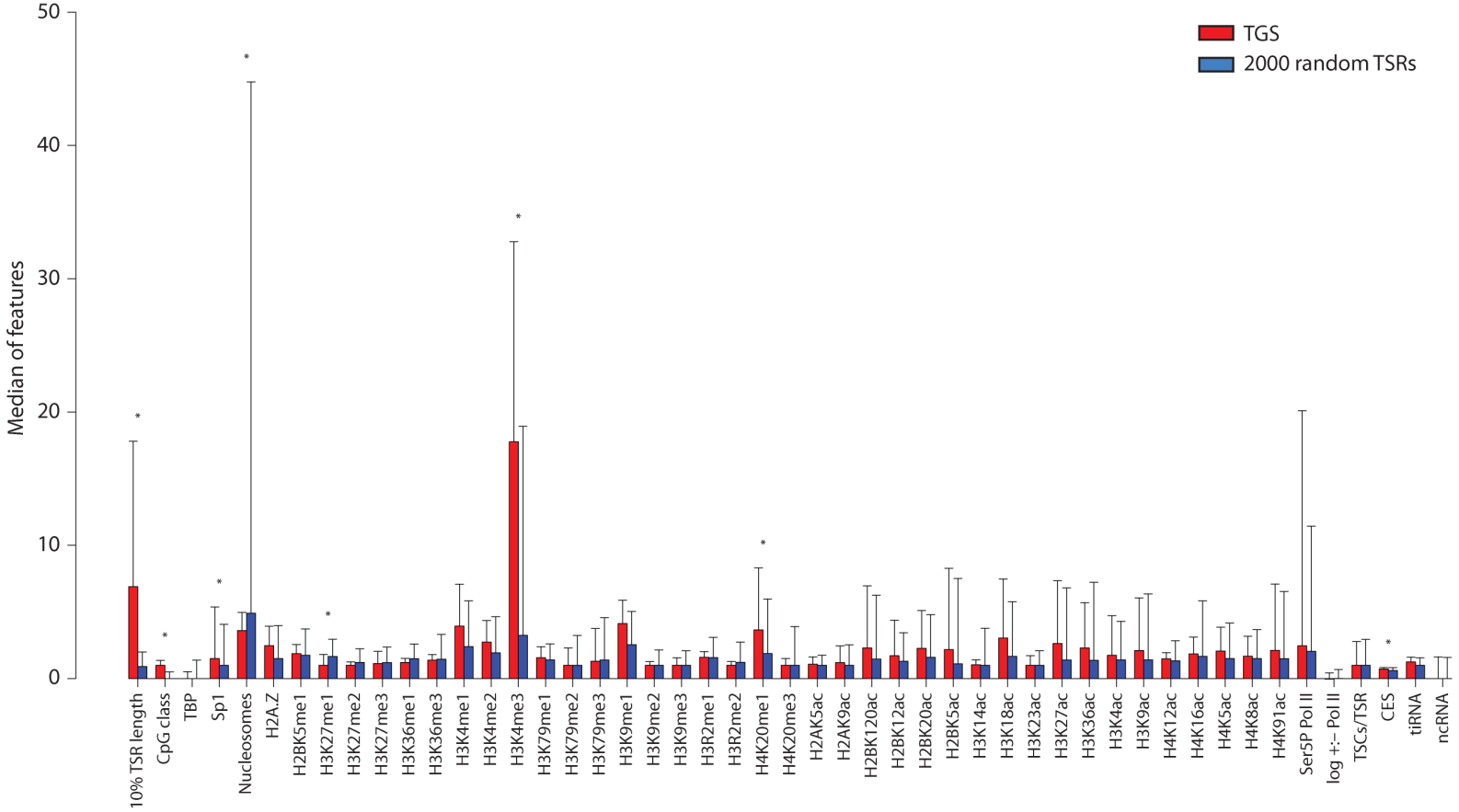


Figure 5.2: The transcriptionally-silenced group promoter features compared with those of 2000 random TSRs. TSR lengths were multiplied by 0.1. Feature values were determined for each TSR or, where appropriate, within a 2,000 bp window centred on the representative TSS of each TSR. Error bars indicate standard deviations. TGS group $n = 14$. *, $p < 0.05$, Mann-Whitney U test.

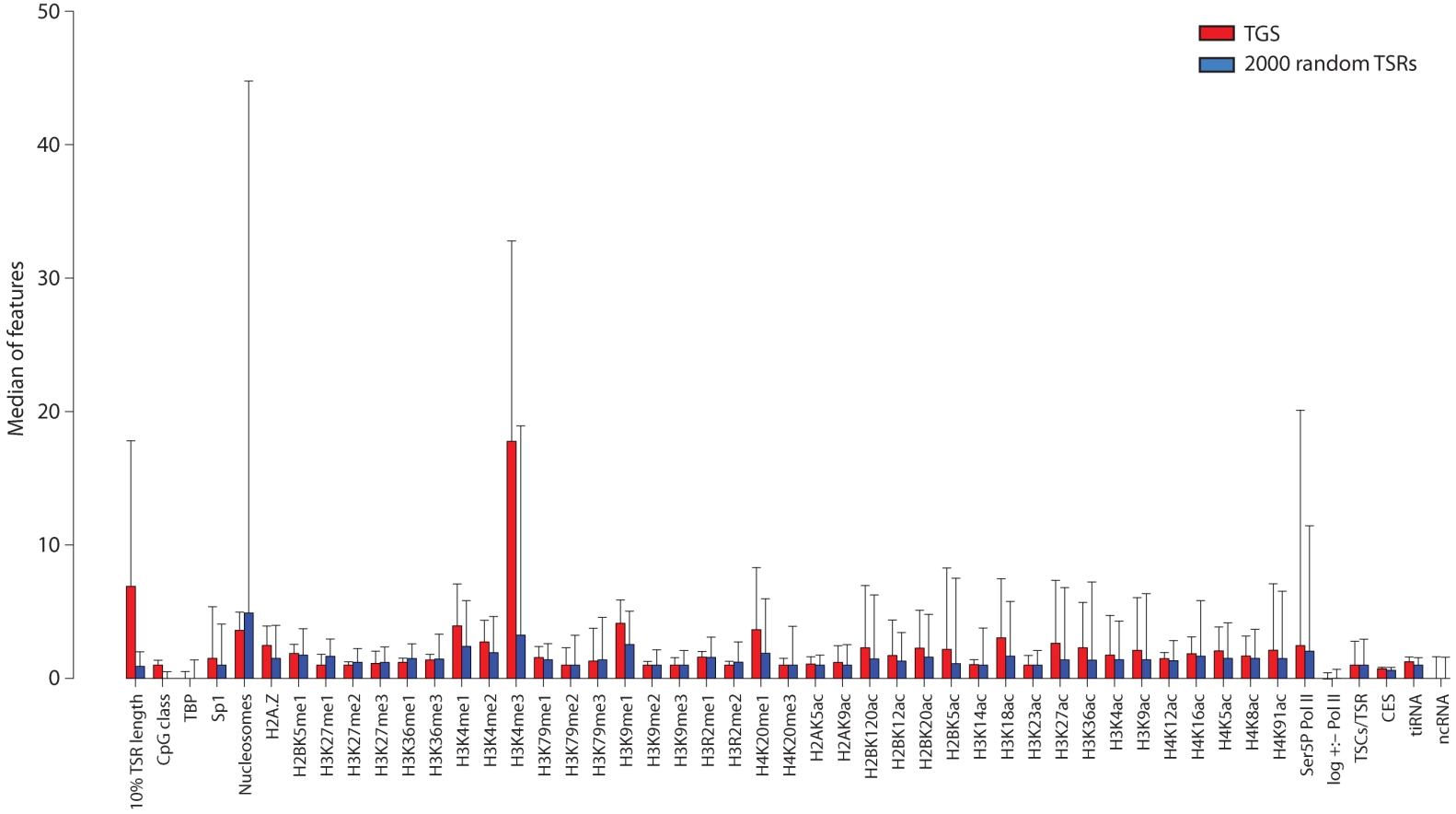


Figure 5.3: The transcriptionally-silenced group promoter features compared with those of 2000 random TSRs and corrected for multiple testing. TSR lengths were multiplied by 0.1. Feature values were determined for each TSR or, where appropriate, within a 2,000 bp window centred on the representative TSS of each TSR. TGS group *n* = 14. Error bars indicate standard deviations. *, fdr <0.05, Benjamini and Hochberg correction.

Statistical significance was initially determined using a Mann-Whitney U test (p <0.05). However, this does not correct for multiple testing. Therefore, the Benjamini and Hochberg correction was applied and statistical significance instead determined on the basis of false discovery rates (fdr <0.05). Following fdr-correction, none of the promoter features significantly differ between the TGS group and the 2,000 random TSRs (Figure 5.3). This is likely to be false as correcting for multiple testing tends to be excessively conservative. Nevertheless, correcting for multiple testing is better than not, and the remainder of the analyses were corrected in this fashion.

5.3.3 The TGS group are enriched for promoter features associated with active transcription compared with all other human TSRs

Having demonstrated that the data extraction and analysis was working well, the TGS group of TSRs (n = 14) was compared to all other TSRs annotated in the human genome (n = 43,159; Figure 5.4). As expected, the results closely mimicked those obtained when the TGS group were compared to 2,000 random TSRs. In other words, there were no significant differences in promoter features between the two groups following fdr-correction. There was increased TSR length, CpG class, CES score and Sp1-binding sites in the TGS group and an enrichment of features associated with active transcription, including increased H3K4me3, H4K20me1 and H2A.Z occupancy and decreased H3K27me1 and nucleosome occupancy.

Using the same rationale described in 5.3.2, it is not surprising to see enrichment for features associated with active transcription in the TGS group compared to all other TSRs in the human genome. However, this is not helpful in identifying promoter features associated with TGS susceptibility. Therefore, the TGS group TSRs were compared to those active in the THP-1 cell line, from which FANTOM4 data was obtained.

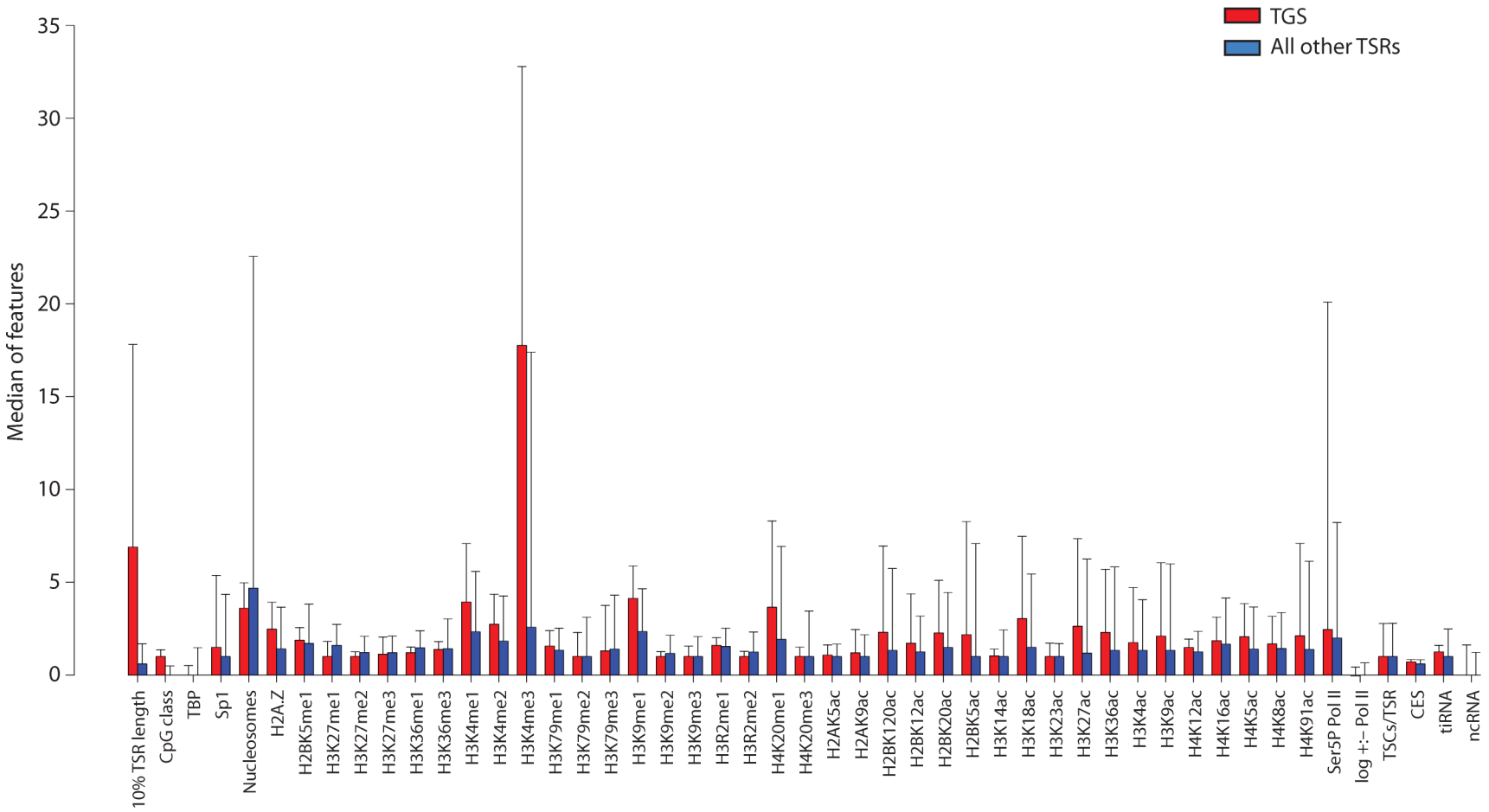


Figure 5.4: The transcriptionally-silenced group promoter features compared with all other TSRs. TSR lengths were multiplied by 0.1. Feature values were determined for each TSR or, where appropriate, within a 2,000 bp window centred on the representative TSS of each TSR. Error bars indicate standard deviations. TGS group *n* = 14; all other TSRs *n* = 43159. *, fdr <0.05, Benjamini and Hochberg correction.

5.3.4 The TGS group show reduced TSR length and number of TSCs per TSR compared with active TSRs in the THP-1 cell line

THP-1 cells are a macrophage-like cancer cell line used in the fourth stage of the FANTOM project. Thus, FANTOM4-derived data, such as the number of tiRNAs, were determined in this cell line. Other FANTOM data, including TSR length, CpG class and the number of TSCs per TSR, were designated on the basis of 56 human CAGE-tag libraries from different tissues (Balwierz *et al.* 2009). However, 18 of these were from FANTOM4 THP-1 samples, introducing a heavy bias toward this cell line. To see if the same promoter features were enriched when the TGS group TSRs were compared to only active TSRs, TSRs associated with expressed genes in the THP-1 cell line were identified – from THP-1 microarray expression data, the top 25% of expressed RefSeq genes, 12 hours post-PMA stimulation, with TSRs overlapping the RefSeq TSS were included (n = 351).

In contrast to the previous analyses, the TSR length of the TGS group was significantly lower than the THP-1 group and the CpG class of both groups was similar (Figure 5.5). There was also little difference in CES between the two groups, although there were significantly less TSCs per TSR in the TGS group compared with the THP-1 group. This suggests that genes active in the THP-1 cell line have proximal alternate promoters, which may be used during different stages of cell activation, as well as different tissues. Several overlapping TSCs may account for the increased length in TSR observed in the THP-1 group.

Several histone modifications associated with active transcription were enriched in the THP-1 group, namely H3K36me1, H3K4me2, H3K4me3, H3K79me1, H3K79me2, H3K79me3, H2BK5ac, H3K27ac, H3K9ac, H4K8ac and H4K91ac (Figure 5.5). In addition, the silent-state histone modification H3K27me3 was enriched in the TGS group compared to the THP-1 group. In contrast, the silent-state histone modification H3K27me1 was also enriched in the THP-1 group. However, the proportion of TSRs in the THP-1 group are more likely to be associated with H3K27me1 than H3K27me2 and H3K27me3, which are more strongly associated with gene silencing. Indeed, H3K27me3 is at a significantly lower level in the THP-1 group than the TGS group, and H3K27ac, which is associated with active transcription, is at a higher level. On the whole, therefore, histone modifications associated with active transcription are enriched in the THP-1 group compared to the TGS group. Indeed, most of the acetylation modifications associated with active TSSs (H2AK9ac, H2BK5ac, H3K18ac, H3K27ac, H3K36ac, H3K9ac and K4K91ac) (Wang *et al.* 2008) were significantly enriched in the THP-1 group, whilst those associated with promoters and actively transcribed regions (H2BK120ac, H2BK20ac, H2BK12ac, H3K4ac, H4K12ac, H4K16ac, H4K5ac, H4K8ac) (Wang *et al.* 2008), though increased, were not enriched significantly, with the exception

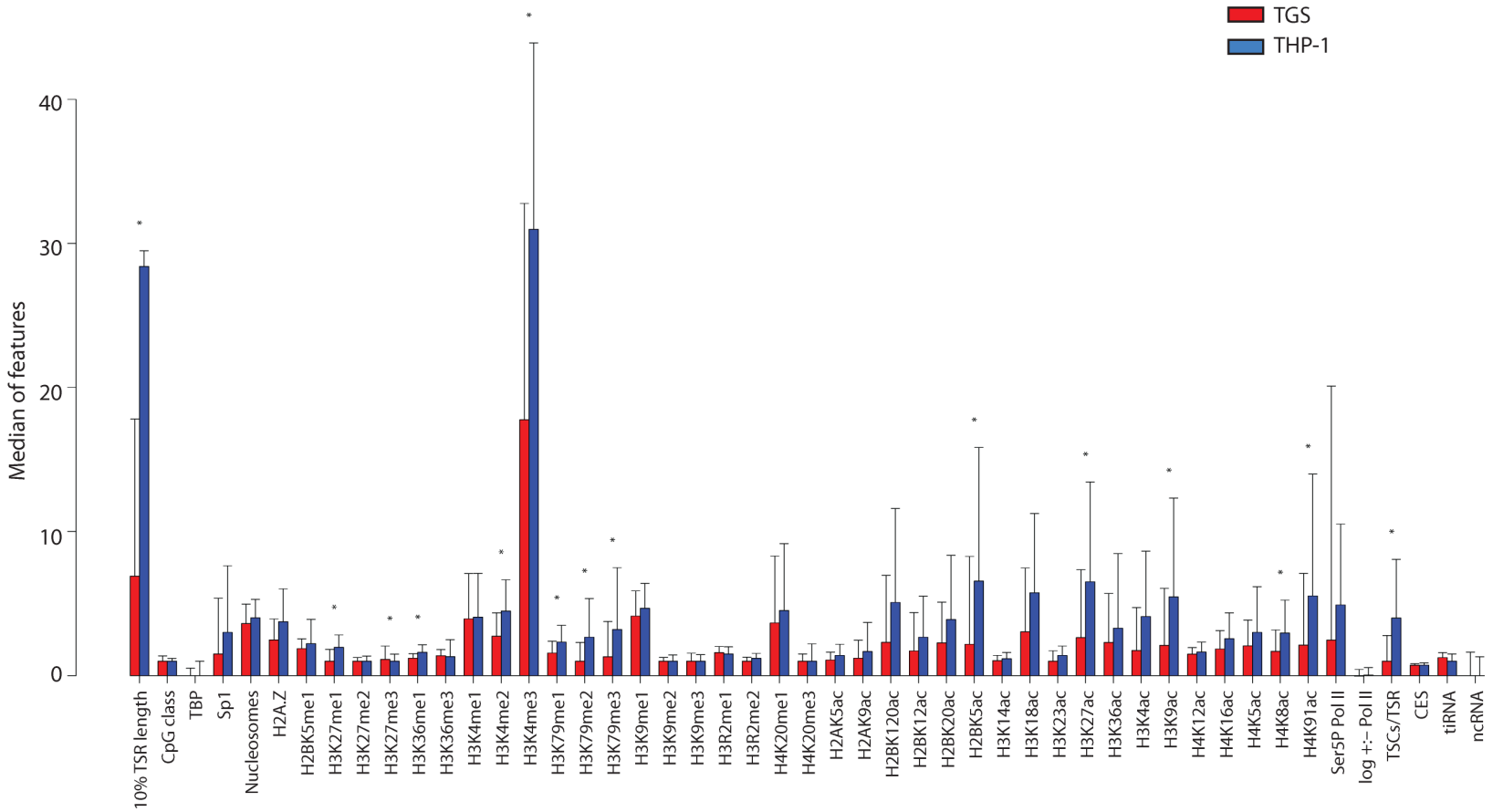


Figure 5.5: The transcriptionally-silenced group promoter features compared with TSRs assigned to genes expressed in THP-1 cells. TSR lengths were multiplied by 0.1. Feature values were determined for each TSR or, where appropriate, within a 2,000 bp window centred on the representative TSS of each TSR. Error bars indicate standard deviations. TGS group *n* = 14; THP-1 dataset *n* = 351. *, fdr <0.05, Benjamini and Hochberg correction.

of H4K8ac. In addition, methylation modifications associated with active promoters (H3K27me1, H3K4me2, H3K4me3, H3K79me1, H3K79me2, H3K79me3, H3K9me1) (Wang *et al.* 2008) were all significantly enriched in the THP-1 group, with the exception of H3K9me1. This demonstrates that assignment of the TSRs to the THP-1 dataset was successful, as active TSS-associated features would be expected to be over-represented in the representative TSS-focused windows used in the analyses. Enrichment for histone modifications associated with active TSSs in the THP-I group also suggests that the majority of the top 25% of expressed genes in THP-1 cells are also expressed in CD4+ T cells, the cell line from which the histone modification data was obtained. Using the THP-1 dataset as a control for active genes, therefore, is reasonable, despite not all the promoter feature data being acquired from these cells.

In contrast to the previous plots, the THP-1 promoter feature values are generally higher than the TGS group, which again probably reflects the difference in the proportion of active genes between these two groups. This makes it difficult to claim the differences seen are due to some genes having the ability to be transcriptionally silenced while other genes are not. Some features are less dependent on gene activity, ie TSR length, CpG class, TBP- and Sp1-binding sites, TSCs per TSR, CES and tiRNA and ncRNA levels, although there presence/levels are still dependent on the tissue/cell line(s) used in data acquisition. To examine whether differences exist in these characteristics between the TGS group and active genes in all tissues, comparisons were made to housekeeping gene-associated TSRs.

5.3.5 The TGS group are enriched for promoter features associated with reduced activity compared with housekeeping gene TSRs

Housekeeping (HK) genes were used as a control group not just because by definition they are expressed in all tissues, but also because they are a set of genes most unlikely to be susceptible to TGS. This is because silencing them would likely result in cell death. HK genes were identified based on a number of physical properties (De Ferrari and Aitken 2006). Those that could be assigned TSRs (where a TSR overlapped the 5' end of the RefSeq gene) were included in the HK gene group (n = 737).

There was significant enrichment for the same histone modifications in the HK gene group as the THP-1 group, plus additional histone modifications associated with active promoters (Figure 5.6). Again this may be attributable to differences in the proportion of active TSRs between the groups. Although the proportion of active TSRs in the HK gene and THP-1 groups in CD4+ T cells are both expected to be close to 100%, TGS group TSRs are more likely to be active in the CD4+ T cell line than constitutively active ie a HK gene. It is therefore unsurprising that there

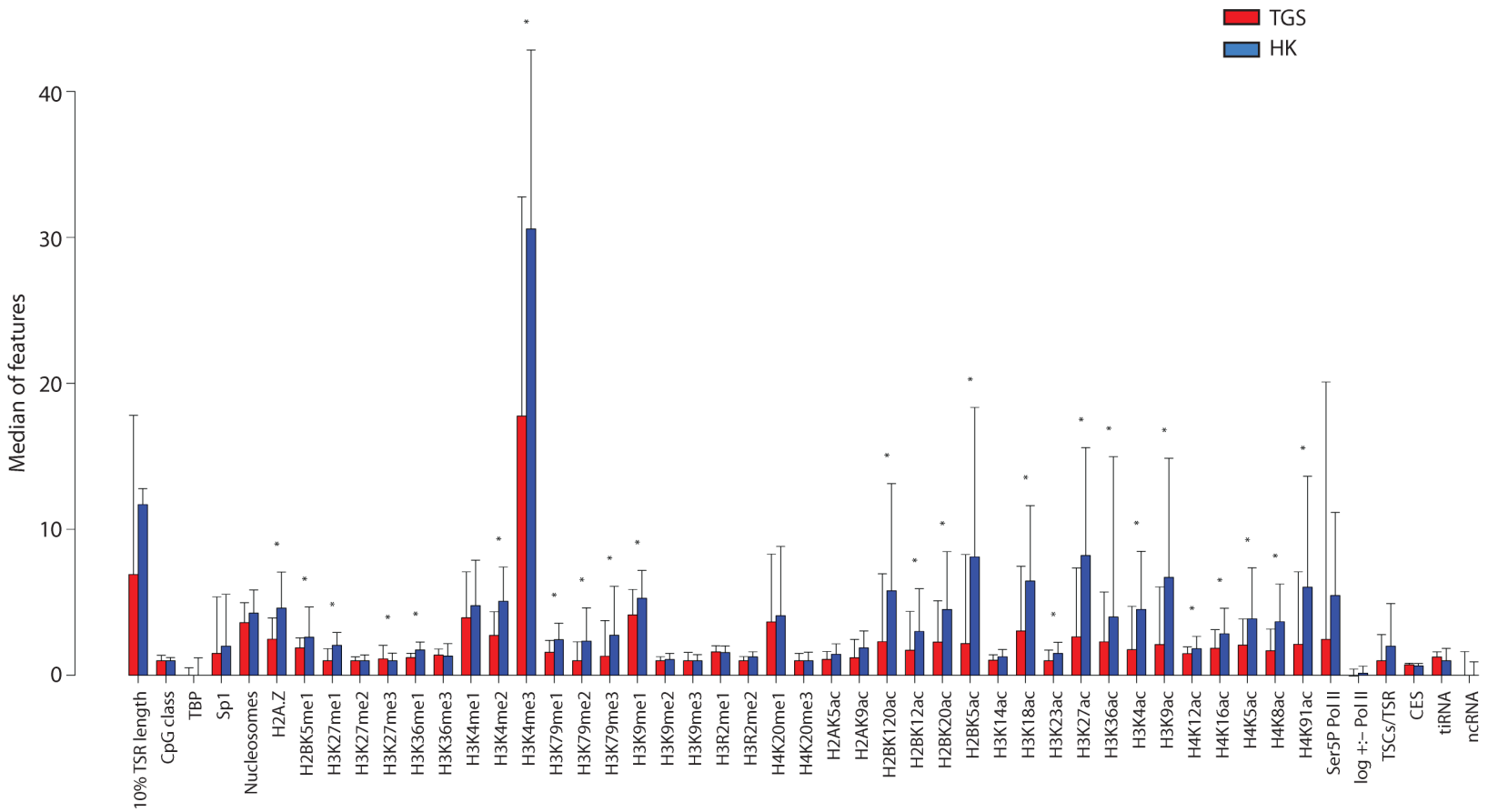


Figure 5.6: The transcriptionally-silenced group promoter features compared with TSRs assigned to housekeeping genes. TSR lengths were multiplied by 0.1. Feature values were determined for each TSR or, where appropriate, within a 2,000 bp window centred on the representative TSS of each TSR. Error bars indicate standard deviations. TGS group *n* = 14; housekeeing (HK) genes *n* = 737. *, fdr <0.05, Benjamini and Hochberg correction.

is a greater difference in active transcription-associated promoter feature values between the HK gene group and the TGS group than the THP-1 group and the TGS group.

With respect to promoter features less dependent of gene activity, the median TSR length and TSC number per TSR were lower in the TGS group than the HK gene group (Figure 5.6), although not statistically significant as was the case with the THP-1 group (Figure 5.5).

To summarise, the proportion of active genes within the respective groups dominates differences in histone modifications and H2A.Z, nucleosome and RNA Pol II occupancy. Thus, inspection of features independent of gene activation may prove more informative with respect to identification of TGS-associated promoter features. TGS group TSR length is greater than the average in the genome but smaller than TSRs associated with HK genes or those active in the THP-1 cell line. In addition, the bulk of TGS group TSRs contain only one TSC, which is the case for the majority of TSRs in the genome, but not those in the THP-1 and HK gene groups. The number of Sp1-binding sites, TSR CpG class and CES for TGS group TSRs are greater than the average TSR but similar to the THP-1 and HK gene groups. Together, this demonstrates that TGS susceptible genes have broad, high CpG class TSRs containing bona fide CpG islands and Sp1-binding sites. However, the TSRs are not as broad as those comprised of multiple TSCs. This suggests that TGS-susceptible genes are characterised by broad TSRs with a single TSC, which may be indicative of a single, broad NDR.

5.3.6 The fraction of TSRs with ncRNA overlap is greatest in the TGS group

A hypothesis stated in 5.1 was that broad TSRs are more susceptible to TGS by favouring production of promoter-associated transcripts. Despite enrichment for broad TSRs with a single TSC in the TGS group, none of the analyses revealed a significant enrichment in tiRNAs and ncRNAs compared to control groups. Nevertheless, the median TGS group tiRNA level is higher than controls throughout, despite the proportion of active genes in each control group differing substantially. This suggests that factors besides gene activation determine tiRNA level.

The ncRNA median for all groups was close to zero. This was because the majority of TSR do not overlap any ncRNAs, which is probably a consequence of the paucity of ncRNA data. A better analysis was, therefore, to compare the fraction of TSRs overlapping a ncRNA in each group. TGS group TSRs do overlap ncRNAs to a greater degree than any of the other datasets, although the *p* values are not significant (Table 5.3). The small number of TGS group TSRs and the scarcity of ncRNA datasets preclude conclusions as to the whether ncRNAs are enriched at TGS-susceptible promoters.

Table 5.3: Fraction of ncRNA overlap with TSRs in the various groups.

TSR Group	Number with ncRNA overlap	Number without ncRNA overlap	Fraction with overlap (%)	Comparison with TGS group (*p*-value)
TGS	5	9	35.71	
All other	7927	35232	18.37	0.094
THP-1	162	737	21.98	0.221
HK	110	241	31.13	0.729

The given *p*-value indicates whether the fraction of 2,000 bp windows centred on the representative TSS of TGS TSRs that overlap a ncRNA is significantly different from the fraction for the various controls. Chi-square test.

5.3.7 Classification of *htatsf1* and *ddx3x* promoter features

To determine which group the *htatsf1* and *ddx3x* promoters most closely resemble, a representative feature vector (RFV) was generated for each group and their Euclidean distance from *htatsf1* and *ddx3x* promoter features calculated. The *htatsf1* promoter feature vector most closely resembles the THP-1 group RFV and the *ddx3x* promoter feature vector most closely resembles the RFV for non-TGS TSRs in the genome (Table 5.4).

Table 5.4: The Euclidean distance between *htatsf1* and *ddx3x* promoter features and various TSR groups.

TSR Group	Euclidean distance	
	HTATSF1	DDX3X
TGS	280.65	92.03
All other	301.95	76.45
THP-1	25.55	344.82
HK	174.66	194.27

The Euclidean distance of promoter feature vectors for *htatsf1* (TSR ID TSR_hg18_v1_chrX_+_135407060) and *ddx3x* (TSR ID TSR ID TSR_hg18_v1_chrX_+_41078358) from the RFV of each group is shown.

A drawback of clustering by this method is that features that by their nature have a large scale will be inordinately important in determining distance to the cluster centre. In this case, this means TSR length will be far more important in assigning membership than, for example, CES, since TSR length can be hundreds of base pairs whilst the CES ranges from zero to one. The lengths of the TSR assigned to *htatsf1* and *ddx3x* are 389 bp and 22 bp, respectively. It is not surprising, therefore, that *htatsf1* and *ddx3x* promoters clustered with the THP-1 and all other TSRs groups, respectively.

5.4 Discussion

The approach used here was to identify promoter features associated with TGS susceptibility by looking for enrichment of features in promoters at which TGS had been reported compared to various control groups. This approach relies on the assumption that some genes are not susceptible to TGS and that there are a sufficient number of these TGS-resistant genes in each control group for relevant differences to become apparent. One could imagine that some genes would have evolved TGS-resistance if transcriptional silencing of the gene by an endogenous TGS mechanism were detrimental to the cell. HK genes are expressed in all tissues because they perform essential cellular functions and so this is the most probable group of TGS-resistant genes. However, one of the TGS group TSRs, associated with eukaryotic translation elongation factor 1 alpha (EEF1A1), is also in the HK gene group. This may be construed as good evidence that this is not a HK gene after all. In other words, that the Naïve Bayes method incorrectly predicted it to be a HK gene (De Ferrari and Aitken 2006). However, EEF1A1 has also been identified as expressed in all conditions tested in a publicly available database of microarray results (Eisenberg and Levanon 2003). Thus, it seems that at least some HK genes are susceptible to TGS.

The analysis was hampered by the extraction of genome-wide promoter data from cell lines different to those in which TGS was reported. For example, tiRNA levels were determined in THP-1 cells, an acute myeloid leukaemia cell line, histone modifications in CD4+ T cells and TSR data from a number of different cell lines and tissues with a bias towards THP-1. Nonetheless, the cosignificance of features retrieved from different sources that are known to associate, such as TSR length, CES and Sp1 binding sites, shows the retrieval and analysis methodology was effective and that different data sources did not interfere with detection of promoter feature associations.

Other observations supported the validity and effectiveness of the approach. Enrichment for active histone modifications in the THP-1 group demonstrates that the majority of highly expressed genes in THP-1 cells are also expressed in CD4+ T cells (Figure 5.5). Perhaps this relates to both the THP-1 group and histone modification data being derived from stimulated haematopoietic cells. This validates the use of the THP-1 group to control for features associated with active gene expression, in particular histone modifications. In addition, the histone acetylation modifications significantly different between the TGS group and the THP-1 group were mainly those associated with the TSS, rather than those associated with both promoters and actively transcribed regions. Compartmentalisation of acetylation patterns relates to the histone acetyltransferase (HAT) associated with RNA Pol II: PCAF is known to associate with elongation-competent RNA Pol II and p300 with initiation-competent RNA Pol II (Cho *et al.* 1998). The enrichment for TSS-associated histone acetylations in the THP-1 group (Figure 5.5) confirms that

assignment of THP-1 TSRs was effective and that the probability of the representative TSS being the most active TSS in stimulated CD4+ T cells is high. Likewise, the number of TSCs per TSR was greater for the HK gene group than the TGS group (Figure 5.6), which agrees with previous findings that even HK genes display alternate promoter usage with tissue-preference (Jacox *et al.* 2010), and suggests that assignment of HK gene TSRs was also effective.

One of the major difficulties of this study is discriminating between enrichment for promoter features that arise as a result of enhanced TGS susceptibility or a difference in the proportion of active genes. Thus, features less dependent on gene activity are probably most informative in discerning features associated with enhanced TGS susceptibility. Such features had the added advantage of being retrieved from sources that used several tissues and cell lines, with the exception of tiRNAs. Interrogation of these features suggests that TGS susceptible genes have broad TSRs comprising a single TSC. One could argue that a high number of TSCs within a TSR may be associated with increased gene susceptibility to TGS, as it implies increased regulation of the region and a propensity for nucleosome remodelling. However, one can speculate that, within a promoter architecture composed of a broad TSR and a single TSC, a stochastic process within the transcription machinery-accessible region determines the TSS usage (Balwierz *et al.* 2009). Hence, a broad TSR with one TSC implies a large accessible region, or NDR. This may favour production of promoter-associated transcripts and thus, increase TGS susceptibility. The TGS group were also characterised by high CpG class and CES. Together, it is unclear how informative this is as the majority of genes have high CpG class promoters, which is correlated with broad TSRs, and the majority of TSRs comprise only one TSC (Balwierz *et al.* 2009). Perhaps, therefore, the majority of human genes are indeed susceptible to TGS, or perhaps these features do not define a TGS susceptible gene at all. Without a larger TGS group it is difficult to make definite conclusions. Regardless, whilst the majority of promoters are high CpG class (Carninci *et al.* 2006; Balwierz *et al.* 2009), the majority of genes is not expressed in any given cell line at one time, indicating that broad TSR, high CpG class promoters are not all constitutively expressed and undergo silencing.

The hypothesis that TGS susceptible genes are characterised by broad TSRs with a single TSC through enhanced promoter-associated transcript production predicts that ncRNAs would be enriched in the TGS group. There was a trend toward increased tiRNAs, a ncRNA implicated in epigenetic silencing, at TGS group TSRs compared with all control groups, irrespective of whether the proportion of active TSRs was greater in the TGS or control groups. It has been reported that although promoters with tiRNAs are more highly expressed than promoters without tiRNAs, there is not a direct correlation (Taft *et al.* 2009). However, there is a correlation between promoters with tiRNAs and promoters with either a broad, general distribution of TSSs or a broad distribution with a dominant TSS peak (Taft *et al.* 2009); in other words, those with a broad TSR

and a single TSC. In contrast, tiRNAs are not correlated with promoters of a bi- or multi-modal distribution (Taft *et al.* 2009), or those with multiple TSCs. These findings are mirrored in this study, where increased tiRNAs in the TGS group are observed compared with the HK genes and THP-1 groups, which have a higher number of TSCs per TSR.

How tiRNAs could influence susceptibility to TGS is unclear. It has been suggested that tiRNAs are TFIIS cleavage products, which arise when RNA Pol II 'backtracks' at certain promoters (Taft *et al.* 2009). It is unlikely that they are simply signatures of transcription, however, as their size and non-random distribution is evolutionarily conserved and the relationship between tiRNA and gene expression level is complex. Small RNA inducers of TGS are effective when targeted to specific regions and can be extremely sensitive to single base-pair shifts (Janowski *et al.* 2005). A role for tiRNAs in chromatin regulation would be consistent with promoter-targeted siRNAs interfering with their production. Alternatively, tiRNAs could be the target with which guide strands hybridise or endogenous TGS-inducers themselves, fine-tuning gene expression. Consequently, their presence may indicate TGS susceptibility, although a larger TGS group is needed to test this hypothesis. Experimental evidence for a role of tiRNAs in TGS is lacking. However, ncRNAs displaying the characteristics of tiRNAs have also been found at splice sites, termed spliRNAs (Taft *et al.* 2010), and small RNA TGS-inducers have also been shown to regulate splicing through chromatin remodelling when targeted to splice sites. This may be coincidental, but the small size of these RNAs and their similarity to the RNAs used in TGS induction may have prohibited detection of a role in TGS to date.

The fraction of TSRs with overlapping ncRNAs was greatest in the TGS group, although there was not a statistically significant enrichment when compared to the various control groups (Table 5.3). ncRNAs have already been shown to be endogenous inducers of TGS. These include small RNAs, such as miRNAs (Gonzalez *et al.* 2008; Kim *et al.* 2008; Tan *et al.* 2009b), and long, antisense ncRNAs (Morris *et al.* 2008; Yu *et al.* 2008; Hawkins and Morris 2010; Su *et al.* 2010). Often, these ncRNAs are encoded across or proximal to the targeted gene (Kim *et al.* 2008; Morris *et al.* 2008; Yu *et al.* 2008; Tan *et al.* 2009b; Hawkins and Morris 2010). Unfortunately, only recently has the extent of non-coding transcription been appreciated and, consequently, the functional import of ncRNAs in regulating gene expression. Although functional genomics strategies are being used to improve ncRNA annotation (Forrest *et al.* 2009; Mattick 2009), current databases likely grossly underestimate their number. Therefore, no significant enrichment for ncRNAs in the TGS group may be due to the group's small size and the current paucity of ncRNA data. Of course, the number of TGS susceptible genes in the control groups may also influence whether significant results are obtained.

TGS was induced in HEK293T cells at *htatsf1* but not *ddx3x* (see Chapter 4). Promoter features of these genes were closest, determined by Euclidean distance, to the RFVs of the THP-1 group and all other TSRs, respectively (Table 5.4). Unfortunately, the TSR length inordinately influenced the clustering analysis. There are more sophisticated methods for clustering, such as the Naïve Bayes algorithm, but for the most part they are not useful when dealing with small sample sizes, which is the case here. Another drawback of this analysis is that the probability of correctly classifying two genes as TGS permissive or not by chance is high at 0.25. In addition, knowing whether a clustering method is effective depends on being certain of *htatsf1* and *ddx3x* TGS susceptibility. It is possible that an appropriate target site was not found in *ddx3x*, or that *ddx3x* is not susceptible to TGS in HEK293T cells in particular. Thus, a more rigorous prediction tool first requires a larger TGS group and this needs to be tested on more genes, which requires a great deal more experimental data.

Manual inspection of the *htatsf1* and *ddx3x* genes reveals TSRs of 389 and 22 bp, respectively. Both are high CpG class and comprise a single TSC. Based on these features alone, one would predict *htatsf1* to be more susceptible to TGS, if indeed broad TSRs favour production of promoter-associated transcripts. However, no experimental data on the relative abundance of promoter-associated transcripts for these genes was obtained.

To conclude, it would be helpful to be able to predict which genes are susceptible to TGS. To this end, there is a need to understand the context-dependent ability of small RNAs to induce TGS, which may depend, in part, on promoter architecture. The small number of genes at which TGS has been reported currently precludes definitive identification of promoter characteristics associated with TGS susceptibility and the classification of genes as TGS-permissive or not. However, this analysis is the first to demonstrate that there is enrichment for certain promoter features at genes where TGS has been reported. These include a high CpG class, broad TSR comprising a single TSC, which may facilitate the production of promoter-associated transcripts that are the targets of TGS-inducing small RNAs. In support of this hypothesis, there is some evidence of an increase in ncRNAs overlapping TGS TSRs, but a larger sample size is required to confirm this association. Nonetheless, this study provides initial guidelines for selection of TGS targets based on the architecture of promoters.

CHAPTER 6

General discussion

5.1 Summary of findings

There is still a need for novel anti-HIV therapies, particularly those that reduce viral escape, toxicity and administration frequency. This thesis examined whether putative HDFs HTATSF1, DDX3X and SMARCB1, are potential therapeutic targets that warrant further investigation and whether they are susceptible to TGS. Through exploitation of the RNAi pathway, it was confirmed for the first time that HTATSF1 is a HDF in T cell-derived SupT1 cells and, on shRNA-mediated suppression, inhibits HIV-1 replication without overt cytotoxicity (Chapters 2 and 3). The approach used here did not reveal a role for HTATSF1 in splicing of HIV-1 transcripts (Chapter 3), contributing toward the controversy surrounding the exact role(s) of HTATSF1 in the viral life cycle (Naito *et al.* 2007; Chen *et al.* 2009; Miller *et al.* 2009). This work also showed that expression of HTATSF1 could be reduced by induction of TGS with small, promoter-targeted siRNAs and that this was sufficient to inhibit viral replication, which ultimately may provide a durable silencing strategy (Chapter 4). The inability to confirm whether DDX3X and SMARCB1 were HDFs highlighted the limitations associated with RNAi-mediated suppression and the use of reporter cell systems for monitoring effects on HIV-1 replication (Chapters 2 and 3). TGS induction was not successful at the *ddx3x* promoter (Chapter 4 and A2), leading to the examination of promoter architectures genome-wide to assess whether certain features are associated with small RNA-induced TGS susceptibility. This novel analysis revealed that promoters at which TGS has been reported displayed broad TSRs comprising a single TSC, a high CES and CpG class and were enriched for overlapping ncRNAs (Chapter 5), which, hopefully, will assist researchers with TGS target selection in the future.

Inhibition of HIV-1 replication in SupT1 cells was not very pronounced on HTATSF1 suppression and may be accompanied by growth arrest (Chapter 3). Thus, the rigorous analysis employed here, whilst validating HTATSF1 as a HDF in T cell-derived lines, does not implicate HTATSF1 as a drug target. However, it should be noted that, although the SupT1 cell line more closely resembles the natural hosts of HIV-1 than the HeLa-derived TZM-bl reporter cells, this is still a model system. Levels of HDFs differ between cell types, which can alter the kinetics of HIV-1 replication (Li *et al.* 2010). To completely exclude the possibility of HTATSF1-targeting as a therapeutic strategy, an improved system is required, such as primary cells and/or a humanised mouse model of HIV-1 infection. The latter will hopefully reveal whether a selective advantage of

cells with HTATSF1 suppression is sufficient to overcome the observed slight decrease in cell growth (Chapter 3) in the context of HIV-1 infection *in vivo*. A biochemical approach to clarify the role(s) of HTATSF1 in HIV-1 infection would also be useful. Regardless of whether such studies implicate HTATSF1 as a drug target or not, the validation of HTATSF1 as a HDF described here may have implications for gene-association studies, as heterogeneity in HTATSF1 expression may affect patient susceptibility to HIV-1 infection and the rate of disease progression.

5.2 Viral escape

The high mutability of HIV-1 is problematic to the generation of therapies with sustained efficacy (see 1.2.5). HTATSF1 suppression inhibited p81A-4 replication in SupT1 cells for up to 14 days (Figure 3.8), although there is still the possibility that, over time, the virus will adapt to another cofactor. The lack of overlap between HDFs identified in genome-wide screens (Brass *et al.* 2008; Konig *et al.* 2008; Zhou *et al.* 2008; Yeung *et al.* 2009b) suggests a redundancy within the processes hijacked by the virus (see 1.3.4), which may facilitate viral escape from HDF suppression. Altered cofactor usage is likely to reduce replication efficiency, although it may also create a selective pressure on the tropism of the virus. For example, HTATSF1 suppression could drive a shift in the cell type in which the virus preferentially replicates to those where HTATSF1 suppression is incomplete, or to those where a redundant, albeit less efficient, cofactor is expressed at high levels. In turn, shifts in tropism could alter the pathogenesis associated with HIV-1 infection.

To date, viral escape in the context of HDF suppression has not been observed, although few studies have examined this phenomenon. Nevertheless, the apparent cure of a HIV-infected patient treated with a special bone marrow transplant containing haematopoietic stem cells with the Δ32 CCR5 allele (Hutter *et al.* 2009; Allers *et al.* 2010), despite the reported ability of HIV-1 to use other coreceptors (Alkhatib and Berger 2007), raises the intriguing possibility that cofactor adaptation in certain circumstances may be impossible. One report has specifically examined the ability of RNAi against cellular factors to mediate durable inhibition of HIV-1: the authors showed that cell lines with suppressed ALIX, TRBP or ATG16 expression were less susceptible to HIV-1 replication for up to two months (Eekels *et al.* 2011), albeit using a virus derived from a molecular clone, not a primary isolate containing multiple variants. Admittedly, the potential for viral escape may vary considerably with the HDF targeted, the cell type infected and the strain of virus involved, necessitating long-term studies in development of HDF-targeting therapies.

Besides increasing the genetic barrier to viral escape, another advantage to HDF targeting is that the accessibility of the target is consistent between cell types infected by HIV-1. In contrast, viral targets may be less accessible in macrophages than T cells because new virions are released into endocytic organelles where they accumulate prior to release as an exosome. Virions with these exosomes evade the immune system and, importantly, are sheltered from antiviral therapies and the RNAi machinery (Verani *et al.* 2005). Incoming viral genomic RNA is also protected from RNAi, as it remains partially encapsidated during reverse transcription (Westerhout *et al.* 2006; Arhel *et al.* 2007; Gao *et al.* 2008).

Of course the disadvantage of targeting a HDF such as HTATSF1 is that the target must be expressed in all cell types infected by HIV and the dosage needed to effectively inhibit HDF function without cytotoxic effects must be similar in all of them. Tissue-specific delivery or *ex vivo* treatments accompanied by engraftments become crucial where systemic inhibition is accompanied by toxic side effects. It is hoped that selection of cellular targets that are dispensable for cellular function or up-regulated specifically in cells infected by HIV will reduce these difficulties, but complete abolition of HDF function may not be desirable. Consequently, HDF-targeted therapies may prove most effective when used in combination with those targeting the virus directly (see 1.3.6).

It has been appreciated for some time that combinatorial strategies are required to minimise viral escape (Vandamme *et al.* 1998; Chang *et al.* 2005; Leonard and Schaffer 2005; ter Brake *et al.* 2006; Aviran *et al.* 2010). The addition of a compound capable of inhibiting HTATSF1 function, even if only able to mediate a partial inhibition of HIV-1 replication by itself, may have a synergistic effect on the long-term efficacy of combinatorial therapy by reducing viral escape. Indeed, several groups are developing a HIV-HDF combinatorial approach. For example, bifunctional siRNAs have been designed that simultaneously direct cleavage of HIV transcripts and CCR5 mRNA (Ehsani *et al.* 2010). In addition, a triple RNA-based combination therapy composed of a siRNA targeting Tat/Rev, a TAR decoy and a CCR5-targeting ribozyme has been shown to inhibit HIV-1 replication in primary haematopoietic cells (Li *et al.* 2005) and SCID-hu mouse-derived T cells (Anderson *et al.* 2007) to such an extent that this treatment has progressed to clinical trials (DiGiusto *et al.* 2010). The addition of small-molecule HDF inhibitors to existing HAART is also under investigation, such as those that abolish p75-IN and CCR5-gp120 interactions [reviewed in (Arhel and Kirchhoff 2010)]. Successful utilisation of HDF inhibitors may open the door to modulation of other cellular factors in HIV-1 treatment, such as over-expressing those functioning as viral restriction factors (eg APOBEC3G and TRIM5α [reviewed in (Strebel *et al.* 2009; Arhel and Kirchhoff 2010)] and inhibiting those involved in maintaining latency.

5.3 Application of HDF-targeted TGS

Small RNA-induced TGS presents a novel therapeutic approach for silencing expression of HDFs, including HTATSF1 (Chapter 4). This gene silencing strategy has the potential to mediate long-term target suppression (Hawkins *et al.* 2009) (see 1.4.1), which may reduce toxicities associated with exploitation of the PTGS arm of the RNAi pathway (Grimm *et al.* 2006; Grimm *et al.* 2010) where continuous guide strand expression is required. The durability of TGS in the absence of promoter-targeted guide strand is likely to vary substantially between genes, however. Although controversial, a number of studies suggest transcription is pervasive (Kapranov *et al.* 2002; Bertone *et al.* 2004; Kampa *et al.* 2004; Schadt *et al.* 2004; Carninci *et al.* 2005; Cheng *et al.* 2005; Carninci *et al.* 2006; Birney *et al.* 2007; Kapranov *et al.* 2007a; Kapranov *et al.* 2007b; van Bakel *et al.* 2010), which may lead to reactivation of the targeted gene or prevent complete silencing. In addition, given the target gene is actively transcribed prior to TGS, other pathways in the cell may favour reactivation of the gene. Duration of TGS may be influenced by long-range chromosomal interactions: coregulated genes are preferentially associated within transcription factories (Schoenfelder *et al.* 2010), but little is known about the regulation of transcription factory association and so it is unclear whether long-term TGS requires a loss of target association with these specialised transcriptional hotspots. Hopefully, future studies will examine how 3D chromosomal interactions are affected by small RNA-induced TGS and the role of these interactions in TGS duration.

The widespread application of TGS is compounded by sparse knowledge of the mechanism(s) involved (see 1.4) and what governs efficacy in selection of both a target gene and a target site within a promoter. Several promoter features were enriched in TGS susceptible genes (Chapter 5), the application of which may facilitate selection of additional HDF targets besides HTATSF1. The significance of these findings is unclear, however, as TGS reports are few and a number describe TGS at sites other than the promoter, including regulatory domains (Gonzalez *et al.* 2008), downstream of 3' UTRs (Yue *et al.* 2010) and alternatively spliced exons (Allo *et al.* 2009; Luco *et al.* 2011). These regions do not display all the promoter features identified as enriched in TGS susceptible genes, suggesting that more experimental data is required before conclusions can be drawn on what discriminates a good TGS target.

Although guidelines for selection of TGS target genes and sites within a promoter have not been defined, it may be possible to increase the potency of TGS by creating a TGS-permissive state within the cell. Nascent RNA Pol II transcripts are required for small RNA-induced transcriptional regulation (Han *et al.* 2007). Therefore, enhanced TGS susceptibility may be

achieved by inhibiting the exosome, which has been shown to increase the life-span of promoter-associated RNAs (Preker *et al.* 2008).

Selective periods of the cell cycle may be preferentially TGS-permissive. In *S. pombe*, brief centromeric repeat transcription during S phase is required to maintain heterochromatin at the region through successive cell divisions by producing the RNAi effectors (Chen *et al.* 2008). Similarly, convergent genes in *S. pombe* employ a cell cycle-regulated transcription termination mechanism that involves dsRNA production and heterochromatin formation (Gullerova and Proudfoot 2008). In human cells, TGS-inducing rRNAs are increased in a cell cycle-regulated manner (Santoro *et al.* 2010). Further, intrachromosomal interactions between enhancer regions and the TSS can be periodic, varying in a ligand- and cell cycle-dependent fashion (Matilainen *et al.* 2010). Such periodicity may affect target gene susceptibility if using enhancer-targeted guide strands. These studies suggest that RNA-induced heterochromatin formation may be favoured during specific periods of the cell cycle, through up-regulation of pathway components and/or enhanced association of the targeted region with the proximal promoter.

Decreasing the elongation rate of RNA Pol II may enhance siRNA-mediated heterochromatin formation. Recently, siRNAs have been shown to inhibit RNA Pol II during elongation in *C. elegans* (Guang *et al*. 2010). In *S. pombe*, 6-azauracil inhibits RNA Pol II elongation and enhances RNA-induced gene silencing (S. I. Grewal, personal communication). Similarly, in human cells, decreased RNA Pol II processivity has been associated with heterochromatin marks and recruitment of HP1 isoforms (Allo *et al.* 2009). Thus, introducing a drug to inhibit RNA Pol II processivity at the same time as introducing/expressing the TGS-inducing small RNA may enhance TGS establishment.

A major concern of a TGS-based gene therapy is the potential for off-target effects. Small RNA TGS effectors resemble miRNAs and so have the potential to induce PTGS of mRNAs with cognate seed-match hexamers (Jackson *et al.* 2003; Jackson and Linsley 2004; Jackson *et al.* 2006). They could also give rise to unwanted changes in alternative splicing (Allo *et al.* 2009; Luco *et al.* 2011) through TGS-mediated affects. Even if care is taken to avoid such target redundancy, the demonstration that similar small RNAs can induce both TGS and TGA (Janowski *et al.* 2007; Turunen *et al.* 2009) (see 1.4.4) in a currently unpredictable manner raises the issue of how to avoid unwanted TGA on systemic administration of TGS effectors. Currently, there is no data on the consequences of HTATSF1 over-expression. No doubt the progress of TGS to the clinic will be assisted by developments in tissue-specific delivery of RNAi mimics [reviewed in (Tiemann and Rossi 2009; Lares *et al.* 2010)].

5.4 Delivery of HDF-targeted TGS-based gene therapy

Although non-encapsulated siRNAs have been delivered directly to cells by gymnosis (Stein *et al.* 2010), this is not generally feasible for most therapeutic applications, including antiviral therapy. siRNAs are inherently unstable, vulnerable to degradation and may trigger an immune response, thus a method of delivery is required to ensure siRNAs remain functional long enough to reach their target cells. Chemical modifications, such as locked nucleic acids and 2'-*O*-methyl-modified uridine or guanine nucleosides, have been used to increase resistance to nuclease activity and target specificity, and decrease inflammation [reviewed in (Shukla *et al.* 2010)]. Importantly, introduction of such modifications did not alter the ability of promoter-targeted agRNAs to induce TGS (Watts *et al.* 2010), although care must be taken as incorporation of too many modifications can cause toxicity and/or reduce efficacy.

Lipid-based delivery vehicles, including liposomes, lipoplexes and nanostructured gyroid cubic lipid matrices (Leal *et al.* 2010), protect from nuclease degradation and mediate high cell transfection efficiency but do not permit tissue-specific delivery of siRNAs [reviewed in (Lares *et al.* 2010)]. Liposomes may be replaces by peptides, such as histidine-lysine polymers, which may be conjugated to ligands for cell type-specific delivery, but the transfection efficiency of such vehicles is generally low (Leng *et al.* 2007). The first Phase I clinical trial involving systemic administration of siRNAs successfully achieved solid tumour-targeting using nanoparticle vehicles composed of a linear, cyclodextrin-based polymer, a human transferring protein ligand to target the transferring receptor and a hydrophilic polymer (polyethylene glycol) to promote nanoparticle stability, as well as the anti-ribonucleotide reductase siRNA.

Other ligand-based targeting methods have been developed specifically for the delivery of siRNAs to HIV-infected cells that avoid the requirement for polymer-based nanoparticles. Aptamers are structured nucleic acids with high affinity to a chosen target. They have the distinct advantage of being relatively easy to produce *in vitro* as siRNA-aptamer chimaeras may be synthesised as a single polymer. Two RNA aptamers that bind the HIV-1 gp120 have been developed: one is an aptamer-siRNA chimaera, the other contains a GC-bridge between the two to facilitate interchange of siRNAs (Zhou *et al.* 2009). Whilst the aptamer alone inhibits HIV replication, the chimaeras provide more potent inhibition through delivery of siRNAs targeting HIV transcripts. An alternative method for delivery of siRNAs to HIV-infected cells is the conjugation of siRNAs to antibody fragments. One approach used a fusion protein, protamine, which binds siRNA, conjugated to the Fab fragment of an HIV-1 envelope antibody (Song *et al.* 2005). However, it is unknown whether delivery strategies employing aptamers and antibodies targeting highly variable HIV envelope proteins will select for resistant virus. Another approach

used a modified scFv antibody, which targets T cell surface protein CD7, to allow conjugation to a nona-D-arginine peptide that engaged with siRNA (Kumar *et al.* 2008). Clearly, great improvements in the tissue-specific delivery of siRNAs are being made, although it has not been examined whether these methods mediate efficient nuclear delivery of siRNAs, which is required for TGS.

Other methods are being developed for the delivery of RNAi expression cassettes, including hijacking the Sleeping Beauty DNA transposon (Fletcher 2010) and viral vectors, in particular lentiviruses. Lentiviruses may infect a number of different cell types, including haematopoietic stem cells and non-dividing resting memory T cells (Yam *et al.* 2002). Improvements have been made to prevent low titres through siRNA production targeting components of the packaging system (Poluri and Sutton 2008; Liu *et al.* 2010), although this is less likely to be an issue with HDF-targeting effectors. In addition, safer vectors are being developed that minimise the risk of recombination with infectious HIV to produce new hybrid viruses (Miyoshi *et al.* 1998; VandenDriessche *et al.* 2002). Insertional mutagenesis is a concern and, although it has not yet been observed with lentiviruses, it may be prevented by targeting integration away from functional genes using an artificial tethering factor (Gijsbers *et al.* 2010; Silvers *et al.* 2010). Strategies to insulate lentiviral vector-expressed transgenes from the influence of neighbouring chromatin have also been described (Ramezani and Hawley 2010).

Tissue-specific expression of guide strand precursors may be obtained by the use of appropriate RNA Pol II promoters. Although products of pri-miRNA mimics have not been shown to mediate TGS, products of shRNAs have (Castanotto *et al.* 2005; Kim *et al.* 2006; Murayama *et al.* 2006; Hawkins *et al.* 2009; Turunen *et al.* 2009; Yamagishi *et al.* 2009), suggesting that incorporation of TGS-inducing guide strands into mimics that require Dicer processing does not ablate efficacy and may be extended to those which also require Drosha processing. This may be facilitated by use of pri-miRNA scaffolds derived from endogenous miRNAs shown to mediate TGS (Gonzalez *et al.* 2008; Kim *et al.* 2008; Tan *et al.* 2009b). Expression specifically in HIV-infected cells may be permitted through the use of mobilisation competent vectors and inducible systems whose activation is stimulated by the presence of Tat (Unwalla and Rossi 2008; Turner *et al.* 2009), although these would be refractory in the context of latent infection.

In summary, many promising methods for delivery of RNAi mimics have been developed that may be applied to the delivery of TGS effectors to the nucleus of their intended targets. The most effective delivery strategy may depend on the nature of the other components of the combinatorial therapy. Providing efficient delivery to HIV-infected cells is achieved, it is unknown the extent to which the efficacy of TGS effectors will be affected by HIV-mediated modulation of the RNAi pathway.

5.5 Modulation of RNAi by HIV

The role of RNAi as an antiviral defence mechanism in mammalian cells remains controversial (Saumet and Lecellier 2006). Several pathogenic human viruses were shown to encode proteins that function as suppressors of RNAi [reviewed in (de Vries and Berkhout 2008)], supporting that RNAi serves as an innate defence response in mammals. In addition, components of the RNAi machinery, including TRBP and PACT, have also been described as inducers of an interferon response [reviewed in (Lares *et al.* 2010)].

An interesting interplay exists between HIV and the RNAi machinery [reviewed in (Corbeau 2008)]. RNAi is capable of restricting HIV-1 replication by targeting viral transcripts with either miRNAs or siRNAs (Klase *et al.* 2007; Liu *et al.* 2009; Nathans *et al.* 2009; Wang *et al.* 2009b; Yeung *et al.* 2009a). Ablation of Dicer (Triboulet *et al.* 2007), or disruption of P body structures (Nathans *et al.* 2009), increases HIV-1 replication in T cells, suggesting an anti-HIV role for RNAi. It is perhaps not surprising, therefore, that HIV counteracts RNAi in T cells: Tat interacts directly with Dicer reducing its ability to produce miRNAs and the viral TAR element sequesters TRBP (Bennasser *et al.* 2005; Bennasser and Jeang 2006). This suppression of RNAi may serve to reduce the production of HIV-targeted host miRNAs (Hariharan *et al.* 2005; Huang *et al.* 2007), as well as reduce the production of miRNAs that regulate HDFs, such as miR-17-5p and miR-20 that silence PCAF (Triboulet *et al.* 2007). Other host miRNAs are up-regulated on HIV infection (Yeung *et al.* 2005; Triboulet *et al.* 2007) but the molecular basis of the up-regulation of some and down-regulation of others is not understood.

A global suppression of RNAi may not be beneficial to the virus as HIV appears to encode a number of its own miRNAs. The earliest direct evidence for a viral miRNA was found by Northern blot analysis (Klase *et al.* 2007; Ouellet *et al.* 2008). More powerful techniques, such as primer-directed sequencing (Klase *et al.* 2009) and pyrosequencing (Yeung *et al.* 2009a), have verified that viral miRNAs are produced. Of note is a viral miRNA derived from the TAR element, produced during productive and latent infections, which serves two known functions in the maintenance of latency: to induce TGS at the LTR (Klase *et al.* 2007; Suzuki *et al.* 2008) and to down-regulate host genes by PTGS with roles in DNA repair, cytokine signalling and apoptosis, thus promoting the survival of latently-infected cells (Klase *et al.* 2009). Another viral miRNA, hiv1-mir-H1, down-regulates an immune response and inhibits expression of Vpr-targeting cellular miR-149 (Kaul *et al.* 2009).

The reciprocal interactions between HIV and RNAi are further complicated by differences between the cell types HIV infects. It has been suggested that miR-198 restricts HIV-1 replication

in monocytes through targeting of the HDF cyclin T1 (Sung and Rice 2009). However, Dicer is absent from monocytes (Klase *et al.* 2007) with low-level miRNA generation in these cells dependent on PIWIL4 (Coley *et al.* 2010), a member of the Ago family of proteins whose function was previously believed restricted to the formation of heterochromatin guided by piwi-interacting RNAs (piRNAs) at transposable elements [reviewed in (Kim *et al.* 2009)]. Dicer expression was detected on differentiation of monocytes into monocyte-derived macrophages, although this was inhibited in the presence of HIV in a manner dependent on Vpr (Coley *et al.* 2010). Intriguingly, this suggests that RNAi is suppressed in both monocytes and T cells by two independent mechanisms.

Thus, the global impact of the reciprocal interactions between HIV and RNAi, and how this impacts the efficacy of TGS-based gene therapy, on *in vivo* infection remains to be determined. Other strategies for TGS induction may need to be considered, including the expression of long, antisense ncRNAs. Endogenous asRNAs have been shown to regulate gene expression and their over-expression may lead to heterochromatin formation and TGS (Morris *et al.* 2008; Yu *et al.* 2008; Hawkins and Morris 2010).

5.6 Concluding remarks

Despite being the most studied infectious agent of this century, HIV remains a major global health concern. The high mutability of the virus has countered attempts to develop a prophylactic vaccine and a cure. The reliance of viruses, including HIV, on their hosts for replication raises the intriguing possibility of targeting a host factor as opposed to the virus itself. Whilst this may engender toxicity, it is hoped that this will reduce the ability of HIV to develop treatment resistance and will also provide a therapy that is broadly inhibitory of HIV subtypes.

The observation that homozygosity for the naturally occurring CCR5 Δ32 allele confers resistance to R5-tropic viruses without any observed detrimental effects to the host, suggests that it may be possible to treat HIV by targeting host proteins required for viral replication without toxic side effects. A number of genome-wide screens has been performed to identify host proteins involved in HIV-1 replication, although all were subject to a number of limitations. Given that levels of HDFs influence HIV replication kinetics and gene expression varies between tissues, it is important that putative HDFs identified by these screens are validated with additional experiments, preferably in cells that are natural substrates of HIV infection.

This thesis presents an effective protocol for the validation of HDFs and elucidation of the therapeutic potential of their targeting. Application of this protocol has contributed to the list of

validated HDFs by confirming that HTATSF1 functions as an HDF in T cell-derived SupT1 cells. Reduced HTATSF1 expression moderately decreased the efficiency of HIV-1 replication without overt cytotoxicity, although this may be accompanied by decreased proliferative capacity of cells. Further studies are required to clarify the significance of HTATSF1 suppression on viral replication *in vivo*, and whether HTATSF1 inhibitors would improve existing combinatorial strategies by increasing the genetic barrier to viral escape. However, the work presented in this thesis suggests that HTATSF1 may not constitute a therapeutic target, as its suppression did not have a dramatic effect on viral replication. Whether a drug target or not, validation of HTATSF1 as a HDF is valuable since its expression level may influence disease progression. Thus, the degree of HTATSF1 heterogeneity within the population should be examined and *htatsf1* should be included in gene-association studies for the detection of HIV risk factors. Most importantly, the methodology employed here was successful at discriminating a validated HDF from a promising drug target. This thesis demonstrates that a thorough analysis of each putative HDF is necessary to confirm its therapeutic potential and, as such, a protocol similar to that executed in this thesis should be applied when validating results from genome-wide screens.

Current HIV treatment regimens suffer from limitations associated with their chronic, daily administration, including toxic side effects and patient compliance. Alternative therapies that reduce intervention frequency and are practical to resource-poor settings, would, therefore, be of benefit. Small RNA-induced TGS is an arm of the RNAi pathway recently identified in mammalian cells that has the potential to provide durable silencing, providing heterochromatin is achieved at the targeted promoter. This thesis demonstrated that it is possible to suppress HTATSF1 expression through induction of TGS at the *htatsf1* promoter, expanding the options for targeting of this HDF. This thesis also provides proof of principle that TGS at a HDF promoter may inhibit HIV-1 replication. Before implementation as a therapeutic strategy, future work must examine the potential for TGS at HDF promoters in T cells and whether this is sufficiently potent to inhibit viral replication over time.

The huge progress made in delivery of RNAi effectors ensures that this should not prove a barrier to the implementation of TGS-based gene therapies. Of greater concern is the potential of HIV to modulate the RNAi pathway and how this may impact TGS efficacy *in vivo*. Furthermore, difficulties in TGS target identification currently precludes widespread exploitation of this gene-silencing phenomenon, although the novel analysis employed here for the identification of promoter features enriched in TGS susceptible genes will provide the groundwork for improvements in this regard. In conclusion, this thesis provides an experimental strategy for elucidation of the therapeutic potential of targeting HDFs and, ultimately, paves the way for a HDF-targeted TGS-based treatment that provides durable, sustained inhibition of the virus.

REFERENCES

Agrawal, L., C. R. Maxwell, P. J. Peters, P. R. Clapham, S. M. Liu, C. R. Mackay and D. S. Strayer (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." J Gen Virol **90**(Pt 3): 710-22.

Aich, P., H. L. Wilson, R. S. Kaushik, A. A. Potter, L. A. Babiuk and P. Griebel (2007). "Comparative analysis of innate immune responses following infection of newborn calves with bovine rotavirus and bovine coronavirus." J Gen Virol **88**(Pt 10): 2749-61.

Alkhatib, G. and E. A. Berger (2007). "HIV coreceptors: from discovery and designation to new paradigms and promise." Eur J Med Res **12**(9): 375-84.

Allers, K., G. Hutter, J. Hofmann, C. Loddenkemper, K. Rieger, E. Thiel and T. Schneider (2010). "Evidence for the cure of HIV infection by CCR5{Delta}32/{Delta}32 stem cell transplantation." Blood.

Allo, M., V. Buggiano, J. P. Fededa, E. Petrillo, I. Schor, M. de la Mata, E. Agirre, M. Plass, E. Eyras, S. A. Elela, R. Klinck, B. Chabot and A. R. Kornblihtt (2009). "Control of alternative splicing through siRNA-mediated transcriptional gene silencing." Nat Struct Mol Biol **16**(7): 717-24.

Amarzguioui, M. and H. Prydz (2004). "An algorithm for selection of functional siRNA sequences." Biochem Biophys Res Commun **316**(4): 1050-8.

An, D. S., R. E. Donahue, M. Kamata, B. Poon, M. Metzger, S. H. Mao, A. Bonifacino, A. E. Krouse, J. L. Darlix, D. Baltimore, F. X. Qin and I. S. Chen (2007). "Stable reduction of CCR5 by RNAi through hematopoietic stem cell transplant in non-human primates." Proc Natl Acad Sci U S A **104**(32): 13110-5.

An, D. S., F. X. Qin, V. C. Auyeung, S. H. Mao, S. K. Kung, D. Baltimore and I. S. Chen (2006). "Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors." Mol Ther **14**(4): 494-504.

An, P. and C. A. Winkler (2010). "Host genes associated with HIV/AIDS: advances in gene discovery." Trends Genet **26**(3): 119-31.

Andersen, J. L., E. Le Rouzic and V. Planelles (2008). "HIV-1 Vpr: mechanisms of G2 arrest and apoptosis." Exp Mol Pathol **85**(1): 2-10.

Anderson, J. and R. Akkina (2005a). "CXCR4 and CCR5 shRNA transgenic CD34+ cell derived macrophages are functionally normal and resist HIV-1 infection." Retrovirology **2**: 53.

Anderson, J. and R. Akkina (2005b). "HIV-1 resistance conferred by siRNA cosuppression of CXCR4 and CCR5 coreceptors by a bispecific lentiviral vector." AIDS Res Ther **2**(1): 1.

Anderson, J., A. Banerjea and R. Akkina (2003a). "Bispecific short hairpin siRNA constructs targeted to CD4, CXCR4, and CCR5 confer HIV-1 resistance." Oligonucleotides **13**(5): 303-12.

Anderson, J., A. Banerjea, V. Planelles and R. Akkina (2003b). "Potent suppression of HIV type 1 infection by a short hairpin anti-CXCR4 siRNA." AIDS Res Hum Retroviruses **19**(8): 699-706.

Anderson, J., M. J. Li, B. Palmer, L. Remling, S. Li, P. Yam, J. K. Yee, J. Rossi, J. Zaia and R. Akkina (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." Mol Ther **15**(6): 1182-8.

Angus, A. G., D. Dalrymple, S. Boulant, D. R. McGivern, R. F. Clayton, M. J. Scott, R. Adair, S. Graham, A. M. Owsianka, P. Targett-Adams, K. Li, T. Wakita, J. McLauchlan, S. M. Lemon and A. H. Patel (2010). "Requirement of cellular DDX3 for hepatitis C virus replication is unrelated to its interaction with the viral core protein." J Gen Virol **91**(Pt 1): 122-32.

Anish, R., M. B. Hossain, R. H. Jacobson and S. Takada (2009). "Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements." PLoS One **4**(4): e5103.

Arhel, N. and F. Kirchhoff (2010). "Host proteins involved in HIV infection: new therapeutic targets." Biochim Biophys Acta **1802**(3): 313-21.

Arhel, N. J. and F. Kirchhoff (2009). "Implications of Nef: host cell interactions in viral persistence and progression to AIDS." Curr Top Microbiol Immunol **339**: 147-75.

Arhel, N. J., S. Souquere-Besse, S. Munier, P. Souque, S. Guadagnini, S. Rutherford, M. C. Prevost, T. D. Allen and P. Charneau (2007). "HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore." EMBO J **26**(12): 3025-37.

Ariumi, Y., F. Serhan, P. Turelli, A. Telenti and D. Trono (2006). "The integrase interactor 1 (INI1) proteins facilitate Tat-mediated human immunodeficiency virus type 1 transcription." Retrovirology **3**: 47.

Aviran, S., P. S. Shah, D. V. Schaffer and A. P. Arkin (2010). "Computational models of HIV-1 resistance to gene therapy elucidate therapy design principles." PLoS Comput Biol **6**(8).

Balwierz, P. J., P. Carninci, C. O. Daub, J. Kawai, Y. Hayashizaki, W. Van Belle, C. Beisel and E. van Nimwegen (2009). "Methods for analyzing deep sequencing expression data: constructing the human and mouse promoterome with deepCAGE data." Genome Biol **10**(7): R79.

Barichievy, S., S. Saayman, P. Arbuthnot and M. S. Weinberg (2009). "RNA interference-based gene expression strategies aimed at sustained therapeutic inhibition of HIV." Curr Top Med Chem **9**(12): 1065-78.

Barichievy, S., S. Saayman, K. J. von Eije, K. V. Morris, P. Arbuthnot and M. S. Weinberg (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**(4): 419-31.

Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-37.

Bartel, D. P. (2009). "MicroRNAs: target recognition and regulatory functions." Cell **136**(2): 215-33.

Bartholomeeusen, K., F. Christ, J. Hendrix, J. C. Rain, S. Emiliani, R. Benarous, Z. Debyser, R. Gijsbers and J. De Rijck (2009). "Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ." J Biol Chem **284**(17): 11467-77.

Bartholomeeusen, K., J. De Rijck, K. Busschots, L. Desender, R. Gijsbers, S. Emiliani, R. Benarous, Z. Debyser and F. Christ (2007). "Differential interaction of HIV-1 integrase and JPO2 with the C terminus of LEDGF/p75." J Mol Biol **372**(2): 407-21.

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 Res **34**(1): 322-33.

Bennasser, Y. and K. T. Jeang (2006). "HIV-1 Tat interaction with Dicer: requirement for RNA." Retrovirology **3**: 95.

Bennasser, Y., S. Y. Le, M. Benkirane and K. T. Jeang (2005). "Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing." Immunity **22**(5): 607-19.

Berezikov, E., W. J. Chung, J. Willis, E. Cuppen and E. C. Lai (2007). "Mammalian mirtron genes." Mol Cell **28**(2): 328-36.

Berkhout, B. (2004). "RNA interference as an antiviral approach: targeting HIV-1." Curr Opin Mol Ther **6**(2): 141-5.

Berkhout, B. (2009). "Toward a durable anti-HIV gene therapy based on RNA interference." Ann N Y Acad Sci **1175**: 3-14.

Berkhout, B., R. H. Silverman and K. T. Jeang (1989). "Tat trans-activates the human immunodeficiency virus through a nascent RNA target." Cell **59**(2): 273-82.

Berkowitz, R., J. Fisher and S. P. Goff (1996). "RNA packaging." Curr Top Microbiol Immunol **214**: 177-218.

Bertone, P., V. Stolc, T. E. Royce, J. S. Rozowsky, A. E. Urban, X. Zhu, J. L. Rinn, W. Tongprasit, M. Samanta, S. Weissman, M. Gerstein and M. Snyder (2004). "Global identification of human transcribed sequences with genome tiling arrays." Science **306**(5705): 2242-6.

Bertrand, E., D. Castanotto, C. Zhou, C. Carbonnelle, N. S. Lee, P. Good, S. Chatterjee, T. Grange, R. Pictet, D. Kohn, D. Engelke and J. J. Rossi (1997). "The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization." RNA **3**(1): 75-88.

Biegel, J. A., J. Y. Zhou, L. B. Rorke, C. Stenstrom, L. M. Wainwright and B. Fogelgren (1999). "Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors." Cancer Res **59**(1): 74-9.

Birney, E., J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M. Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermuller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W. K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C. L. Wei, Y. Ruan, K. Struhl, M. Gerstein, S. E. Antonarakis, Y. Fu, E. D. Green, U. Karaoz, A. Siepel, J. Taylor, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, G. M. Cooper, G. Asimenos, C. N. Dewey, M. Hou, S. Nikolaev, J. I. Montoya-Burgos, A. Loytynoja, S. Whelan, F. Pardi, T. Massingham, H. Huang, N. R. Zhang, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, A. Sidow, N. D. Trinklein, Z. D. Zhang, L. Barrera, R. Stuart, D. C. King, A. Ameur, S. Enroth, M. C. Bieda, J. Kim, A. A. Bhinge, N. Jiang, J. Liu, F. Yao, V. B. Vega, C. W. Lee, P. Ng, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, J. C. Fowler, P. Couttet, A. W. Bruce, O. M. Dovey, P. D. Ellis, C. F. Langford, D. A. Nix, G. Euskirchen, S. Hartman, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. H. Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, M. Xu, J. N. Haidar, Y. Yu, V. R. Iyer, R. D. Green, C. Wadelius, P. J. Farnham, B. Ren, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, L. Armengol, C. P. Bird, P. I. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, A. Woodroffe, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrimsdottir, J. Huppert, M. C. Zody, G. R. Abecasis, X. Estivill, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V. Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R. W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E. R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu and P. J. de Jong (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." Nature **447**(7146): 799-816.

Bock, C., J. Walter, M. Paulsen and T. Lengauer (2007). "CpG island mapping by epigenome prediction." PLoS Comput Biol **3**(6): e110.

Boden, D., O. Pusch, F. Lee, L. Tucker and B. Ramratnam (2003). "Human immunodeficiency virus type 1 escape from RNA interference." J Virol **77**(21): 11531-5.

Boden, D., O. Pusch, R. Silbermann, F. Lee, L. Tucker and B. Ramratnam (2004). "Enhanced gene silencing of HIV-1 specific siRNA using microRNA designed hairpins." Nucleic Acids Res **32**(3): 1154-8.

Boese, A., P. Sommer, D. Holzer, R. Maier and U. Nehrbass (2009). "Integrase interactor 1 (Ini1/hSNF5) is a repressor of basal human immunodeficiency virus type 1 promoter activity." J Gen Virol **90**(Pt 10): 2503-12.

Bogerd, H. P., A. Echarri, T. M. Ross and B. R. Cullen (1998). "Inhibition of human immunodeficiency virus Rev and human T-cell leukemia virus Rex function, but not Mason-Pfizer monkey virus constitutive transport element activity, by a mutant human nucleoporin targeted to Crm1." J Virol **72**(11): 8627-35.

Bogerd, H. P., H. W. Karnowski, X. Cai, J. Shin, M. Pohlers and B. R. Cullen (2010). "A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs." Mol Cell **37**(1): 135-42.

Borchert, G. M., W. Lanier and B. L. Davidson (2006). "RNA polymerase III transcribes human microRNAs." Nat Struct Mol Biol **13**(12): 1097-101.

Borner, K., J. Hermle, C. Sommer, N. P. Brown, B. Knapp, B. Glass, J. Kunkel, G. Torralba, J. Reymann, N. Beil, J. Beneke, R. Pepperkok, R. Schneider, T. Ludwig, M. Hausmann, F. Hamprecht, H. Erfle, L. Kaderali, H. G. Krausslich and M. J. Lehmann (2010). "From experimental setup to bioinformatics: an RNAi screening platform to identify host factors involved in HIV-1 replication." Biotechnol J **5**(1): 39-49.

Boudreau, R. L., A. M. Monteys and B. L. Davidson (2008). "Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs." RNA **14**(9): 1834-44.

Bourgo, R. J., H. Siddiqui, S. Fox, D. Solomon, C. G. Sansam, M. Yaniv, C. Muchardt, D. Metzger, P. Chambon, C. W. Roberts and E. S. Knudsen (2009). "SWI/SNF deficiency results in aberrant chromatin organization, mitotic failure, and diminished proliferative capacity." Mol Biol Cell **20**(14): 3192-9.

Brass, A. L., D. M. Dykxhoorn, Y. Benita, N. Yan, A. Engelman, R. J. Xavier, J. Lieberman and S. J. Elledge (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." Science **319**(5865): 921-6.

Briggs, J. A., J. D. Riches, B. Glass, V. Bartonova, G. Zanetti and H. G. Krausslich (2009). "Structure and assembly of immature HIV." Proc Natl Acad Sci U S A **106**(27): 11090-5.

Briggs, J. A., T. Wilk, R. Welker, H. G. Krausslich and S. D. Fuller (2003). "Structural organization of authentic, mature HIV-1 virions and cores." EMBO J **22**(7): 1707-15.

Broder, S. (2010). "The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic." Antiviral Res **85**(1): 1-18.

Brooks, J. T., J. E. Kaplan, K. K. Holmes, C. Benson, A. Pau and H. Masur (2009). "HIV-associated opportunistic infections--going, going, but not gone: the continued need for prevention and treatment guidelines." Clin Infect Dis **48**(5): 609-11.

Brown, P. O., B. Bowerman, H. E. Varmus and J. M. Bishop (1987). "Correct integration of retroviral DNA in vitro." Cell **49**(3): 347-56.

Buhler, M. and D. Moazed (2007). "Transcription and RNAi in heterochromatic gene silencing." Nat Struct Mol Biol **14**(11): 1041-1048.

Buhler, M., A. Verdel and D. Moazed (2006). "Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing." Cell **125**(5): 873-86.

Bukrinsky, M. (2006). "SNFing HIV transcription." Retrovirology **3**: 49.

Bushman, F. D., N. Malani, J. Fernandes, I. D'Orso, G. Cagney, T. L. Diamond, H. Zhou, D. J. Hazuda, A. S. Espeseth, R. Konig, S. Bandyopadhyay, T. Ideker, S. P. Goff, N. J. Krogan, A. D. Frankel, J. A. Young and S. K. Chanda (2009). "Host cell factors in HIV replication: meta-analysis of genome-wide studies." PLoS Pathog **5**(5): e1000437.

Butler, J. E. and J. T. Kadonaga (2002). "The RNA polymerase II core promoter: a key component in the regulation of gene expression." Genes Dev **16**(20): 2583-92.

Cai, X., C. H. Hagedorn and B. R. Cullen (2004). "Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs." RNA **10**(12): 1957-66.

Cairns, B. R. (2009). "The logic of chromatin architecture and remodelling at promoters." Nature **461**(7261): 193-8.

Campbell, E. M. and T. J. Hope (2008). "Live cell imaging of the HIV-1 life cycle." Trends Microbiol **16**(12): 580-7.

Cano, J. and G. V. Kalpana (2010). "Inhibition of early stages of HIV-1 assembly by INI1/hSNF5 transdominant negative mutant S6." J Virol.

Carlson, L. A., J. A. Briggs, B. Glass, J. D. Riches, M. N. Simon, M. C. Johnson, B. Muller, K. Grunewald and H. G. Krausslich (2008). "Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis." Cell Host Microbe **4**(6): 592-9.

Carmona, S., A. Ely, C. Crowther, N. Moolla, F. H. Salazar, P. L. Marion, N. Ferry, M. S. Weinberg and P. Arbuthnot (2006). "Effective inhibition of HBV replication in vivo by anti-HBx short hairpin RNAs." Mol Ther **13**(2): 411-21.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M. Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A. Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai and Y. Hayashizaki (2005). "The transcriptional landscape of the mammalian genome." Science **309**(5740): 1559-63.

Carninci, P., A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. Semple, M. S. Taylor, P. G. Engstrom, M. C. Frith, A. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume and Y. Hayashizaki (2006). "Genome-wide analysis of mammalian promoter architecture and evolution." Nat Genet **38**(6): 626-35.

Carr, J. K., M. Avila, M. Gomez Carrillo, H. Salomon, J. Hierholzer, V. Watanaveeradej, M. A. Pando, M. Negrete, K. L. Russell, J. Sanchez, D. L. Birx, R. Andrade, J. Vinoles and F. E. McCutchan

(2001). "Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America." AIDS **15**(15): F41-7.

Castanotto, D., H. Li and J. J. Rossi (2002). "Functional siRNA expression from transfected PCR products." Rna **8**(11): 1454-60.

Castanotto, D., K. Sakurai, R. Lingeman, H. Li, L. Shively, L. Aagaard, H. Soifer, A. Gatignol, A. Riggs and J. J. Rossi (2007). "Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC." Nucleic Acids Res **35**(15): 5154-64.

Castanotto, D., S. Tommasi, M. Li, H. Li, S. Yanow, G. P. Pfeifer and J. J. Rossi (2005). "Short hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells." Mol Ther **12**(1): 179-83.

Chan, A. Y., F. T. Vreede, M. Smith, O. G. Engelhardt and E. Fodor (2006). "Influenza virus inhibits RNA polymerase II elongation." Virology **351**(1): 210-7.

Chang, L. J., X. Liu and J. He (2005). "Lentiviral siRNAs targeting multiple highly conserved RNA sequences of human immunodeficiency virus type 1." Gene Ther **12**(14): 1133-44.

Chen, E. S., K. Zhang, E. Nicolas, H. P. Cam, M. Zofall and S. I. Grewal (2008). "Cell cycle control of centromeric repeat transcription and heterochromatin assembly." Nature **451**(7179): 734-7.

Chen, Y., Y. Yamaguchi, Y. Tsugeno, J. Yamamoto, T. Yamada, M. Nakamura, K. Hisatake and H. Handa (2009). "DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation." Genes Dev **23**(23): 2765-77.

Chendrimada, T. P., R. I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura and R. Shiekhattar (2005). "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." Nature **436**(7051): 740-4.

Cheng, J., P. Kapranov, J. Drenkow, S. Dike, S. Brubaker, S. Patel, J. Long, D. Stern, H. Tammana, G. Helt, V. Sementchenko, A. Piccolboni, S. Bekiranov, D. K. Bailey, M. Ganesh, S. Ghosh, I. Bell, D. S. Gerhard and T. R. Gingeras (2005). "Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution." Science **308**(5725): 1149-54.

Cheng, S. W., K. P. Davies, E. Yung, R. J. Beltran, J. Yu and G. V. Kalpana (1999). "c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function." Nat Genet **22**(1): 102-5.

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 Mol Biol **408**: 223-41.

Cherepanov, P., G. Maertens, P. Proost, B. Devreese, J. Van Beeumen, Y. Engelborghs, E. De Clercq and Z. Debyser (2003). "HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells." J Biol Chem **278**(1): 372-81.

Chesebro, B., K. Wehrly, J. Nishio and S. Perryman (1992). "Macrophage-tropic human immunodeficiency virus isolates from different patients exhibit unusual V3 envelope sequence homogeneity in comparison with T-cell-tropic isolates: definition of critical amino acids involved in cell tropism." J Virol **66**(11): 6547-54.

Chiu, Y. L., H. Cao, J. M. Jacque, M. Stevenson and T. M. Rana (2004). "Inhibition of human immunodeficiency virus type 1 replication by RNA interference directed against human transcription elongation factor P-TEFb (CDK9/CyclinT1)." J Virol **78**(5): 2517-29.

Chiu, Y. L. and T. M. Rana (2002). "RNAi in human cells: basic structural and functional features of small interfering RNA." Mol Cell **10**(3): 549-61.

Cho, H., G. Orphanides, X. Sun, X. J. Yang, V. Ogryzko, E. Lees, Y. Nakatani and D. Reinberg (1998). "A human RNA polymerase II complex containing factors that modify chromatin structure." Mol Cell Biol **18**(9): 5355-63.

Chou, T. C. and P. Talalay (1984). "Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors." Adv Enzyme Regul **22**: 27-55.

Christ, F., A. Voet, A. Marchand, S. Nicolet, B. A. Desimmie, D. Marchand, D. Bardiot, N. J. Van der Veken, B. Van Remoortel, S. V. Strelkov, M. De Maeyer, P. Chaltin and Z. Debyser (2010).

"Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication." Nat Chem Biol **6**(6): 442-8.

Chu, Y., X. Yue, S. T. Younger, B. A. Janowski and D. R. Corey (2010). "Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter." Nucleic Acids Res.

Chun, T. W., J. S. Justement, D. Murray, C. W. Hallahan, J. Maenza, A. C. Collier, P. M. Sheth, R. Kaul, M. Ostrowski, S. Moir, C. Kovacs and A. S. Fauci (2010). "Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: implications for eradication." AIDS **24**(18): 2803-8.

Cifuentes, D., H. Xue, D. W. Taylor, H. Patnode, Y. Mishima, S. Cheloufi, E. Ma, S. Mane, G. J. Hannon, N. D. Lawson, S. A. Wolfe and A. J. Giraldez (2010). "A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity." Science **328**(5986): 1694-8.

Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker and F. Bushman (2005). "A role for LEDGF/p75 in targeting HIV DNA integration." Nat Med **11**(12): 1287-9.

Clouse, K. N., M. J. Luo, Z. Zhou and R. Reed (2001). "A Ran-independent pathway for export of spliced mRNA." Nat Cell Biol **3**(1): 97-9.

Cogburn, L. A., T. E. Porter, M. J. Duclos, J. Simon, S. C. Burgess, J. J. Zhu, H. H. Cheng, J. B. Dodgson and J. Burnside (2007). "Functional genomics of the chicken--a model organism." Poult Sci **86**(10): 2059-94.

Cole, C., A. Sobala, C. Lu, S. R. Thatcher, A. Bowman, J. W. Brown, P. J. Green, G. J. Barton and G. Hutvagner (2009). "Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs." RNA **15**(12): 2147-60.

Coley, W., R. Van Duyne, L. Carpio, I. Guendel, K. Kehn-Hall, S. Chevalier, A. Narayanan, T. Luu, N. Lee, Z. Klase and F. Kashanchi (2010). "Absence of DICER in monocytes and its regulation by HIV-1." J Biol Chem **285**(42): 31930-43.

Connor, R. I., B. K. Chen, S. Choe and N. R. Landau (1995). "Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes." Virology **206**(2): 935-44.

Corbeau, P. (2008). "Interfering RNA and HIV: reciprocal interferences." PLoS Pathog **4**(9): e1000162.

Core, L. J., J. J. Waterfall and J. T. Lis (2008). "Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters." Science **322**(5909): 1845-8.

Craig, E., Z. K. Zhang, K. P. Davies and G. V. Kalpana (2002). "A masked NES in INI1/hSNF5 mediates hCRM1-dependent nuclear export: implications for tumorigenesis." Embo J **21**(1-2): 31-42.

Cullen, B. R. (1993). "Does HIV-1 Tat induce a change in viral initiation rights?" Cell **73**(3): 417-20.

Czech, B. and G. J. Hannon (2011). "Small RNA sorting: matchmaking for Argonautes." Nat Rev Genet **12**(1): 19-31.

Das, S., J. Cano and G. V. Kalpana (2009). "Multimerization and DNA binding properties of INI1/hSNF5 and its functional significance." J Biol Chem **284**(30): 19903-14.

David, L., W. Huber, M. Granovskaia, J. Toedling, C. J. Palm, L. Bofkin, T. Jones, R. W. Davis and L. M. Steinmetz (2006). "A high-resolution map of transcription in the yeast genome." Proc Natl Acad Sci U S A **103**(14): 5320-5.

Dayton, A. I., J. G. Sodroski, C. A. Rosen, W. C. Goh and W. A. Haseltine (1986). "The trans-activator gene of the human T cell lymphotropic virus type III is required for replication." Cell **44**(6): 941-7.

De Ferrari, L. and S. Aitken (2006). "Mining housekeeping genes with a Naive Bayes classifier." BMC Genomics **7**: 277.

De Luca, L., S. Ferro, R. Gitto, M. L. Barreca, S. Agnello, F. Christ, Z. Debyser and A. Chimirri (2010). "Small molecules targeting the interaction between HIV-1 integrase and LEDGF/p75 cofactor." Bioorg Med Chem **18**(21): 7515-21.

De Rijck, J., K. Bartholomeeusen, H. Ceulemans, Z. Debyser and R. Gijsbers (2010). "High-resolution profiling of the LEDGF/p75 chromatin interaction in the ENCODE region." Nucleic Acids Res **38**(18): 6135-47.

de Vries, W. and B. Berkhout (2008). "RNAi suppressors encoded by pathogenic human viruses." Int J Biochem Cell Biol **40**(10): 2007-12.

DeCristofaro, M. F., B. L. Betz, W. Wang and B. E. Weissman (1999). "Alteration of hSNF5/INI1/BAF47 detected in rhabdoid cell lines and primary rhabdomyosarcomas but not Wilms' tumors." Oncogene **18**(52): 7559-65.

Denli, A. M., B. B. Tops, R. H. Plasterk, R. F. Ketting and G. J. Hannon (2004). "Processing of primary microRNAs by the Microprocessor complex." Nature **432**(7014): 231-5.

Derdeyn, C. A., J. M. Decker, F. Bibollet-Ruche, J. L. Mokili, M. Muldoon, S. A. Denham, M. L. Heil, F. Kasolo, R. Musonda, B. H. Hahn, G. M. Shaw, B. T. Korber, S. Allen and E. Hunter (2004). "Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission." Science **303**(5666): 2019-22.

Derdeyn, C. A., J. M. Decker, J. N. Sfakianos, X. Wu, W. A. O'Brien, L. Ratner, J. C. Kappes, G. M. Shaw and E. Hunter (2000). "Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120." J Virol **74**(18): 8358-67.

Dieudonne, M., P. Maiuri, C. Biancotto, A. Knezevich, A. Kula, M. Lusic and A. Marcello (2009). "Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery." EMBO J **28**(15): 2231-2243.

DiGiusto, D. L., A. Krishnan, L. Li, H. Li, S. Li, A. Rao, S. Mi, P. Yam, S. Stinson, M. Kalos, J. Alvarnas, S. F. Lacey, J. K. Yee, M. Li, L. Couture, D. Hsu, S. J. Forman, J. J. Rossi and J. A. Zaia (2010). "RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma." Sci Transl Med **2**(36): 36ra43.

Douek, D. C. (2003). "Disrupting T-cell homeostasis: how HIV-1 infection causes disease." AIDS Rev **5**(3): 172-7.

Dube, M., M. G. Bego, C. Paquay and E. A. Cohen (2010a). "Modulation of HIV-1-host interaction: role of the Vpu accessory protein." Retrovirology **7**(1): 114.

Dube, M., B. B. Roy, P. Guiot-Guillain, J. Binette, J. Mercier, A. Chiasson and E. A. Cohen (2010b). "Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment." PLoS Pathog **6**(4): e1000856.

Eekels, J. J., D. Geerts, R. E. Jeeninga and B. Berkhout (2011). "Long-term inhibition of HIV-1 replication with RNA interference against cellular co-factors." Antiviral Res **89**(1): 43-53.

Ehsani, A., P. Saetrom, J. Zhang, J. Alluin, H. Li, O. Snove, Jr., L. Aagaard and J. J. Rossi (2010). "Rational design of micro-RNA-like bifunctional siRNAs targeting HIV and the HIV coreceptor CCR5." Mol Ther **18**(4): 796-802.

Eisenberg, E. and E. Y. Levanon (2003). "Human housekeeping genes are compact." Trends Genet **19**(7): 362-5.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl (2001). "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**(6836): 494-8.

Emerman, M. and M. H. Malim (1998). "HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology." Science **280**(5371): 1880-4.

Emmerth, S., H. Schober, D. Gaidatzis, T. Roloff, K. Jacobeit and M. Buhler (2010). "Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly." Dev Cell **18**(1): 102-13.

Ender, C., A. Krek, M. R. Friedlander, M. Beitzinger, L. Weinmann, W. Chen, S. Pfeffer, N. Rajewsky and G. Meister (2008). "A human snoRNA with microRNA-like functions." Mol Cell **32**(4): 519-28.

Fanales-Belasio, E., M. Raimondo, B. Suligoi and S. Butto (2010). "HIV virology and pathogenetic mechanisms of infection: a brief overview." Ann Ist Super Sanita **46**(1): 5-14.

Fassati, A., D. Gorlich, I. Harrison, L. Zaytseva and J. M. Mingot (2003). "Nuclear import of HIV-1 intracellular reverse transcription complexes is mediated by importin 7." EMBO J **22**(14): 3675-85.

Feinberg, M. B., D. Baltimore and A. D. Frankel (1991). "The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation." Proc Natl Acad Sci U S A **88**(9): 4045-9.

Filardi, P. P., S. Paolillo, C. Marciano, A. Iorio, T. Losco, F. Marsico, O. Scala, D. Ruggiero, S. Ferraro and M. Chiariello (2008). "Cardiovascular effects of antiretroviral drugs: clinical review." Cardiovasc Hematol Disord Drug Targets **8**(4): 238-44.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans." Nature **391**(6669): 806-11.

Fisher, A. G., M. B. Feinberg, S. F. Josephs, M. E. Harper, L. M. Marselle, G. Reyes, M. A. Gonda, A. Aldovini, C. Debouk, R. C. Gallo and et al. (1986). "The trans-activator gene of HTLV-III is essential for virus replication." Nature **320**(6060): 367-71.

Fletcher, B. S. (2010). "Delivery of small interfering RNA (siRNA) using the sleeping beauty transposon." Cold Spring Harb Protoc **2010**(11): pdb prot5521.

Fong, Y. W. and Q. Zhou (2001). "Stimulatory effect of splicing factors on transcriptional elongation." Nature **414**(6866): 929-33.

Fornerod, M., M. Ohno, M. Yoshida and I. W. Mattaj (1997). "CRM1 is an export receptor for leucine-rich nuclear export signals." Cell **90**(6): 1051-60.

Forrest, A. R., R. F. Abdelhamid and P. Carninci (2009). "Annotating non-coding transcription using functional genomics strategies." Brief Funct Genomic Proteomic **8**(6): 437-43.

Frankel, A. D. and J. A. Young (1998). "HIV-1: fifteen proteins and an RNA." Annu Rev Biochem **67**: 1-25.

Gallo, R. C., P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch and M. Popovic (1983). "Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS)." Science **220**(4599): 865-7.

Ganser-Pornillos, B. K., M. Yeager and W. I. Sundquist (2008). "The structural biology of HIV assembly." Curr Opin Struct Biol **18**(2): 203-17.

Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp and B. H. Hahn (1999). "Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes." Nature **397**(6718): 436-41.

Gao, Y., M. A. Lobritz, J. Roth, M. Abreha, K. N. Nelson, I. Nankya, D. M. Moore-Dudley, A. Abraha, S. L. Gerson and E. J. Arts (2008). "Targets of small interfering RNA restriction during human immunodeficiency virus type 1 replication." J Virol **82**(6): 2938-51.

Garcia-Sastre, A. and C. A. Biron (2006). "Type 1 interferons and the virus-host relationship: a lesson in detente." Science **312**(5775): 879-82.

Gardiner-Garden, M. and M. Frommer (1987). "CpG islands in vertebrate genomes." J Mol Biol **196**(2): 261-82.

Ge, H., Y. Si and R. G. Roeder (1998). "Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation." EMBO J **17**(22): 6723-9.

Ge, Q., H. Ilves, A. Dallas, P. Kumar, J. Shorenstein, S. A. Kazakov and B. H. Johnston (2010). "Minimal-length short hairpin RNAs: the relationship of structure and RNAi activity." RNA **16**(1): 106-17.

Geiduschek, E. P. and G. A. Kassavetis (2001). "The RNA polymerase III transcription apparatus." J Mol Biol **310**(1): 1-26.

Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nat Rev Genet **10**(2): 94-108.

Giering, J. C., D. Grimm, T. A. Storm and M. A. Kay (2008). "Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic." Mol Ther **16**(9): 1630-6.

Gijsbers, R., K. Ronen, S. Vets, N. Malani, J. De Rijck, M. McNeely, F. D. Bushman and Z. Debyser (2010). "LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin." Mol Ther **18**(3): 552-60.

Giles, K. E., R. Ghirlando and G. Felsenfeld (2010). "Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism." Nat Cell Biol **12**(1): 94-9; sup pp 1-6.

Goff, S. P. (2008). "Knockdown screens to knockout HIV-1." Cell **135**(3): 417-20.

Gonzalez, S., D. G. Pisano and M. Serrano (2008). "Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs." Cell Cycle **7**(16): 2601-8.

Gregory, R. I., K. P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch and R. Shiekhattar (2004). "The Microprocessor complex mediates the genesis of microRNAs." Nature **432**(7014): 235-40.

Gresh, L., B. Bourachot, A. Reimann, B. Guigas, L. Fiette, S. Garbay, C. Muchardt, L. Hue, M. Pontoglio, M. Yaniv and A. Klochendler-Yeivin (2005). "The SWI/SNF chromatin-remodeling complex subunit SNF5 is essential for hepatocyte differentiation." EMBO J **24**(18): 3313-24.

Grewal, S. I. and S. C. Elgin (2007). "Transcription and RNA interference in the formation of heterochromatin." Nature **447**(7143): 399-406.

Grimm, D. and M. A. Kay (2007). "Combinatorial RNAi: a winning strategy for the race against evolving targets?" Mol Ther **15**(5): 878-88.

Grimm, D., K. L. Streetz, C. L. Jopling, T. A. Storm, K. Pandey, C. R. Davis, P. Marion, F. Salazar and M. A. Kay (2006). "Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways." Nature **441**(7092): 537-41.

Grimm, D., L. Wang, J. S. Lee, N. Schurmann, S. Gu, K. Borner, T. A. Storm and M. A. Kay (2010). "Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver." J Clin Invest **120**(9): 3106-19.

Guang, S., A. F. Bochner, K. B. Burkhart, N. Burton, D. M. Pavelec and S. Kennedy (2010). "Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription." Nature **465**(7301): 1097-101.

Guang, S., A. F. Bochner, D. M. Pavelec, K. B. Burkhart, S. Harding, J. Lachowiec and S. Kennedy (2008). "An Argonaute transports siRNAs from the cytoplasm to the nucleus." Science **321**(5888): 537-41.

Gullerova, M. and N. J. Proudfoot (2008). "Cohesin complex promotes transcriptional termination between convergent genes in S. pombe." Cell **132**(6): 983-95.

Guo, H., N. T. Ingolia, J. S. Weissman and D. P. Bartel (2010). "Mammalian microRNAs predominantly act to decrease target mRNA levels." Nature **466**(7308): 835-40.

Hahn, M., S. Dambacher and G. Schotta (2010). "Heterochromatin dysregulation in human diseases." J Appl Physiol **109**(1): 232-42.

Han, J., D. Kim and K. V. Morris (2007). "Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells." Proc Natl Acad Sci U S A **104**(30): 12422-7.

Han, J., Y. Lee, K. H. Yeom, Y. K. Kim, H. Jin and V. N. Kim (2004). "The Drosha-DGCR8 complex in primary microRNA processing." Genes Dev **18**(24): 3016-27.

Han, J., Y. Lee, K. H. Yeom, J. W. Nam, I. Heo, J. K. Rhee, S. Y. Sohn, Y. Cho, B. T. Zhang and V. N. Kim (2006). "Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex." Cell **125**(5): 887-901.

Hanahan, D. (1983). "Studies on transformation of Escherichia coli with plasmids." J Mol Biol **166**(4): 557-80.

Harborth, J., S. M. Elbashir, K. Vandenburgh, H. Manninga, S. A. Scaringe, K. Weber and T. Tuschl (2003). "Sequence, chemical, and structural variation of small interfering RNAs and short

hairpin RNAs and the effect on mammalian gene silencing." Antisense Nucleic Acid Drug Dev **13**(2): 83-105.

Hariharan, M., V. Scaria, B. Pillai and S. K. Brahmachari (2005). "Targets for human encoded microRNAs in HIV genes." Biochem Biophys Res Commun **337**(4): 1214-8.

Harris, M. E. and T. J. Hope (2000). "RNA export: insights from viral models." Essays Biochem **36**: 115-27.

Hawkins, P. G. and K. V. Morris (2010). "Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5." Transcr **1**(3): 165-175.

Hawkins, P. G., S. Santoso, C. Adams, V. Anest and K. V. Morris (2009). "Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells." Nucleic Acids Res **37**(9): 2984-95.

Hayouka, Z., M. Hurevich, A. Levin, H. Benyamini, A. Iosub, M. Maes, D. E. Shalev, A. Loyter, C. Gilon and A. Friedler (2010a). "Cyclic peptide inhibitors of HIV-1 integrase derived from the LEDGF/p75 protein." Bioorg Med Chem **18**(23): 8388-95.

Hayouka, Z., A. Levin, M. Maes, E. Hadas, D. E. Shalev, D. J. Volsky, A. Loyter and A. Friedler (2010b). "Mechanism of action of the HIV-1 integrase inhibitory peptide LEDGF 361-370." Biochem Biophys Res Commun **394**(2): 260-5.

He, G., L. Ylisastigui and D. M. Margolis (2002). "The regulation of HIV-1 gene expression: the emerging role of chromatin." DNA Cell Biol **21**(10): 697-705.

He, J., S. Choe, R. Walker, P. Di Marzio, D. O. Morgan and N. R. Landau (1995). "Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity." J Virol **69**(11): 6705-11.

Hehl, E. A., P. Joshi, G. V. Kalpana and V. R. Prasad (2004). "Interaction between human immunodeficiency virus type 1 reverse transcriptase and integrase proteins." J Virol **78**(10): 5056-67.

Hendrix, J., R. Gijsbers, J. De Rijck, A. Voet, J. I. Hotta, M. McNeely, J. Hofkens, Z. Debyser and Y. Engelborghs (2010). "The transcriptional co-activator LEDGF/p75 displays a dynamic scan-and-lock mechanism for chromatin tethering." Nucleic Acids Res.

Hill, M., G. Tachedjian and J. Mak (2005). "The packaging and maturation of the HIV-1 Pol proteins." Curr HIV Res **3**(1): 73-85.

Hirsch, A. J. (2010). "The use of RNAi-based screens to identify host proteins involved in viral replication." Future Microbiol **5**(2): 303-11.

Hirsch, M. S., H. F. Gunthard, J. M. Schapiro, F. Brun-Vezinet, B. Clotet, S. M. Hammer, V. A. Johnson, D. R. Kuritzkes, J. W. Mellors, D. Pillay, P. G. Yeni, D. M. Jacobsen and D. D. Richman (2008). "Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society-USA panel." Top HIV Med **16**(3): 266-85.

Ho, D. D. (1997). "Perspectives series: host/pathogen interactions. Dynamics of HIV-1 replication in vivo." J Clin Invest **99**(11): 2565-7.

Hoffer, P., S. Ivashuta, O. Pontes, A. Vitins, C. Pikaard, A. Mroczka, N. Wagner and T. Voelker (2011). "Posttranscriptional gene silencing in nuclei." Proc Natl Acad Sci U S A **108**(1): 409-14.

Hombrouck, A., J. De Rijck, J. Hendrix, L. Vandekerckhove, A. Voet, M. De Maeyer, M. Witvrouw, Y. Engelborghs, F. Christ, R. Gijsbers and Z. Debyser (2007). "Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV." PLoS Pathog **3**(3): e47.

Hong, D., W. Lu, F. Ye, Y. Hu and X. Xie (2009). "Gene silencing of HPV16 E6/E7 induced by promoter-targeting siRNA in SiHa cells." Br J Cancer **101**(10): 1798-804.

Hornung, V., J. Ellegast, S. Kim, K. Brzozka, A. Jung, H. Kato, H. Poeck, S. Akira, K. K. Conzelmann, M. Schlee, S. Endres and G. Hartmann (2006). "5'-Triphosphate RNA is the ligand for RIG-I." Science **314**(5801): 994-7.

Huang da, W., B. T. Sherman and R. A. Lempicki (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc **4**(1): 44-57.

Huang, J., F. Wang, E. Argyris, K. Chen, Z. Liang, H. Tian, W. Huang, K. Squires, G. Verlinghieri and H. Zhang (2007). "Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes." Nat Med **13**(10): 1241-7.

Huang, Y., W. A. Paxton, S. M. Wolinsky, A. U. Neumann, L. Zhang, T. He, S. Kang, D. Ceradini, Z. Jin, K. Yazdanbakhsh, K. Kunstman, D. Erickson, E. Dragon, N. R. Landau, J. Phair, D. D. Ho and R. A. Koup (1996). "The role of a mutant CCR5 allele in HIV-1 transmission and disease progression." Nat Med **2**(11): 1240-3.

Hutter, G., D. Nowak, M. Mossner, S. Ganepola, A. Mussig, K. Allers, T. Schneider, J. Hofmann, C. Kucherer, O. Blau, I. W. Blau, W. K. Hofmann and E. Thiel (2009). "Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation." N Engl J Med **360**(7): 692-8.

Hutvagner, G., J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschl and P. D. Zamore (2001). "A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA." Science **293**(5531): 834-8.

Iida, T., J. Nakayama and D. Moazed (2008). "siRNA-mediated heterochromatin establishment requires HP1 and is associated with antisense transcription." Mol Cell **31**(2): 178-89.

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

Isakoff, M. S., C. G. Sansam, P. Tamayo, A. Subramanian, J. A. Evans, C. M. Fillmore, X. Wang, J. A. Biegel, S. L. Pomeroy, J. P. Mesirov and C. W. Roberts (2005). "Inactivation of the Snf5 tumor suppressor stimulates cell cycle progression and cooperates with p53 loss in oncogenic transformation." Proc Natl Acad Sci U S A **102**(49): 17745-50.

Ishaq, M., J. Hu, X. Wu, Q. Fu, Y. Yang, Q. Liu and D. Guo (2008). "Knockdown of cellular RNA helicase DDX3 by short hairpin RNAs suppresses HIV-1 viral replication without inducing apoptosis." Mol Biotechnol **39**(3): 231-8.

Jablonski, J. A., A. L. Amelio, M. Giacca and M. Caputi (2010). "The transcriptional transactivator Tat selectively regulates viral splicing." Nucleic Acids Res **38**(4): 1249-60.

Jackson, A. L., S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley (2003). "Expression profiling reveals off-target gene regulation by RNAi." Nat Biotechnol **21**(6): 635-7.

Jackson, A. L., J. Burchard, J. Schelter, B. N. Chau, M. Cleary, L. Lim and P. S. Linsley (2006). "Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity." RNA **12**(7): 1179-87.

Jackson, A. L. and P. S. Linsley (2004). "Noise amidst the silence: off-target effects of siRNAs?" Trends Genet **20**(11): 521-4.

Jacox, E., V. Gotea, I. Ovcharenko and L. Elnitski (2010). "Tissue-specific and ubiquitous expression patterns from alternative promoters of human genes." PLoS One **5**(8): e12274.

Jacque, J. M. and M. Stevenson (2006). "The inner-nuclear-envelope protein emerin regulates HIV-1 infectivity." Nature **441**(7093): 641-5.

Jain, A., M. Magistri, S. Napoli, G. M. Carbone and C. V. Catapano (2010). "Mechanisms of triplex DNA-mediated inhibition of transcription initiation in cells." Biochimie **92**(3): 317-20.

Jannot, G., M. E. Boisvert, I. H. Banville and M. J. Simard (2008). "Two molecular features contribute to the Argonaute specificity for the microRNA and RNAi pathways in C. elegans." RNA **14**(5): 829-35.

Janowski, B. A., K. E. Huffman, J. C. Schwartz, R. Ram, D. Hardy, D. S. Shames, J. D. Minna and D. R. Corey (2005). "Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs." Nat Chem Biol **1**(4): 216-22.

Janowski, B. A., K. E. Huffman, J. C. Schwartz, R. Ram, R. Nordsell, D. S. Shames, J. D. Minna and D. R. Corey (2006). "Involvement of AGO1 and AGO2 in mammalian transcriptional silencing." Nat Struct Mol Biol **13**(9): 787-92.

Janowski, B. A., S. T. Younger, D. B. Hardy, R. Ram, K. E. Huffman and D. R. Corey (2007). "Activating gene expression in mammalian cells with promoter-targeted duplex RNAs." Nat Chem Biol **3**(3): 166-73.

Jeang, K. T., Y. Chang, B. Berkhout, M. L. Hammarskjold and D. Rekosh (1991). "Regulation of HIV expression: mechanisms of action of Tat and Rev." AIDS **5 Suppl 2**: S3-14.

Jere, A., M. Fujita, A. Adachi and M. Nomaguchi (2010). "Role of HIV-1 Nef protein for virus replication in vitro." Microbes Infect **12**(1): 65-70.

Jiang, C. and B. F. Pugh (2009). "Nucleosome positioning and gene regulation: advances through genomics." Nat Rev Genet **10**(3): 161-72.

Jiang, C., Z. Xuan, F. Zhao and M. Q. Zhang (2007). "TRED: a transcriptional regulatory element database, new entries and other development." Nucleic Acids Res **35**(Database issue): D137-40.

Juven-Gershon, T., J. Y. Hsu, J. W. Theisen and J. T. Kadonaga (2008). "The RNA polymerase II core promoter - the gateway to transcription." Curr Opin Cell Biol **20**(3): 253-9.

Kahan, B. D., S. Gibbons, N. Tejpal, T. C. Chou and S. Stepkowski (1991). "Synergistic effect of the rapamycin-cyclosporine combination: median effect analysis of in vitro immune performances by human T lymphocytes in PHA, CD3, and MLR proliferative and cytotoxicity assays." Transplant Proc **23**(1 Pt 2): 1090-1.

Kalpana, G. V., S. Marmon, W. Wang, G. R. Crabtree and S. P. Goff (1994). "Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5." Science **266**(5193): 2002-6.

Kalverda, A. P., G. S. Thompson, A. Vogel, M. Schroder, A. G. Bowie, A. R. Khan and S. W. Homans (2009). "Poxvirus K7 protein adopts a Bcl-2 fold: biochemical mapping of its interactions with human DEAD box RNA helicase DDX3." J Mol Biol **385**(3): 843-53.

Kampa, D., J. Cheng, P. Kapranov, M. Yamanaka, S. Brubaker, S. Cawley, J. Drenkow, A. Piccolboni, S. Bekiranov, G. Helt, H. Tammana and T. R. Gingeras (2004). "Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22." Genome Res **14**(3): 331-42.

Kao, S. Y., A. F. Calman, P. A. Luciw and B. M. Peterlin (1987). "Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product." Nature **330**(6147): 489-93.

Kapranov, P., S. E. Cawley, J. Drenkow, S. Bekiranov, R. L. Strausberg, S. P. Fodor and T. R. Gingeras (2002). "Large-scale transcriptional activity in chromosomes 21 and 22." Science **296**(5569): 916-9.

Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermuller, I. L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana and T. R. Gingeras (2007a). "RNA maps reveal new RNA classes and a possible function for pervasive transcription." Science **316**(5830): 1484-8.

Kapranov, P., A. T. Willingham and T. R. Gingeras (2007b). "Genome-wide transcription and the implications for genomic organization." Nat Rev Genet **8**(6): 413-23.

Karpala, A. J., T. J. Doran and A. G. Bean (2005). "Immune responses to dsRNA: implications for gene silencing technologies." Immunol Cell Biol **83**(3): 211-6.

Kaul, D., A. Ahlawat and S. D. Gupta (2009). "HIV-1 genome-encoded hiv1-mir-H1 impairs cellular responses to infection." Mol Cell Biochem **323**(1-2): 143-8.

Kawaji, H., M. C. Frith, S. Katayama, A. Sandelin, C. Kai, J. Kawai, P. Carninci and Y. Hayashizaki (2006a). "Dynamic usage of transcription start sites within core promoters." Genome Biol **7**(12): R118.

Kawaji, H., T. Kasukawa, S. Fukuda, S. Katayama, C. Kai, J. Kawai, P. Carninci and Y. Hayashizaki (2006b). "CAGE Basic/Analysis Databases: the CAGE resource for comprehensive promoter analysis." Nucleic Acids Res **34**(Database issue): D632-6.

Kawaji, H., J. Severin, M. Lizio, A. R. Forrest, E. van Nimwegen, M. Rehli, K. Schroder, K. Irvine, H. Suzuki, P. Carninci, Y. Hayashizaki and C. O. Daub (2010). "Update of the FANTOM web

resource: from mammalian transcriptional landscape to its dynamic regulation." Nucleic Acids Res.

Kawaji, H., J. Severin, M. Lizio, A. Waterhouse, S. Katayama, K. M. Irvine, D. A. Hume, A. R. Forrest, H. Suzuki, P. Carninci, Y. Hayashizaki and C. O. Daub (2009). "The FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation." Genome Biol **10**(4): R40.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler and D. Haussler (2002). "The human genome browser at UCSC." Genome Res **12**(6): 996-1006.

Khvorova, A., A. Reynolds and S. D. Jayasena (2003). "Functional siRNAs and miRNAs exhibit strand bias." Cell **115**(2): 209-16.

Kim, D. H., P. Saetrom, O. Snove, Jr. and J. J. Rossi (2008). "MicroRNA-directed transcriptional gene silencing in mammalian cells." Proc Natl Acad Sci U S A **105**(42): 16230-5.

Kim, D. H., L. M. Villeneuve, K. V. Morris and J. J. Rossi (2006). "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells." Nat Struct Mol Biol **13**(9): 793-7.

Kim, J. B., Y. Yamaguchi, T. Wada, H. Handa and P. A. Sharp (1999). "Tat-SF1 protein associates with RAP30 and human SPT5 proteins." Mol Cell Biol **19**(9): 5960-8.

Kim, J. W., Y. H. Zhang, M. A. Zern, J. J. Rossi and J. Wu (2007). "Short hairpin RNA causes the methylation of transforming growth factor-beta receptor II promoter and silencing of the target gene in rat hepatic stellate cells." Biochem Biophys Res Commun **359**(2): 292-7.

Kim, T. H., L. O. Barrera, M. Zheng, C. Qu, M. A. Singer, T. A. Richmond, Y. Wu, R. D. Green and B. Ren (2005). "A high-resolution map of active promoters in the human genome." Nature **436**(7052): 876-80.

Kim, V. N., J. Han and M. C. Siomi (2009). "Biogenesis of small RNAs in animals." Nat Rev Mol Cell Biol **10**(2): 126-39.

Kim, Y. K. and V. N. Kim (2007). "Processing of intronic microRNAs." EMBO J **26**(3): 775-83.

Klase, Z., P. Kale, R. Winograd, M. V. Gupta, M. Heydarian, R. Berro, T. McCaffrey and F. Kashanchi (2007). "HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR." BMC Mol Biol **8**: 63.

Klase, Z., R. Winograd, J. Davis, L. Carpio, R. Hildreth, M. Heydarian, S. Fu, T. McCaffrey, E. Meiri, M. Ayash-Rashkovsky, S. Gilad, Z. Bentwich and F. Kashanchi (2009). "HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression." Retrovirology **6**: 18.

Kleinman, M. E., K. Yamada, A. Takeda, V. Chandrasekaran, M. Nozaki, J. Z. Baffi, R. J. Albuquerque, S. Yamasaki, M. Itaya, Y. Pan, B. Appukuttan, D. Gibbs, Z. Yang, K. Kariko, B. K. Ambati, T. A. Wilgus, L. A. DiPietro, E. Sakurai, K. Zhang, J. R. Smith, E. W. Taylor and J. Ambati (2008). "Sequence- and target-independent angiogenesis suppression by siRNA via TLR3." Nature **452**(7187): 591-7.

Knudson, A. G., Jr. (1971). "Mutation and cancer: statistical study of retinoblastoma." Proc Natl Acad Sci U S A **68**(4): 820-3.

Komano, J., Y. Futahashi, E. Urano, K. Miyauchi, T. Murakami, Z. Matsuda and N. Yamamoto (2005). "The interaction of HIV-1 with the host factors." Jpn J Infect Dis **58**(3): 125-30.

Konig, R., Y. Zhou, D. Elleder, T. L. Diamond, G. M. Bonamy, J. T. Irelan, C. Y. Chiang, B. P. Tu, P. D. De Jesus, C. E. Lilley, S. Seidel, A. M. Opaluch, J. S. Caldwell, M. D. Weitzman, K. L. Kuhen, S. Bandyopadhyay, T. Ideker, A. P. Orth, L. J. Miraglia, F. D. Bushman, J. A. Young and S. K. Chanda (2008). "Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication." Cell **135**(1): 49-60.

Korber, B. and S. Gnanakaran (2009). "The implications of patterns in HIV diversity for neutralizing antibody induction and susceptibility." Curr Opin HIV AIDS **4**(5): 408-17.

Kornberg, R. D. and A. Klug (1981). "The nucleosome." Sci Am **244**(2): 52-64.

Krishnan, V. and S. L. Zeichner (2004). "Alterations in the expression of DEAD-box and other RNA binding proteins during HIV-1 replication." Retrovirology **1**: 42.

Kuhmann, S. E., E. J. Platt, S. L. Kozak and D. Kabat (2000). "Cooperation of multiple CCR5 coreceptors is required for infections by human immunodeficiency virus type 1." J Virol **74**(15): 7005-15.

Kumar, P., H. S. Ban, S. S. Kim, H. Wu, T. Pearson, D. L. Greiner, A. Laouar, J. Yao, V. Haridas, K. Habiro, Y. G. Yang, J. H. Jeong, K. Y. Lee, Y. H. Kim, S. W. Kim, M. Peipp, G. H. Fey, N. Manjunath, L. D. Shultz, S. K. Lee and P. Shankar (2008). "T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice." Cell **134**(4): 577-86.

Lai, M. C., W. C. Chang, S. Y. Shieh and W. Y. Tarn (2010). "DDX3 regulates cell growth through translational control of cyclin E1." Mol Cell Biol **30**(22): 5444-53.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blocker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V. Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi and Y. J. Chen (2001). "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.

Landthaler, M., A. Yalcin and T. Tuschl (2004). "The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis." Curr Biol **14**(23): 2162-7.

Lange, C. G., M. M. Lederman, K. Medvik, R. Asaad, M. Wild, R. Kalayjian and H. Valdez (2003). "Nadir CD4+ T-cell count and numbers of CD28+ CD4+ T-cells predict functional responses to immunizations in chronic HIV-1 infection." AIDS **17**(14): 2015-23.

Lares, M. R., J. J. Rossi and D. L. Ouellet (2010). "RNAi and small interfering RNAs in human disease therapeutic applications." Trends Biotechnol **28**(11): 570-9.

Laspia, M. F., A. P. Rice and M. B. Mathews (1989). "HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation." Cell **59**(2): 283-92.

Leal, C., N. F. Bouxsein, K. K. Ewert and C. R. Safinya (2010). "Highly efficient gene silencing activity of siRNA embedded in a nanostructured gyroid cubic lipid matrix." J Am Chem Soc **132**(47): 16841-7.

Lee, D., H. Sohn, G. V. Kalpana and J. Choe (1999). "Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA." Nature **399**(6735): 487-91.

Lee, N. S., D. H. Kim, J. Alluin, M. Robbins, S. Gu, H. Li, J. Kim, P. M. Salvaterra and J. J. Rossi (2008). "Functional and intracellular localization properties of U6 promoter-expressed siRNAs, shRNAs, and chimeric VA1 shRNAs in mammalian cells." RNA **14**(9): 1823-33.

Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75**(5): 843-54.

Lee, Y., I. Hur, S. Y. Park, Y. K. Kim, M. R. Suh and V. N. Kim (2006). "The role of PACT in the RNA silencing pathway." Embo J **25**(3): 522-32.

Lee, Y., K. Jeon, J. T. Lee, S. Kim and V. N. Kim (2002). "MicroRNA maturation: stepwise processing and subcellular localization." EMBO J **21**(17): 4663-70.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek and V. N. Kim (2004). "MicroRNA genes are transcribed by RNA polymerase II." EMBO J **23**(20): 4051-60.

Lee, Y. S., Y. Shibata, A. Malhotra and A. Dutta (2009). "A novel class of small RNAs: tRNA-derived RNA fragments (tRFs)." Genes Dev **23**(22): 2639-49.

Lehrman, G., I. B. Hogue, S. Palmer, C. Jennings, C. A. Spina, A. Wiegand, A. L. Landay, R. W. Coombs, D. D. Richman, J. W. Mellors, J. M. Coffin, R. J. Bosch and D. M. Margolis (2005). "Depletion of latent HIV-1 infection in vivo: a proof-of-concept study." Lancet **366**(9485): 549-55.

Leng, Q., L. Goldgeier, J. Zhu, P. Cambell, N. Ambulos and A. J. Mixson (2007). "Histidine-lysine peptides as carriers of nucleic acids." Drug News Perspect **20**(2): 77-86.

Leonard, J. N. and D. V. Schaffer (2005). "Computational design of antiviral RNA interference strategies that resist human immunodeficiency virus escape." J Virol **79**(3): 1645-54.

Levin, A., Z. Hayouka, A. Friedler and A. Loyter (2010a). "Peptides derived from the HIV-1 integrase promote HIV-1 infection and multi-integration of viral cDNA in LEDGF/p75-knockdown cells." Virol J **7**: 177.

Levin, A., J. Rosenbluh, Z. Hayouka, A. Friedler and A. Loyter (2010b). "Integration of HIV-1 DNA is regulated by interplay between viral rev and cellular LEDGF/p75 proteins." Mol Med **16**(1-2): 34-44.

Li, J., Y. Liu, T. Kim, R. Min and Z. Zhang (2010). "Gene expression variability within and between human populations and implications toward disease susceptibility." PLoS Comput Biol **6**(8).

Li, L., J. M. Olvera, K. E. Yoder, R. S. Mitchell, S. L. Butler, M. Lieber, S. L. Martin and F. D. Bushman (2001). "Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection." EMBO J **20**(12): 3272-81.

Li, L. C., S. T. Okino, H. Zhao, D. Pookot, R. F. Place, S. Urakami, H. Enokida and R. Dahiya (2006). "Small dsRNAs induce transcriptional activation in human cells." Proc Natl Acad Sci U S A **103**(46): 17337-42.

Li, M. J., J. Kim, S. Li, J. Zaia, J. K. Yee, J. Anderson, R. Akkina and J. J. Rossi (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." Mol Ther **12**(5): 900-9.

Li, X. Y. and M. R. Green (1998). "The HIV-1 Tat cellular coactivator Tat-SF1 is a general transcription elongation factor." Genes Dev **12**(19): 2992-6.

Lichterfeld, M., D. E. Kaufmann, X. G. Yu, S. K. Mui, M. M. Addo, M. N. Johnston, D. Cohen, G. K. Robbins, E. Pae, G. Alter, A. Wurcel, D. Stone, E. S. Rosenberg, B. D. Walker and M. Altfeld (2004). "Loss of HIV-1-specific CD8+ T cell proliferation after acute HIV-1 infection and restoration by vaccine-induced HIV-1-specific CD4+ T cells." J Exp Med **200**(6): 701-12.

Lim, H. G., K. Suzuki, D. A. Cooper and A. D. Kelleher (2008). "Promoter-targeted siRNAs induce gene silencing of simian immunodeficiency virus (SIV) infection in vitro." Mol Ther **16**(3): 565-70.

Liu, R., W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup and N. R. Landau (1996). "Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection." Cell **86**(3): 367-77.

Liu, S. Q., S. X. Liu and Y. X. Fu (2008). "Molecular motions of human HIV-1 gp120 envelope glycoproteins." J Mol Model **14**(9): 857-70.

Liu, Y. P., J. Gruber, J. Haasnoot, P. Konstantinova and B. Berkhout (2009). "RNAi-mediated inhibition of HIV-1 by targeting partially complementary viral sequences." Nucleic Acids Res **37**(18): 6194-204.

Liu, Y. P., J. Haasnoot and B. Berkhout (2007). "Design of extended short hairpin RNAs for HIV-1 inhibition." Nucleic Acids Res **35**(17): 5683-93.

Liu, Y. P., M. A. Vink, J. T. Westerink, E. Ramirez de Arellano, P. Konstantinova, O. Ter Brake and B. Berkhout (2010). "Titers of lentiviral vectors encoding shRNAs and miRNAs are reduced by different mechanisms that require distinct repair strategies." RNA **16**(7): 1328-39.

Llano, M., D. T. Saenz, A. Meehan, P. Wongthida, M. Peretz, W. H. Walker, W. Teo and E. M. Poeschla (2006a). "An essential role for LEDGF/p75 in HIV integration." Science **314**(5798): 461-4.

Llano, M., M. Vanegas, O. Fregoso, D. Saenz, S. Chung, M. Peretz and E. M. Poeschla (2004). "LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes." J Virol **78**(17): 9524-37.

Llano, M., M. Vanegas, N. Hutchins, D. Thompson, S. Delgado and E. M. Poeschla (2006b). "Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75." J Mol Biol **360**(4): 760-73.

Loseke, S., E. Grage-Griebenow, H. Heine, A. Wagner, S. Akira, S. Bauer and A. Bufe (2006). "In vitro-generated viral double-stranded RNA in contrast to polyinosinic:polycytidylic acid induces interferon-alpha in human plasmacytoid dendritic cells." Scand J Immunol **63**(4): 264-74.

Luco, R. F., M. Allo, I. E. Schor, A. R. Kornblihtt and T. Misteli (2011). "Epigenetics in Alternative Pre-mRNA Splicing." Cell **144**(1): 16-26.

Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature **389**(6648): 251-60.

Lusso, P. (2000). "Chemokines and viruses: the dearest enemies." Virology **273**(2): 228-40.

Ma, J. B., Y. R. Yuan, G. Meister, Y. Pei, T. Tuschl and D. J. Patel (2005). "Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein." Nature **434**(7033): 666-70.

MacRae, I. J., E. Ma, M. Zhou, C. V. Robinson and J. A. Doudna (2008). "In vitro reconstitution of the human RISC-loading complex." Proc Natl Acad Sci U S A **105**(2): 512-7.

Maeda, N., H. Nishiyori, M. Nakamura, C. Kawazu, M. Murata, H. Sano, K. Hayashida, S. Fukuda, M. Tagami, A. Hasegawa, K. Murakami, K. Schroder, K. Irvine, D. Hume, Y. Hayashizaki, P. Carninci and H. Suzuki (2008). "Development of a DNA barcode tagging method for monitoring dynamic changes in gene expression by using an ultra high-throughput sequencer." Biotechniques **45**(1): 95-7.

Maertens, G., P. Cherepanov, W. Pluymers, K. Busschots, E. De Clercq, Z. Debyser and Y. Engelborghs (2003). "LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells." J Biol Chem **278**(35): 33528-39.

Mahmoudi, T., M. Parra, R. G. Vries, S. E. Kauder, C. P. Verrijzer, M. Ott and E. Verdin (2006). "The SWI/SNF chromatin-remodeling complex is a cofactor for Tat transactivation of the HIV promoter." J Biol Chem **281**(29): 19960-8.

Majid, S., A. A. Dar, S. Saini, S. Yamamura, H. Hirata, Y. Tanaka, G. Deng and R. Dahiya (2010). "MicroRNA-205-directed transcriptional activation of tumor suppressor genes in prostate cancer." Cancer.

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

Marciniak, R. A., B. J. Calnan, A. D. Frankel and P. A. Sharp (1990). "HIV-1 Tat protein trans-activates transcription in vitro." Cell **63**(4): 791-802.

Maroun, M., O. Delelis, G. Coadou, T. Bader, E. Segeral, G. Mbemba, C. Petit, P. Sonigo, J. C. Rain, J. F. Mouscadet, R. Benarous and S. Emiliani (2006). "Inhibition of early steps of HIV-1 replication by SNF5/Ini1." J Biol Chem **281**(32): 22736-43.

Marques, J. T. and B. R. Williams (2005). "Activation of the mammalian immune system by siRNAs." Nat Biotechnol **23**(11): 1399-405.

Marshall, H. M., K. Ronen, C. Berry, M. Llano, H. Sutherland, D. Saenz, W. Bickmore, E. Poeschla and F. D. Bushman (2007). "Role of PSIP1/LEDGF/p75 in lentiviral infectivity and integration targeting." PLoS ONE **2**(12): e1340.

Martianov, I., A. Ramadass, A. Serra Barros, N. Chow and A. Akoulitchev (2007). "Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript." Nature **445**(7128): 666-70.

Martinez, M. A., A. Gutierrez, M. Armand-Ugon, J. Blanco, M. Parera, J. Gomez, B. Clotet and J. A. Este (2002). "Suppression of chemokine receptor expression by RNA interference allows for inhibition of HIV-1 replication." AIDS **16**(18): 2385-90.

Mata, R. C., P. Viciana, A. de Alarcon, L. F. Lopez-Cortes, J. Gomez-Vera, M. Trastoy and J. M. Cisneros (2005). "Discontinuation of antiretroviral therapy in patients with chronic HIV infection: clinical, virologic, and immunologic consequences." AIDS Patient Care STDS **19**(9): 550-62.

Matilainen, J. M., T. Husso, S. Toropainen, S. Seuter, M. P. Turunen, P. Gynther, S. Yla-Herttuala, C. Carlberg and S. Vaisanen (2010). "Primary effect of 1alpha,25(OH)(2)D(3) on IL-10 expression in monocytes is short-term down-regulation." Biochim Biophys Acta **1803**(11): 1276-86.

Matranga, C., Y. Tomari, C. Shin, D. P. Bartel and P. D. Zamore (2005). "Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes." Cell **123**(4): 607-20.

Mattick, J. S. (2009). "The genetic signatures of noncoding RNAs." PLoS Genet **5**(4): e1000459.

Matzke, M. A. and J. A. Birchler (2005). "RNAi-mediated pathways in the nucleus." Nat Rev Genet **6**(1): 24-35.

McKee, C. J., J. J. Kessl, N. Shkriabai, M. J. Dar, A. Engelman and M. Kvaratskhelia (2008). "Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein." J Biol Chem **283**(46): 31802-12.

McKenna, E. S., C. G. Sansam, Y. J. Cho, H. Greulich, J. A. Evans, C. S. Thom, L. A. Moreau, J. A. Biegel, S. L. Pomeroy and C. W. Roberts (2008). "Loss of the epigenetic tumor suppressor SNF5 leads to cancer without genomic instability." Mol Cell Biol **28**(20): 6223-33.

McManus, M. T., C. P. Petersen, B. B. Haines, J. Chen and P. A. Sharp (2002). "Gene silencing using micro-RNA designed hairpins." RNA **8**(6): 842-50.

McManus, M. T. and P. A. Sharp (2002). "Gene silencing in mammals by small interfering RNAs." Nat Rev Genet **3**(10): 737-47.

McMichael, A., M. Mwau and T. Hanke (2002). "HIV T cell vaccines, the importance of clades." Vaccine **20**(15): 1918-21.

Meister, G., M. Landthaler, A. Patkaniowska, Y. Dorsett, G. Teng and T. Tuschl (2004). "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs." Mol Cell **15**(2): 185-97.

Mellors, J. W., C. R. Rinaldo, Jr., P. Gupta, R. M. White, J. A. Todd and L. A. Kingsley (1996). "Prognosis in HIV-1 infection predicted by the quantity of virus in plasma." Science **272**(5265): 1167-70.

Metivier, R., R. Gallais, C. Tiffoche, C. Le Peron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon and G. Salbert (2008). "Cyclical DNA methylation of a transcriptionally active promoter." Nature **452**(7183): 45-50.

Michel, F., C. Crucifix, F. Granger, S. Eiler, J. F. Mouscadet, S. Korolev, J. Agapkina, R. Ziganshin, M. Gottikh, A. Nazabal, S. Emiliani, R. Benarous, D. Moras, P. Schultz and M. Ruff (2009). "Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor." EMBO J **28**(7): 980-91.

Miller, H. B., K. O. Saunders, G. D. Tomaras and M. A. Garcia-Blanco (2009). "Tat-SF1 is not required for Tat transactivation but does regulate the relative levels of unspliced and spliced HIV-1 RNAs." PLoS One **4**(5): e5710.

Milne, L., Y. Xu, D. M. Perrin and D. S. Sigman (2000). "An approach to gene-specific transcription inhibition using oligonucleotides complementary to the template strand of the open complex." Proc Natl Acad Sci U S A **97**(7): 3136-41.

Miyoshi, H., U. Blomer, M. Takahashi, F. H. Gage and I. M. Verma (1998). "Development of a self-inactivating lentivirus vector." J Virol **72**(10): 8150-7.

Mizutani, T., A. Ishizaka, M. Tomizawa, T. Okazaki, N. Yamamichi, A. Kawana-Tachikawa, A. Iwamoto and H. Iba (2009). "Loss of the Brm-type SWI/SNF chromatin remodeling complex is a strong barrier to the Tat-independent transcriptional elongation of human immunodeficiency virus type 1 transcripts." J Virol **83**(22): 11569-80.

Moore, M. J. and N. J. Proudfoot (2009). "Pre-mRNA processing reaches back to transcription and ahead to translation." Cell **136**(4): 688-700.

Morlando, M., M. Ballarino, N. Gromak, F. Pagano, I. Bozzoni and N. J. Proudfoot (2008). "Primary microRNA transcripts are processed co-transcriptionally." Nat Struct Mol Biol **15**(9): 902-9.

Morozov, A., E. Yung and G. V. Kalpana (1998). "Structure-function analysis of integrase interactor 1/hSNF5L1 reveals differential properties of two repeat motifs present in the highly conserved region." Proc Natl Acad Sci U S A **95**(3): 1120-5.

Morris, K. V., S. W. Chan, S. E. Jacobsen and D. J. Looney (2004). "Small interfering RNA-induced transcriptional gene silencing in human cells." Science **305**(5688): 1289-92.

Morris, K. V., S. Santoso, A. M. Turner, C. Pastori and P. G. Hawkins (2008). "Bidirectional transcription directs both transcriptional gene activation and suppression in human cells." PLoS Genet **4**(11): e1000258.

Moulard, M. and E. Decroly (2000). "Maturation of HIV envelope glycoprotein precursors by cellular endoproteases." Biochim Biophys Acta **1469**(3): 121-32.

Mulhern, O. and A. G. Bowie (2010). "Unexpected roles for DEAD-box protein 3 in viral RNA sensing pathways." Eur J Immunol **40**(4): 933-5.

Murayama, A., K. Sakura, M. Nakama, K. Yasuzawa-Tanaka, E. Fujita, Y. Tateishi, Y. Wang, T. Ushijima, T. Baba, K. Shibuya, A. Shibuya, Y. Kawabe and J. Yanagisawa (2006). "A specific CpG site demethylation in the human interleukin 2 gene promoter is an epigenetic memory." EMBO J **25**(5): 1081-92.

Nair, M. P., J. L. Reynolds, S. D. Mahajan, S. A. Schwartz, R. Aalinkeel, B. Bindukumar and D. Sykes (2005). "RNAi-directed inhibition of DC-SIGN by dendritic cells: prospects for HIV-1 therapy." AAPS J **7**(3): E572-8.

Naito, T., Y. Kiyasu, K. Sugiyama, A. Kimura, R. Nakano, A. Matsukage and K. Nagata (2007). "An influenza virus replicon system in yeast identified Tat-SF1 as a stimulatory host factor for viral RNA synthesis." Proc Natl Acad Sci U S A **104**(46): 18235-40.

Naito, Y., T. Yamada, K. Ui-Tei, S. Morishita and K. Saigo (2004). "siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference." Nucleic Acids Res **32**(Web Server issue): W124-9.

Naito, Y., J. Yoshimura, S. Morishita and K. Ui-Tei (2009). "siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect." BMC Bioinformatics **10**: 392.

Nallagatla, S. R., J. Hwang, R. Toroney, X. Zheng, C. E. Cameron and P. C. Bevilacqua (2007). "5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops." Science **318**(5855): 1455-8.

Napoli, S., C. Pastori, M. Magistri, G. M. Carbone and C. V. Catapano (2009). "Promoter-specific transcriptional interference and c-myc gene silencing by siRNAs in human cells." EMBO J **28**(12): 1708-19.

Nathans, R., C. Y. Chu, A. K. Serquina, C. C. Lu, H. Cao and T. M. Rana (2009). "Cellular microRNA and P bodies modulate host-HIV-1 interactions." Mol Cell **34**(6): 696-709.

Neil, S. J., T. Zang and P. D. Bieniasz (2008). "Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu." Nature **451**(7177): 425-30.

Nguyen, D. G., K. C. Wolff, H. Yin, J. S. Caldwell and K. L. Kuhen (2006). ""UnPAKing" human immunodeficiency virus (HIV) replication: using small interfering RNA screening to identify novel cofactors and elucidate the role of group I PAKs in HIV infection." J Virol **80**(1): 130-7.

Nomaguchi, M. and A. Adachi (2009). "[Accessory proteins of HIV and innate anti-retroviral factors]." Uirusu **59**(1): 67-74.

Noto, T., H. M. Kurth, K. Kataoka, L. Aronica, L. V. DeSouza, K. W. Siu, R. E. Pearlman, M. A. Gorovsky and K. Mochizuki (2010). "The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus." Cell **140**(5): 692-703.

Novina, C. D., M. F. Murray, D. M. Dykxhoorn, P. J. Beresford, J. Riess, S. K. Lee, R. G. Collman, J. Lieberman, P. Shankar and P. A. Sharp (2002). "siRNA-directed inhibition of HIV-1 infection." Nat Med **8**(7): 681-6.

Oda, S., M. Schroder and A. R. Khan (2009). "Structural basis for targeting of human RNA helicase DDX3 by poxvirus protein K7." Structure **17**(11): 1528-37.

Okamura, K., J. W. Hagen, H. Duan, D. M. Tyler and E. C. Lai (2007). "The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila." Cell **130**(1): 89-100.

Okano, M., D. W. Bell, D. A. Haber and E. Li (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-57.

Oshiumi, H., M. Ikeda, M. Matsumoto, A. Watanabe, O. Takeuchi, S. Akira, N. Kato, K. Shimotohno and T. Seya (2010a). "Hepatitis C Virus Core Protein Abrogates the DDX3 Function That Enhances IPS-1-Mediated IFN-Beta Induction." PLoS One **5**(12): e14258.

Oshiumi, H., K. Sakai, M. Matsumoto and T. Seya (2010b). "DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-beta-inducing potential." Eur J Immunol **40**(4): 940-8.

Ouellet, D. L., I. Plante, P. Landry, C. Barat, M. E. Janelle, L. Flamand, M. J. Tremblay and P. Provost (2008). "Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element." Nucleic Acids Res **36**(7): 2353-65.

Overhoff, M., M. Alken, R. K. Far, M. Lemaitre, B. Lebleu, G. Sczakiel and I. Robbins (2005). "Local RNA target structure influences siRNA efficacy: a systematic global analysis." J Mol Biol **348**(4): 871-81.

Pache, L., R. Konig and S. K. Chanda (2010). "Identifying HIV-1 host cell factors by genome-scale RNAi screening." Methods.

Pachkov, M., I. Erb, N. Molina and E. van Nimwegen (2007). "SwissRegulon: a database of genome-wide annotations of regulatory sites." Nucleic Acids Res **35**(Database issue): D127-31.

Palanichamy, J. K., M. Mehndiratta, M. Bhagat, P. Ramalingam, B. Das, P. Das, S. Sinha and P. Chattopadhyay (2010). "Silencing of integrated human papillomavirus-16 oncogenes by small interfering RNA-mediated heterochromatization." Mol Cancer Ther **9**(7): 2114-22.

Pan, C. and R. Russell (2010). "Roles of DEAD-box proteins in RNA and RNP Folding." RNA Biol **7**(6): 28-37.

Parada, C. A. and R. G. Roeder (1996). "Enhanced processivity of RNA polymerase II triggered by Tat-induced phosphorylation of its carboxy-terminal domain." Nature **384**(6607): 375-8.

Parada, C. A. and R. G. Roeder (1999). "A novel RNA polymerase II-containing complex potentiates Tat-enhanced HIV-1 transcription." Embo J **18**(13): 3688-701.

Park, C. W., Y. Zeng and C. J. Steer (2010). "Human DNA methyltransferase 3a does not associate with microRNAs in the regulation of DNA methylation." J Cardiovasc Transl Res **3**(3): 290-5.

Parker, J. S., S. M. Roe and D. Barford (2005). "Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex." Nature **434**(7033): 663-6.

Passman, M., M. Weinberg, M. Kew and P. Arbuthnot (2000). "In situ demonstration of inhibitory effects of hammerhead ribozymes that are targeted to the hepatitis Bx sequence in cultured cells." Biochem Biophys Res Commun **268**(3): 728-33.

Pastor, F., D. Kolonias, P. H. Giangrande and E. Gilboa (2010). "Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay." Nature **465**(7295): 227-30.

Patarca, R., C. Heath, G. J. Goldenberg, C. A. Rosen, J. G. Sodroski, W. A. Haseltine and U. M. Hansen (1987). "Transcription directed by the HIV long terminal repeat in vitro." AIDS Res Hum Retroviruses **3**(1): 41-55.

Patzel, V., S. Rutz, I. Dietrich, C. Koberle, A. Scheffold and S. H. Kaufmann (2005). "Design of siRNAs producing unstructured guide-RNAs results in improved RNA interference efficiency." Nat Biotechnol **23**(11): 1440-4.

Paul, C. P., P. D. Good, I. Winer and D. R. Engelke (2002). "Effective expression of small interfering RNA in human cells." Nat Biotechnol **20**(5): 505-8.

Pawlicki, J. M. and J. A. Steitz (2008). "Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production." J Cell Biol **182**(1): 61-76.

Pebernard, S. and R. D. Iggo (2004). "Determinants of interferon-stimulated gene induction by RNAi vectors." Differentiation **72**(2-3): 103-11.

Perales, R. and D. Bentley (2009). ""Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions." Mol Cell **36**(2): 178-91.

Place, R. F., L. C. Li, D. Pookot, E. J. Noonan and R. Dahiya (2008). "MicroRNA-373 induces expression of genes with complementary promoter sequences." Proc Natl Acad Sci U S A **105**(5): 1608-13.

Plantier, J. C., M. Leoz, J. E. Dickerson, F. De Oliveira, F. Cordonnier, V. Lemee, F. Damond, D. L. Robertson and F. Simon (2009). "A new human immunodeficiency virus derived from gorillas." Nat Med **15**(8): 871-2.

Platt, E. J., K. Wehrly, S. E. Kuhmann, B. Chesebro and D. Kabat (1998). "Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1." J Virol **72**(4): 2855-64.

Poluri, A. and R. E. Sutton (2008). "Titers of HIV-based vectors encoding shRNAs are reduced by a dicer-dependent mechanism." Mol Ther **16**(2): 378-86.

Preker, P., J. Nielsen, S. Kammler, S. Lykke-Andersen, M. S. Christensen, C. K. Mapendano, M. H. Schierup and T. H. Jensen (2008). "RNA exosome depletion reveals transcription upstream of active human promoters." Science **322**(5909): 1851-4.

Pruitt, K., R. L. Zinn, J. E. Ohm, K. M. McGarvey, S. H. Kang, D. N. Watkins, J. G. Herman and S. B. Baylin (2006). "Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation." PLoS Genet **2**(3): e40.

Pulukuri, S. M. and J. S. Rao (2007). "Small interfering RNA directed reversal of urokinase plasminogen activator demethylation inhibits prostate tumor growth and metastasis." Cancer Res **67**(14): 6637-46.

Qin, X. F., D. S. An, I. S. Chen and D. Baltimore (2003). "Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5." Proc Natl Acad Sci U S A **100**(1): 183-8.

Raghavendra, N. K., N. Shkriabai, R. Graham, S. Hess, M. Kvaratskhelia and L. Wu (2010). "Identification of host proteins associated with HIV-1 preintegration complexes isolated from infected CD4+ cells." Retrovirology **7**: 66.

Rain, J. C., A. Cribier, A. Gerard, S. Emiliani and R. Benarous (2009). "Yeast two-hybrid detection of integrase-host factor interactions." Methods.

Rajkowitsch, L., D. Chen, S. Stampfl, K. Semrad, C. Waldsich, O. Mayer, M. F. Jantsch, R. Konrat, U. Blasi and R. Schroeder (2007). "RNA chaperones, RNA annealers and RNA helicases." RNA Biol **4**(3): 118-30.

Ramezani, A. and R. G. Hawley (2010). "Strategies to insulate lentiviral vector-expressed transgenes." Methods Mol Biol **614**: 77-100.

Rand, T. A., S. Petersen, F. Du and X. Wang (2005). "Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation." Cell **123**(4): 621-9.

Rato, S., S. Maia, P. M. Brito, L. Resende, C. F. Pereira, C. Moita, R. P. Freitas, J. Moniz-Pereira, N. Hacohen, L. F. Moita and J. Goncalves (2010). "Novel HIV-1 knockdown targets identified by an enriched kinases/phosphatases shRNA library using a long-term iterative screen in Jurkat T-cells." PLoS One **5**(2): e9276.

Ray, A., S. N. Mir, G. Wani, Q. Zhao, A. Battu, Q. Zhu, Q. E. Wang and A. A. Wani (2009). "Human SNF5/INI1, a component of the human SWI/SNF chromatin remodeling complex, promotes nucleotide excision repair by influencing ATM recruitment and downstream H2AX phosphorylation." Mol Cell Biol **29**(23): 6206-19.

Ren, G. L., Y. Fang, H. H. Ma, Y. F. Lei, D. Wang, M. C. Xu, P. Z. Wang, C. X. Huang, O. H. Nie, Y. T. Sun and X. F. Bai (2007). "The short hairpin RNA driven by polymerase II suppresses both wild-type and lamivudine-resistant hepatitis B virus strains." Antivir Ther **12**(6): 865-76.

Reynolds, A., E. M. Anderson, A. Vermeulen, Y. Fedorov, K. Robinson, D. Leake, J. Karpilow, W. S. Marshall and A. Khvorova (2006). "Induction of the interferon response by siRNA is cell type- and duplex length-dependent." RNA **12**(6): 988-93.

Reynolds, A., D. Leake, Q. Boese, S. Scaringe, W. S. Marshall and A. Khvorova (2004). "Rational siRNA design for RNA interference." Nat Biotechnol **22**(3): 326-30.

Richter, S., H. Cao and T. M. Rana (2002). "Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation." Biochemistry **41**(20): 6391-7.

Robbins, M. A., M. Li, I. Leung, H. Li, D. V. Boyer, Y. Song, M. A. Behlke and J. J. Rossi (2006). "Stable expression of shRNAs in human CD34+ progenitor cells can avoid induction of interferon responses to siRNAs in vitro." Nat Biotechnol **24**(5): 566-71.

Roebuck, K. A. and M. Saifuddin (1999). "Regulation of HIV-1 transcription." Gene Expr **8**(2): 67-84.

Rosenberg, Y. J., A. O. Anderson and R. Pabst (1998). "HIV-induced decline in blood CD4/CD8 ratios: viral killing or altered lymphocyte trafficking?" Immunol Today **19**(1): 10-7.

Rossi, J. J. (2008). "Expression strategies for short hairpin RNA interference triggers." Hum Gene Ther **19**(4): 313-7.

Ruby, J. G., C. H. Jan and D. P. Bartel (2007). "Intronic microRNA precursors that bypass Drosha processing." Nature **448**(7149): 83-6.

Sabin, C. A., S. W. Worm, R. Weber, P. Reiss, W. El-Sadr, F. Dabis, S. De Wit, M. Law, A. D'Arminio Monforte, N. Friis-Moller, O. Kirk, C. Pradier, I. Weller, A. N. Phillips and J. D. Lundgren (2008). "Use of nucleoside reverse transcriptase inhibitors and risk of myocardial infarction in HIV-infected patients enrolled in the D:A:D study: a multi-cohort collaboration." Lancet **371**(9622): 1417-26.

Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart and M. Parmentier (1996). "Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene." Nature **382**(6593): 722-5.

Sandelin, A., P. Carninci, B. Lenhard, J. Ponjavic, Y. Hayashizaki and D. A. Hume (2007). "Mammalian RNA polymerase II core promoters: insights from genome-wide studies." Nat Rev Genet **8**(6): 424-36.

Santoro, R., K. M. Schmitz, J. Sandoval and I. Grummt (2010). "Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans." EMBO Rep **11**(1): 52-8.

Sarafianos, S. G., B. Marchand, K. Das, D. M. Himmel, M. A. Parniak, S. H. Hughes and E. Arnold (2009). "Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition." J Mol Biol **385**(3): 693-713.

Sarma, K. and D. Reinberg (2005). "Histone variants meet their match." Nat Rev Mol Cell Biol **6**(2): 139-49.

Saumet, A. and C. H. Lecellier (2006). "Anti-viral RNA silencing: do we look like plants?" Retrovirology **3**: 3.

Schadt, E. E., S. W. Edwards, D. GuhaThakurta, D. Holder, L. Ying, V. Svetnik, A. Leonardson, K. W. Hart, A. Russell, G. Li, G. Cavet, J. Castle, P. McDonagh, Z. Kan, R. Chen, A. Kasarskis, M. Margarint, R. M. Caceres, J. M. Johnson, C. D. Armour, P. W. Garrett-Engele, N. F. Tsinoremas and D. D. Shoemaker (2004). "A comprehensive transcript index of the human genome generated using microarrays and computational approaches." Genome Biol **5**(10): R73.

Schoenfelder, S., T. Sexton, L. Chakalova, N. F. Cope, A. Horton, S. Andrews, S. Kurukuti, J. A. Mitchell, D. Umlauf, D. S. Dimitrova, C. H. Eskiw, Y. Luo, C. L. Wei, Y. Ruan, J. J. Bieker and P. Fraser (2010). "Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells." Nat Genet **42**(1): 53-61.

Schones, D. E., K. Cui, S. Cuddapah, T. Y. Roh, A. Barski, Z. Wang, G. Wei and K. Zhao (2008). "Dynamic regulation of nucleosome positioning in the human genome." Cell **132**(5): 887-98.

Schramm, L. and N. Hernandez (2002). "Recruitment of RNA polymerase III to its target promoters." Genes Dev **16**(20): 2593-620.

Schroder, M. (2010). "Human DEAD-box protein 3 has multiple functions in gene regulation and cell cycle control and is a prime target for viral manipulation." Biochem Pharmacol **79**(3): 297-306.

Schubert, S., A. Grunweller, V. A. Erdmann and J. Kurreck (2005). "Local RNA target structure influences siRNA efficacy: systematic analysis of intentionally designed binding regions." J Mol Biol **348**(4): 883-93.

Schwartz, J. C., S. T. Younger, N. B. Nguyen, D. B. Hardy, B. P. Monia, D. R. Corey and B. A. Janowski (2008). "Antisense transcripts are targets for activating small RNAs." Nat Struct Mol Biol **15**(8): 842-8.

Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo and G. N. Pavlakis (1990). "Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1." J Virol **64**(6): 2519-29.

Schwarz, D. S., G. Hutvagner, T. Du, Z. Xu, N. Aronin and P. D. Zamore (2003). "Asymmetry in the assembly of the RNAi enzyme complex." Cell **115**(2): 199-208.

Seckler, J. M., K. J. Howard, M. D. Barkley and P. L. Wintrode (2009). "Solution structural dynamics of HIV-1 reverse transcriptase heterodimer." Biochemistry **48**(32): 7646-55.

Seila, A. C., J. M. Calabrese, S. S. Levine, G. W. Yeo, P. B. Rahl, R. A. Flynn, R. A. Young and P. A. Sharp (2008). "Divergent transcription from active promoters." Science **322**(5909): 1849-51.

Shimizu, S., P. Hong, B. Arumugam, L. Pokomo, J. Boyer, N. Koizumi, P. Kittipongdaja, A. Chen, G. Bristol, Z. Galic, J. A. Zack, O. Yang, I. S. Chen, B. Lee and D. S. An (2010). "A highly efficient short hairpin RNA potently down-regulates CCR5 expression in systemic lymphoid organs in the hu-BLT mouse model." Blood **115**(8): 1534-44.

Shimizu, S., M. Kamata, P. Kittipongdaja, K. N. Chen, S. Kim, S. Pang, J. Boyer, F. X. Qin, D. S. An and I. S. Chen (2009). "Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5." Genet Vaccines Ther **7**: 8.

Shukla, S., C. S. Sumaria and P. I. Pradeepkumar (2010). "Exploring chemical modifications for siRNA therapeutics: a structural and functional outlook." ChemMedChem **5**(3): 328-49.

Shun, M. C., N. K. Raghavendra, N. Vandegraaff, J. E. Daigle, S. Hughes, P. Kellam, P. Cherepanov and A. Engelman (2007). "LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration." Genes Dev **21**(14): 1767-78.

Silverberg, M. J., W. Leyden, L. Hurley, A. S. Go, C. P. Quesenberry, Jr., D. Klein and M. A. Horberg (2009). "Response to newly prescribed lipid-lowering therapy in patients with and without HIV infection." Ann Intern Med **150**(5): 301-13.

Silvers, R. M., J. A. Smith, M. Schowalter, S. Litwin, Z. Liang, K. Geary and R. Daniel (2010). "Modification of integration site preferences of an HIV-1-based vector by expression of a novel synthetic protein." Hum Gene Ther **21**(3): 337-49.

Simon, V. and D. D. Ho (2003). "HIV-1 dynamics in vivo: implications for therapy." Nat Rev Microbiol **1**(3): 181-90.

Siomi, H. and M. C. Siomi (2009). "On the road to reading the RNA-interference code." Nature **457**(7228): 396-404.

Skalka, A. M. and R. A. Katz (2005). "Retroviral DNA integration and the DNA damage response." Cell Death Differ **12 Suppl 1**: 971-8.

Skupsky, R., J. C. Burnett, J. E. Foley, D. V. Schaffer and A. P. Arkin (2010). "HIV promoter integration site primarily modulates transcriptional burst size rather than frequency." PLoS Comput Biol **6**(9).

Smalheiser, N. R., G. Lugli, J. Thimmapuram, E. H. Cook and J. Larson (2011). "Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training." RNA **17**(1): 166-81.

Song, E., S. K. Lee, D. M. Dykxhoorn, C. Novina, D. Zhang, K. Crawford, J. Cerny, P. A. Sharp, J. Lieberman, N. Manjunath and P. Shankar (2003). "Sustained small interfering RNA-mediated human immunodeficiency virus type 1 inhibition in primary macrophages." J Virol **77**(13): 7174-81.

Song, E., P. Zhu, S. K. Lee, D. Chowdhury, S. Kussman, D. M. Dykxhoorn, Y. Feng, D. Palliser, D. B. Weiner, P. Shankar, W. A. Marasco and J. Lieberman (2005). "Antibody mediated in vivo delivery of small interfering RNAs via cell-surface receptors." Nat Biotechnol **23**(6): 709-17.

Sorin, M., J. Cano, S. Das, S. Mathew, X. Wu, K. P. Davies, X. Shi, S. W. Cheng, D. Ott and G. V. Kalpana (2009). "Recruitment of a SAP18-HDAC1 complex into HIV-1 virions and its requirement for viral replication." PLoS Pathog **5**(6): e1000463.

Sorin, M., E. Yung, X. Wu and G. V. Kalpana (2006). "HIV-1 replication in cell lines harboring INI1/hSNF5 mutations." Retrovirology **3**: 56.

Staal, F. J., M. Roederer and L. A. Herzenberg (1990). "Intracellular thiols regulate activation of nuclear factor kappa B and transcription of human immunodeficiency virus." Proc Natl Acad Sci U S A **87**(24): 9943-7.

Stark, G. R., I. M. Kerr, B. R. Williams, R. H. Silverman and R. D. Schreiber (1998). "How cells respond to interferons." Annu Rev Biochem **67**: 227-64.

Stein, C. A., J. B. Hansen, J. Lai, S. Wu, A. Voskresenskiy, A. Hog, J. Worm, M. Hedtjarn, N. Souleimanian, P. Miller, H. S. Soifer, D. Castanotto, L. Benimetskaya, H. Orum and T. Koch (2010). "Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents." Nucleic Acids Res **38**(1): e3.

Stojanova, A. and L. Z. Penn (2009). "The role of INI1/hSNF5 in gene regulation and cancer." Biochem Cell Biol **87**(1): 163-77.

Strahl, B. D. and C. D. Allis (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-5.

Strebel, K., J. Luban and K. T. Jeang (2009). "Human cellular restriction factors that target HIV-1 replication." BMC Med **7**: 48.

Streeck, H. and D. F. Nixon (2010). "T cell immunity in acute HIV-1 infection." J Infect Dis **202 Suppl 2**: S302-8.

Struhl, K. (1985). "Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast." Proc Natl Acad Sci U S A **82**(24): 8419-23.

Su, W. Y., H. Xiong and J. Y. Fang (2010). "Natural antisense transcripts regulate gene expression in an epigenetic manner." Biochem Biophys Res Commun **396**(2): 177-81.

Sun, C., C. T. Pager, G. Luo, P. Sarnow and J. H. Cate (2010). "Hepatitis C virus core-derived peptides inhibit genotype 1b viral genome replication via interaction with DDX3X." PLoS One **5**(9).

Sung, T. L. and A. P. Rice (2009). "miR-198 inhibits HIV-1 gene expression and replication in monocytes and its mechanism of action appears to involve repression of cyclin T1." PLoS Pathog **5**(1): e1000263.

Surabhi, R. M. and R. B. Gaynor (2002). "RNA interference directed against viral and cellular targets inhibits human immunodeficiency Virus Type 1 replication." J Virol **76**(24): 12963-73.

Sutherland, H. G., K. Newton, D. G. Brownstein, M. C. Holmes, C. Kress, C. A. Semple and W. A. Bickmore (2006). "Disruption of Ledgf/Psip1 results in perinatal mortality and homeotic skeletal transformations." Mol Cell Biol **26**(19): 7201-10.

Suzuki, K., T. Juelich, H. Lim, T. Ishida, T. Watanebe, D. A. Cooper, S. Rao and A. D. Kelleher (2008). "Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region." J Biol Chem **283**(34): 23353-63.

Suzuki, M. M. and A. Bird (2008). "DNA methylation landscapes: provocative insights from epigenomics." Nat Rev Genet **9**(6): 465-76.

Swan, C. H., B. Buhler, P. Steinberger, M. P. Tschan, C. F. Barbas, 3rd and B. E. Torbett (2006). "T-cell protection and enrichment through lentiviral CCR5 intrabody gene delivery." Gene Ther **13**(20): 1480-92.

Taft, R. J., E. A. Glazov, N. Cloonan, C. Simons, S. Stephen, G. J. Faulkner, T. Lassmann, A. R. Forrest, S. M. Grimmond, K. Schroder, K. Irvine, T. Arakawa, M. Nakamura, A. Kubosaki, K. Hayashida, C. Kawazu, M. Murata, H. Nishiyori, S. Fukuda, J. Kawai, C. O. Daub, D. A. Hume, H. Suzuki, V. Orlando, P. Carninci, Y. Hayashizaki and J. S. Mattick (2009). "Tiny RNAs associated with transcription start sites in animals." Nat Genet **41**(5): 572-8.

Taft, R. J., C. Simons, S. Nahkuri, H. Oey, D. J. Korbie, T. R. Mercer, J. Holst, W. Ritchie, J. J. Wong, J. E. Rasko, D. S. Rokhsar, B. M. Degnan and J. S. Mattick (2010). "Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans." Nat Struct Mol Biol **17**(8): 1030-4.

Takeda, K. and S. Akira (2005). "Toll-like receptors in innate immunity." Int Immunol **17**(1): 1-14.

Tam, O. H., A. A. Aravin, P. Stein, A. Girard, E. P. Murchison, S. Cheloufi, E. Hodges, M. Anger, R. Sachidanandam, R. M. Schultz and G. J. Hannon (2008). "Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes." Nature **453**(7194): 534-8.

Tan, G. S., B. G. Garchow, X. Liu, J. Yeung, J. P. t. Morris, T. L. Cuellar, M. T. McManus and M. Kiriakidou (2009a). "Expanded RNA-binding activities of mammalian Argonaute 2." Nucleic Acids Res **37**(22): 7533-45.

Tan, Y., B. Zhang, T. Wu, G. Skogerbo, X. Zhu, X. Guo, S. He and R. Chen (2009b). "Transcriptional inhibiton of Hoxd4 expression by miRNA-10a in human breast cancer cells." BMC Mol Biol **10**: 12.

Taylor, B. S., M. E. Sobieszczyk, F. E. McCutchan and S. M. Hammer (2008). "The challenge of HIV-1 subtype diversity." N Engl J Med **358**(15): 1590-602.

Ter Brake, O. and B. Berkhout (2005). "A novel approach for inhibition of HIV-1 by RNA interference: counteracting viral escape with a second generation of siRNAs." J RNAi Gene Silencing **1**(2): 56-65.

ter Brake, O., P. Konstantinova, M. Ceylan and B. Berkhout (2006). "Silencing of HIV-1 with RNA interference: a multiple shRNA approach." Mol Ther **14**(6): 883-92.

Tiemann, K. and J. J. Rossi (2009). "RNAi-based therapeutics-current status, challenges and prospects." EMBO Mol Med **1**(3): 142-51.

Ting, A. H., K. E. Schuebel, J. G. Herman and S. B. Baylin (2005). "Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation." Nat Genet **37**(8): 906-10.

Tomari, Y., T. Du and P. D. Zamore (2007). "Sorting of Drosophila small silencing RNAs." Cell **130**(2): 299-308.

Treand, C., I. du Chene, V. Bres, R. Kiernan, R. Benarous, M. Benkirane and S. Emiliani (2006). "Requirement for SWI/SNF chromatin-remodeling complex in Tat-mediated activation of the HIV-1 promoter." Embo J **25**(8): 1690-9.

Triboulet, R., B. Mari, Y. L. Lin, C. Chable-Bessia, Y. Bennasser, K. Lebrigand, B. Cardinaud, T. Maurin, P. Barbry, V. Baillat, J. Reynes, P. Corbeau, K. T. Jeang and M. Benkirane (2007). "Suppression of microRNA-silencing pathway by HIV-1 during virus replication." Science **315**(5818): 1579-82.

Trinklein, N. D., S. F. Aldred, S. J. Hartman, D. I. Schroeder, R. P. Otillar and R. M. Myers (2004). "An abundance of bidirectional promoters in the human genome." Genome Res **14**(1): 62-6.

Turelli, P., V. Doucas, E. Craig, B. Mangeat, N. Klages, R. Evans, G. Kalpana and D. Trono (2001). "Cytoplasmic recruitment of INI1 and PML on incoming HIV preintegration complexes: interference with early steps of viral replication." Mol Cell **7**(6): 1245-54.

Turner, A. M., J. De La Cruz and K. V. Morris (2009). "Mobilization-competent Lentiviral Vector-mediated Sustained Transcriptional Modulation of HIV-1 Expression." Mol Ther **17**(2): 360-8.

Turunen, M. P., T. Lehtola, S. E. Heinonen, G. S. Assefa, P. Korpisalo, R. Girnary, C. K. Glass, S. Vaisanen and S. Yla-Herttuala (2009). "Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenetherapy." Circ Res **105**(6): 604-9.

Tuschl, T. and A. Borkhardt (2002). "Small interfering RNAs: a revolutionary tool for the analysis of gene function and gene therapy." Mol Interv **2**(3): 158-67.

Tyagi, M., R. J. Pearson and J. Karn (2010). "Establishment of HIV latency in primary CD4+ cells is due to epigenetic transcriptional silencing and P-TEFb restriction." J Virol **84**(13): 6425-37.

Ui-Tei, K., Y. Naito, K. Nishi, A. Juni and K. Saigo (2008). "Thermodynamic stability and Watson-Crick base pairing in the seed duplex are major determinants of the efficiency of the siRNA-based off-target effect." Nucleic Acids Res **36**(22): 7100-9.

Ui-Tei, K., Y. Naito, F. Takahashi, T. Haraguchi, H. Ohki-Hamazaki, A. Juni, R. Ueda and K. Saigo (2004). "Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference." Nucleic Acids Res **32**(3): 936-48.

Unwalla, H. and J. J. Rossi (2008). "Tat-regulated expression of RNA interference: triggers for the treatment of HIV infection." Curr HIV/AIDS Rep **5**(1): 40-3.

van Bakel, H., C. Nislow, B. J. Blencowe and T. R. Hughes (2010). "Most "dark matter" transcripts are associated with known genes." PLoS Biol **8**(5): e1000371.

Van Heuverswyn, F. and M. Peeters (2007). "The origins of HIV and implications for the global epidemic." Curr Infect Dis Rep **9**(4): 338-46.

van Nimwegen, E. (2007). "Finding regulatory elements and regulatory motifs: a general probabilistic framework." BMC Bioinformatics **8 Suppl 6**: S4.

Vandamme, A. M., K. Van Vaerenbergh and E. De Clercq (1998). "Anti-human immunodeficiency virus drug combination strategies." Antivir Chem Chemother **9**(3): 187-203.

Vandekerckhove, L., F. Christ, B. Van Maele, J. De Rijck, R. Gijsbers, C. Van den Haute, M. Witvrouw and Z. Debyser (2006). "Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus." J Virol **80**(4): 1886-96.

VandenDriessche, T., L. Thorrez, L. Naldini, A. Follenzi, L. Moons, Z. Berneman, D. Collen and M. K. Chuah (2002). "Lentiviral vectors containing the human immunodeficiency virus type-1 central polypurine tract can efficiently transduce nondividing hepatocytes and antigen-presenting cells in vivo." Blood **100**(3): 813-22.

Verani, A., G. Gras and G. Pancino (2005). "Macrophages and HIV-1: dangerous liaisons." Mol Immunol **42**(2): 195-212.

Verdel, A., S. Jia, S. Gerber, T. Sugiyama, S. Gygi, S. I. Grewal and D. Moazed (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." Science **303**(5658): 672-6.

Versteege, I., N. Sevenet, J. Lange, M. F. Rousseau-Merck, P. Ambros, R. Handgretinger, A. Aurias and O. Delattre (1998). "Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer." Nature **394**(6689): 203-6.

Vries, R. G., V. Bezrookove, L. M. Zuijderduijn, S. K. Kia, A. Houweling, I. Oruetxebarria, A. K. Raap and C. P. Verrijzer (2005). "Cancer-associated mutations in chromatin remodeler hSNF5 promote chromosomal instability by compromising the mitotic checkpoint." Genes Dev **19**(6): 665-70.

Wain, L. V., E. Bailes, F. Bibollet-Ruche, J. M. Decker, B. F. Keele, F. Van Heuverswyn, Y. Li, J. Takehisa, E. M. Ngole, G. M. Shaw, M. Peeters, B. H. Hahn and P. M. Sharp (2007). "Adaptation of HIV-1 to its human host." Mol Biol Evol **24**(8): 1853-60.

Walensky, R. P., A. D. Paltiel, E. Losina, L. M. Mercincavage, B. R. Schackman, P. E. Sax, M. C. Weinstein and K. A. Freedberg (2006). "The survival benefits of AIDS treatment in the United States." J Infect Dis **194**(1): 11-9.

Walter, B. L., K. Wehrly, R. Swanstrom, E. Platt, D. Kabat and B. Chesebro (2005). "Role of low CD4 levels in the influence of human immunodeficiency virus type 1 envelope V1 and V2 regions on entry and spread in macrophages." J Virol **79**(8): 4828-37.

Wang, B., S. Li, H. H. Qi, D. Chowdhury, Y. Shi and C. D. Novina (2009a). "Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins." Nat Struct Mol Biol **16**(12): 1259-66.

Wang, H. and W. S. Ryu (2010). "Hepatitis B virus polymerase blocks pattern recognition receptor signaling via interaction with DDX3: implications for immune evasion." PLoS Pathog **6**(7): e1000986.

Wang, W., J. Cote, Y. Xue, S. Zhou, P. A. Khavari, S. R. Biggar, C. Muchardt, G. V. Kalpana, S. P. Goff, M. Yaniv, J. L. Workman and G. R. Crabtree (1996). "Purification and biochemical heterogeneity of the mammalian SWI-SNF complex." Embo J **15**(19): 5370-82.

Wang, X., Y. Feng, L. Pan, Y. Wang, X. Xu, J. Lu and B. Huang (2007). "The proximal GC-rich region of p16(INK4a) gene promoter plays a role in its transcriptional regulation." Mol Cell Biochem **301**(1-2): 259-66.

Wang, X., L. Ye, W. Hou, Y. Zhou, Y. J. Wang, D. S. Metzger and W. Z. Ho (2009b). "Cellular microRNA expression correlates with susceptibility of monocytes/macrophages to HIV-1 infection." Blood **113**(3): 671-4.

Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, W. Peng, M. Q. Zhang and K. Zhao (2008). "Combinatorial patterns of histone acetylations and methylations in the human genome." Nat Genet **40**(7): 897-903.

Watanabe, T., Y. Totoki, A. Toyoda, M. Kaneda, S. Kuramochi-Miyagawa, Y. Obata, H. Chiba, Y. Kohara, T. Kono, T. Nakano, M. A. Surani, Y. Sakaki and H. Sasaki (2008). "Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes." Nature **453**(7194): 539-43.

Watts, J. K., D. Yu, K. Charisse, C. Montaillier, P. Potier, M. Manoharan and D. R. Corey (2010). "Effect of chemical modifications on modulation of gene expression by duplex antigene RNAs that are complementary to non-coding transcripts at gene promoters." Nucleic Acids Res **38**(15): 5242-59.

Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Paabo, M. Rebhan and D. Schubeler (2007). "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." Nat Genet **39**(4): 457-66.

Wei, J. X., J. Yang, J. F. Sun, L. T. Jia, Y. Zhang, H. Z. Zhang, X. Li, Y. L. Meng, L. B. Yao and A. G. Yang (2009). "Both strands of siRNA have potential to guide posttranscriptional gene silencing in mammalian cells." PLoS One **4**(4): e5382.

Wei, X., J. M. Decker, H. Liu, Z. Zhang, R. B. Arani, J. M. Kilby, M. S. Saag, X. Wu, G. M. Shaw and J. C. Kappes (2002). "Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy." Antimicrob Agents Chemother **46**(6): 1896-905.

Weinberg, M. S., S. Barichievy, L. Schaffer, J. Han and K. V. Morris (2007). "An RNA targeted to the HIV-1 LTR promoter modulates indiscriminate off-target gene activation." Nucleic Acids Res **35**(21): 7303-12.

Weinberg, M. S., L. M. Villeneuve, A. Ehsani, M. Amarzguioui, L. Aagaard, Z. X. Chen, A. D. Riggs, J. J. Rossi and K. V. Morris (2006). "The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells." Rna **12**(2): 256-62.

Weinmann, L., J. Hock, T. Ivacevic, T. Ohrt, J. Mutze, P. Schwille, E. Kremmer, V. Benes, H. Urlaub and G. Meister (2009). "Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs." Cell **136**(3): 496-507.

Westerhout, E. M., O. ter Brake and B. Berkhout (2006). "The virion-associated incoming HIV-1 RNA genome is not targeted by RNA interference." Retrovirology **3**: 57.

Whittle, H., J. Morris, J. Todd, T. Corrah, S. Sabally, J. Bangali, P. T. Ngom, M. Rolfe and A. Wilkins (1994). "HIV-2-infected patients survive longer than HIV-1-infected patients." AIDS **8**(11): 1617-20.

Wiegers, K., G. Rutter, H. Kottler, U. Tessmer, H. Hohenberg and H. G. Krausslich (1998). "Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites." J Virol **72**(4): 2846-54.

Wightman, B., I. Ha and G. Ruvkun (1993). "Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans." Cell **75**(5): 855-62.

Williams, P., J. Wu, S. Cohn, S. Koletar, J. McCutchan, R. Murphy and J. Currier (2009). "Improvement in lipid profiles over 6 years of follow-up in adults with AIDS and immune reconstitution." HIV Med **10**(5): 290-301.

Williams, S. A., H. Kwon, L. F. Chen and W. C. Greene (2007). "Sustained induction of NF-kappa B is required for efficient expression of latent human immunodeficiency virus type 1." J Virol **81**(11): 6043-56.

Wiznerowicz, M. and D. Trono (2003). "Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference." J Virol **77**(16): 8957-61.

Woolcock, K. J., D. Gaidatzis, T. Punga and M. Buhler (2011). "Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe." Nat Struct Mol Biol **18**(1): 94-9.

Wu, D. Y., G. V. Kalpana, S. P. Goff and W. H. Schubach (1996). "Epstein-Barr virus nuclear protein 2 (EBNA2) binds to a component of the human SNF-SWI complex, hSNF5/Ini1." J Virol **70**(9): 6020-8.

Wu, X., H. Liu, H. Xiao, J. A. Conway, E. Hehl, G. V. Kalpana, V. Prasad and J. C. Kappes (1999). "Human immunodeficiency virus type 1 integrase protein promotes reverse transcription through specific interactions with the nucleoprotein reverse transcription complex." J Virol **73**(3): 2126-35.

Yam, P. Y., S. Li, J. Wu, J. Hu, J. A. Zaia and J. K. Yee (2002). "Design of HIV vectors for efficient gene delivery into human hematopoietic cells." Mol Ther **5**(4): 479-84.

Yamagishi, M., T. Ishida, A. Miyake, D. A. Cooper, A. D. Kelleher, K. Suzuki and T. Watanabe (2009). "Retroviral delivery of promoter-targeted shRNA induces long-term silencing of HIV-1 transcription." Microbes Infect **11**(4): 500-8.

Yan, D., R. Perriman, H. Igel, K. J. Howe, M. Neville and M. Ares, Jr. (1998). "CUS2, a yeast homolog of human Tat-SF1, rescues function of misfolded U2 through an unusual RNA recognition motif." Mol Cell Biol **18**(9): 5000-9.

Yang, J. S. and E. C. Lai (2010). "Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates." Cell Cycle **9**(22): 4455-60.

Yang, J. S., T. Maurin, N. Robine, K. D. Rasmussen, K. L. Jeffrey, R. Chandwani, E. P. Papapetrou, M. Sadelain, D. O'Carroll and E. C. Lai (2010). "Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis." Proc Natl Acad Sci U S A **107**(34): 15163-8.

Yedavalli, V. S., C. Neuveut, Y. H. Chi, L. Kleiman and K. T. Jeang (2004). "Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function." Cell **119**(3): 381-92.

Yeung, M. L., Y. Bennasser, T. G. Myers, G. Jiang, M. Benkirane and K. T. Jeang (2005). "Changes in microRNA expression profiles in HIV-1-transfected human cells." Retrovirology **2**: 81.

Yeung, M. L., Y. Bennasser, K. Watashi, S. Y. Le, L. Houzet and K. T. Jeang (2009a). "Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid." Nucleic Acids Res **37**(19): 6575-86.

Yeung, M. L., L. Houzet, V. S. Yedavalli and K. T. Jeang (2009b). "A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication." J Biol Chem **284**(29): 19463-73.

Yi, R., Y. Qin, I. G. Macara and B. R. Cullen (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes Dev **17**(24): 3011-6.

Yoshinaka, Y. and R. B. Luftig (1977). "Murine leukemia virus morphogenesis: cleavage of P70 in vitro can be accompanied by a shift from a concentrically coiled internal strand ("immature") to a collapsed ("mature") form of the virus core." Proc Natl Acad Sci U S A **74**(8): 3446-50.

Yu, J. Y., S. L. DeRuiter and D. L. Turner (2002). "RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells." Proc Natl Acad Sci U S A **99**(9): 6047-52.

Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A. P. Feinberg and H. Cui (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA." Nature **451**(7175): 202-6.

Yue, X., J. C. Schwartz, Y. Chu, S. T. Younger, K. T. Gagnon, S. Elbashir, B. A. Janowski and D. R. Corey (2010). "Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini." Nat Chem Biol **6**(8): 621-9.

Yung, E., M. Sorin, A. Pal, E. Craig, A. Morozov, O. Delattre, J. Kappes, D. Ott and G. V. Kalpana (2001). "Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1." Nat Med **7**(8): 920-6.

Yung, E., M. Sorin, E. J. Wang, S. Perumal, D. Ott and G. V. Kalpana (2004). "Specificity of interaction of INI1/hSNF5 with retroviral integrases and its functional significance." J Virol **78**(5): 2222-31.

Zaitseva, L., P. Cherepanov, L. Leyens, S. J. Wilson, J. Rasaiyaah and A. Fassati (2009). "HIV-1 exploits importin 7 to maximize nuclear import of its DNA genome." Retrovirology **6**: 11.

Zaitseva, L., R. Myers and A. Fassati (2006). "tRNAs promote nuclear import of HIV-1 intracellular reverse transcription complexes." PLoS Biol **4**(10): e332.

Zeng, Y. and B. R. Cullen (2004). "Structural requirements for pre-microRNA binding and nuclear export by Exportin 5." Nucleic Acids Res **32**(16): 4776-85.

Zeng, Y. and B. R. Cullen (2005). "Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences." J Biol Chem **280**(30): 27595-603.

Zeng, Y., E. J. Wagner and B. R. Cullen (2002). "Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells." Mol Cell **9**(6): 1327-33.

Zhang, J. and J. J. Rossi (2010). "Strategies in designing multigene expression units to downregulate HIV-1." Methods Mol Biol **623**: 123-36.

Zhao, F., Z. Xuan, L. Liu and M. Q. Zhang (2005). "TRED: a Transcriptional Regulatory Element Database and a platform for in silico gene regulation studies." Nucleic Acids Res **33**(Database issue): D103-7.

Zhou, H., M. Xu, Q. Huang, A. T. Gates, X. D. Zhang, J. C. Castle, E. Stec, M. Ferrer, B. Strulovici, D. J. Hazuda and A. S. Espeseth (2008). "Genome-scale RNAi screen for host factors required for HIV replication." Cell Host Microbe **4**(5): 495-504.

Zhou, J., P. Swiderski, H. Li, J. Zhang, C. P. Neff, R. Akkina and J. J. Rossi (2009). "Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells." Nucleic Acids Res **37**(9): 3094-109.

Zhou, M., L. Deng, V. Lacoste, H. U. Park, A. Pumfery, F. Kashanchi, J. N. Brady and A. Kumar (2004a). "Coordination of transcription factor phosphorylation and histone methylation by

the P-TEFb kinase during human immunodeficiency virus type 1 transcription." J Virol **78**(24): 13522-33.

Zhou, N., J. Fang, M. Mukhtar, E. Acheampong and R. J. Pomerantz (2004b). "Inhibition of HIV-1 fusion with small interfering RNAs targeting the chemokine coreceptor CXCR4." Gene Ther **11**(23): 1703-12.

Zhou, Q., D. Chen, E. Pierstorff and K. Luo (1998). "Transcription elongation factor P-TEFb mediates Tat activation of HIV-1 transcription at multiple stages." EMBO J **17**(13): 3681-91.

Zhou, Q. and P. A. Sharp (1996). "Tat-SF1: cofactor for stimulation of transcriptional elongation by HIV-1 Tat." Science **274**(5287): 605-10.

Zhu, K., C. Dobard and S. A. Chow (2004). "Requirement for integrase during reverse transcription of human immunodeficiency virus type 1 and the effect of cysteine mutations of integrase on its interactions with reverse transcriptase." J Virol **78**(10): 5045-55.

APPENDICES

A1 Standard laboratory practices

A1.1 Bacterial procedures

A1.1.1 Bacterial growth

All of the work used the bacterial strain *Escherichia coli* DH5α (Hanahan 1983). Cultures were grown in LB broth (10% tryptone, 5% yeast extract and 10% NaCl (Merck, Darmstadt, Germany)) at 37 °C with shaking at 350 rpm in an orbital shaker (New Scientific Co, Inc., NJ, USA). Colonies were grown on LB agar plates (LB with 10% agar (Merck, Darmstadt, Germany)) supplemented with 200 μg/ml ampicillin (Sigma-Aldrich, MO, USA).

TA cloning required blue/white screening of colonies. In this case, ampicillin LB agar plates were coated with 800 μg of both 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal) (Sigma, MO, USA) and isopropyl-β-D-1-thiogalactopyranoside (IPTG) (Roche, Mannheim, Germany) and dried prior to plating of transformed bacteria.

A.1.1.2 Preparation of transformation-competent DH5α

Chemically competent cells for use in transformations were prepared by growing DH5α in 100 ml LB until log phase, or OD_{600} 0.4-0.6. Cells were centrifuged for 12 min at 59.5 x *g* and the cell pellets resuspended in transformation buffer (60 mM $CaCl_2$, 15% glycerol and 10 mM PIPES, pH 7.0) and incubated on ice for 30 min. Cells were centrifuged at 4 °C for 10 min at 59.5 x *g* and resuspended in 2 ml transformation buffer. Aliquots of 100 μl of competent cells were stored at -80 °C.

A1.1.3 Transformation of DH5α

DNA was added to 100 μl of competent cells in a volume that did not exceed 10 μl. Cell-DNA mixtures were incubated on ice for 30 min, heat-shocked at 42 °C for 90 sec and then returned to ice for 5 min. Pre-warmed LB (1 ml) was added to the transformed cells which were then incubated at 37 °C for 1 hour to allow expression of antibiotic resistance genes. Cells were then plated onto LB agar plates supplemented with ampicillin, as described in A1.1.1.

A1.1.4 Storage of transformed DH5α

Samples of transformed DH5α were stored at -80 °C by adding 1 ml of log phase cultures to 0.3 ml of 100% glycerol. New cultures were inoculated by adding scrapings of the frozen stock to LB supplemented with ampicillin, as described in A1.1.1.

A1.2 DNA procedures

A1.2.1 Small-scale plasmid DNA preparation

Plasmid DNA was isolated from 4 ml bacterial cultures using the High Pure Plasmid Isolation Kit (Roche, Mannheim, Germany). Cultures were centrifuged at 3,000 x *g* for 10 min and cell pellets resuspended in 250 μl suspension buffer (50 mM Tris.HCl, 10 mM EDTA, pH 8.5 and 2.5 mg RNase A). To suspensions, 250 μl lysis buffer (0.2 M NaOH and 1% SDS) was added and samples were inverted several times to mix prior to incubation at room temperature for 5 min. Three hundred and fifty microliters binding buffer (4 M guanidine hydrochloride, 0.5 M potassium acetate, pH 4.2) was added and samples incubated on ice for 5 min prior to centrifugation at 16,000 x *g* for 1 min. Supernatant was transferred to a High Pure filter tube that was centrifuged for 1 min at 16,000 x *g*. Flow-through was discarded and 700 μl wash buffer (20 mM NaCl, 2 mM Tris.HCl, pH 7.5) added and samples were again centrifuged at 16,000 x *g* for 1 min. After discarding the flow-through once more, another centrifugation step was performed to remove residual wash buffer from the filter. Filters were then transferred to a clean tube and 100 μl of elution buffer (10 mM Tris.HCl, pH 8.5) applied. Samples were allowed to stand at room temperature for 2 min prior to a final centrifugation at 16,000 x *g* to elute the filter-bound plasmid DNA.

The concentration of all DNA preparations was determined by spectrophotometry on a NanoDrop (Themo Fisher Scientific, Inc., MA, USA) where 1 absorbance unit at a wavelength of 260 nm (A_{260}) equates to 50 μg/ml dsDNA. Purity of samples was also determined by spectrophotometry ie proximity of the A_{260}/A_{280} ratio to 1.8.

A1.2.2 Bulk preparation of plasmid DNA

Bulk plasmid DNA preparation from 300 ml LB bacterial cultures used the QIAGEN Plasmid Maxi Kit (QIAGEN, CA, USA). Cells were centrifuged at 3,000 x *g* for 20 min and bacterial pellets resuspended in 10 ml buffer P1 (50 mM Tris.HCl, pH 8.0, 10 mM EDTA and 100 μg/ml RNase A). Ten millilitres buffer P2 (200 mM NAOH and 1% SDS) was then added, samples thoroughly mixed by inversion and incubated at 42 °C for 15 min. Finally, 10 ml of buffer P3 (3 M potassium acetate, pH 5.5) was added to the lysate and samples were centrifuged at 3,000 x *g* for 30 min. Supernatant was transferred to a QIAGEN-tip 500 column, previously equilibrated with 10 ml QBT buffer (750 mM NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol (v/v), 0.15% Triton® X-100 (v/v)). After the lysate had passed through the column, the column was washed twice with 30 ml QC buffer (1 M NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol (v/v)). Bound DNA was eluted into 10 ml isopropanol with addition of 15 ml QF buffer (1.25 M NaCl, 50 mM Tris.HCl, pH 8.5, 15% isopropanol (v/v)) and then stored at -20 °C overnight. The following day, DNA precipitate was collected by centrifugation at 8,000 x *g* at 4 °C for 60 min. DNA pellets were air-dried for 5 min before resuspension in 250 μl nuclease-free water.

A1.2.3 DNA gel electrophoresis

Agarose gels were prepared by heating the appropriate amounts of agarose (Sigma-Aldrich, MO, USA) and TAE buffer (40 mM Tris acetate and 1 mM EDTA) until the agarose was dissolved. Prior to pouring, 1 μl of ethidium bromide was added per 50 ml of agarose-TAE solution. DNA samples were mixed with 6x orange loading dye (Fermentas, WI, USA) before loading on the gel. Five microlitres of O'GeneRuler™ DNA ladder (Fermentas, WI, USA) was included in each gel. Electrophoresis was carried out at ~100 volts.

A1.2.4 DNA extraction from agarose gels

The MiniElute™ Gel Extraction Kit (QIAGEN, CA, USA) was used to extract DNA fragments from agarose gels. Desired fragments were excised using a scalpel, weighed and a volume of QG buffer (proprietary information) equal to 3x the weight of the gel fragment added to the sample. Gel slices were dissolved by incubation at 50 °C for 10 min. One gel volume of isopropanol was added and the mixture then transferred to a QIAquick spin column that was centrifuged for 90 sec at 16,000 x *g*. Flow-through was discarded and 500 μl QG buffer added to the column. Columns were again centrifuged and flow-through discarded as before. Filters were washed with 750 μl PE buffer (proprietary information) and then an additional centrifugation step was performed to dry

the column. DNA was eluted into a clean tube on application of 50 μl EB buffer (10 mM Tris.HCl, pH 8.5) to the filter and centrifugation at 16,000 x *g* for 2 min.

A1.2.5 Calcium chloride method for DNA transfection

DNA (<20 μg) was mixed with 62.5 μl of buffer 1 (2 M $CaCl_2$) and then made up to 500 μl with distilled water. Gradually, the DNA-$CaCl_2$ mix was added to 500 μl buffer 2 (280 nM NaCl, 50 mM HEPES and 1.5 mM Na_2HPO_4, pH 7.05) and incubated at room temperature for 20 min. The DNA mix was added drop-wise to cells and left overnight at 37 °C, 5% CO_2, pH 5.

A1.3 RNA procedures

A1.3.1 RNA extraction from mammalian cells

TriReagent™ (Sigma-Aldrich, MO, USA) was added to PBS-washed cells that were then left to stand at room temperature for 5 min. A volume of TriReagent™ was used sufficient to just cover adherent cells ie ~1 ml per 15 cm^2, or to lyse suspension cells ie ~1 ml per 1 x 10^7 cells. 200 μl chloroform per ml TriReagent™ was added to samples that were then vortexed for 10 sec. Samples stood at room temperature for 5 min prior to centrifugation at 16,000 x *g* for 10 min at 4 °C. The RNA-containing aqueous layer was transferred to 700 μl isopropanol per ml TriReagent™ and samples stored overnight at -80 °C to precipitate RNA. The following day, samples were centrifuged at 16,000 x *g* for 12 min at 4 °C. The RNA pellets were washed with 75% ethanol and centrifuged again at 16,000 x *g* for 8 min at 4 °C. RNA was air-dried for 5 min before resuspension in DEPC-treated nuclease-free water (Ambion, TX, USA). Typically, RNA isolated from 24-well plate cultures were resuspended in 30 μl.

A1.3.2 DNase treatment and reverse transcription

RNA isolated from HEK293 or TZM-bl cells was prepared for qPCR by DNase treatment and reverse transcription using the QuantiTect Reverse Transcription Kit (QIAGEN, CA, USA). Genomic DNA elimination reactions contained ~100 ng total RNA, 1 μl gDNA wipeout buffer (proprietary information) and nuclease-free water to a total volume of 7 μl, which were incubated at 42 °C for 2 min before transfer to ice. To each sample, 0.5 μl Quantiscript reverse transcriptase, 2 μl of Quantiscript RT buffer (proprietary information) and 0.5 μl of RT primer mix was added,

followed by incubation at 42 °C for 15 min. Subsequent incubation at 95 °C for 3 min was performed to inactivate the RT enzyme. The resultant cDNA was stored at -20 °C if not used immediately in qPCR.

A1.3.3 Preparation of radiolabelled Decade™ Marker

Radiolabelled Decade™ Marker (Ambion, TX, USA) was prepared by mixing 100 ng Decade™ Marker RNA, 20 pmol γ-^{32}P-ATP (~6,000 Ci/mmol), 10 U T4 polynucleotide kinase and kinase reaction buffer (proprietary information) in a volume of 10 μl and incubation at 37 °C for 1 hour. Subsequently, cleavage reagent (proprietary information) was added to reactions made up to 20 μl and incubated for 5 min at room temperature.

A1.3.4 Quantitative polymerase chain reaction (qPCR)

cDNA generated from HEK293 or TZM-bl cells was subject to qPCR reactions using the Lightcycler® FastStart DNA Master^{PLUS} SYBR Green I Kit (Roche, Mannheim, Germany). Twenty microlitre reactions contained 1 μl cDNA (see A1.3.2), 20 μmol of each primer and 4 μl of Master Mix (proprietary information). Reactions were set up in chilled capillaries (Roche, Mannheim, Germany) that were centrifuged at 700 x *g* for 15 sec prior to loading into the Carousel-Based Lightcycler® V.2 System (Roche, Mannheim, Germany).

cDNA generated from SupT1 cells was subject to qPCR reactions using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK). As above, the reactions were 20 μl in volume and assembled in capillaries. Reactions contained 1 μl cDNA, 4 μl SensiMix™ (proprietary information), 1.5 μl enzyme (proprietary information), 0.4 μl SYBR solution (proprietary information) and 4 μmol of each primer.

A1.4 Viral procedures

A1.4.1 Propagation of HIV-1 subtype C isolate FV5

A thawed aliquot of FV5 was added to a 70% confluent 75 cm^2 flask of TZM-bl cells in the presence of 15 μg/ml DEAE-D. Twenty-four hours later, the media was removed and 10 ml of fresh media added to cells. Another 24 hours later, the media was collected, filtered, aliquoted and stored at -80 °C.

A.1.4.2 Generation of HIV-1 subtype B molecular clone p81A-4

A 25 cm^2 flask was seeded with 1.2 x 10^6 HEK293T cells. 24 hours later, 4 μg of p81A-4 was mixed with 150 μl serum-free DMEM containing 40 μl PolyFect transfection reagent (QIAGEN, CA, USA). The DNA mix was allowed to stand at room temperature for 5 min prior to addition to cells. Twenty-four hours later, the culture media was replaced with fresh media. Another 24 hours later, media was collected, filtered, made up to 20% FCS, aliquoted and stored at -80 °C.

A.1.4.3 Median tissue culture infectious dose ($TCID_{50}$) from luciferase output

TZM-bl cells were seeded at 1 x 10^4 cells per well in a 96-well. The remaining wells were filled with PBS to minimise plate dehydration. Twenty-four hours later, all media was removed. For subtype B molecular clones $pSG3_{\Delta}^{env}$.ZM53M and p81A-4, an 8-fold dilution of virus was prepared and added to the cells in triplicate, along with 100 μl media containing 30 μg/ml DEAE-D, resulting in a final concentration of DEAE-D of 15 μg/ml and a final dilution of virus of 4^{-2}. To the other wells, 150 μl media was added, containing 15 μg/ml DEAE-D. 50 μl of virus-containing media from the three 4^{-2} wells was transferred to another 3 wells ie a 4-fold dilution, 4^{-3}. This transfer of 50 μl was repeated across the plate until a virus dilution of 4^{-8} was obtained. Another set of wells received 150 μl of media only ie no virus. A similar dilution series was used in determining the $TCID_{50}$ for subtype C isolate FV5, but in this case an 8-fold dilution series was used, spanning 2^{-4} to 2^{-22}.

Twenty-four hours post-infection, cells were washed three times with 300 μl PBS and then received 150 μl media. Another 24 hours later, Tat-induced luciferase activity was quantified using the Bright-Glo™ Luciferase Assay System (Promega,WI, USA), as described in 2.2.10. Samples were considered luciferase positive if the luminescence signal was greater than that of the mean of the no virus samples plus two standard deviations.

$TCID_{50}$ was calculated using the Spearman-Karber method (Chou and Talalay 1984; Kahan *et al.* 1991):

$$M = k - [d(0.5 - (\frac{s}{n}))] \qquad \text{(A1.1)}$$

where *M* is the viral dilution at which 50% of cells are luciferase positive, *k* is the dose of the highest dilution (*k* = -8 for p81A-4 and *k* = -22 for FV5), *d* is the spacing between dilutions (*d* =

1 for p81A-4 and *d* = 3 for FV5), *s* is the number of negative samples and *n* is the wells per dilution (*n* = 3). The 50% endpoint was converted to a viral titre as follows:

$$TCID_{50}/ml = 10^{\log\left(\frac{1}{v}\right)} \times 10^{-M \log r} \quad \text{(A1.2)}$$

where *v* is the culture volume (ml) and *r* is the radix for the dilution series (*r* = 4 for p81A-4 and *r* = 2 for FV5).

A1.4.4 HIV-1 p24 ELISA

Culture supernatant isolated from HIV-1-infected samples were analysed for levels of the HIV-1 antigen p24 by ELISA using the Murex HIV antigen mAb (Murex Biotech Ltd, Kent, UK). One hundred microlitres of sample was incubated with 100 μl of conjugate 1 (biotinylated murine anti-p24 monoclonal antibodies, bovine albumin, Proclin® preservative, aggregated human IgG in PBS, heat inactivated mouse serum, and NP40, bovine casein and bovine aprotinin; concentrations unreported) in wells pre-coated with human anti-HIV polyclonal antibodies at 37 °C for 1 hour. In each assay, a negative control (human serum negative for HIV antigens) was included, as well as a positive control (recombinant p24) at various dilutions in matrix (PBS containing bovine proteins) to enable generation of a p24 standard curve and absolute determination of p24 levels in samples.

Following the incubation with conjugate 1, wells were washed five times with 300 μl wash fluid (PBS with Proclin® preservative). Two hundred microlitres conjugate 2 (peroxidase conjugated streptavidin, PBS, Proclin® preservative, bovine casein and bovine aprotinin) was then added to each well followed by an incubation at 37 °C for 30 min. Wells were washed as previously prior to addition of 200 μl substrate solution (tetramethylbenzidine, DMSO, thiomersal, phosphate citrate and hydrogen peroxide) and incubation at room temperature for 30 min. Finally, 50 μl of 2 M sulphuric acid was added to each well which were then left at room temperature for 5 to 15 min. Absorbance at a wavelength of 450 nm with 655 nm as reference was determined in a Model 680 microplate reader (BioRad, CA, USA).

Some samples were first analysed at various dilutions to determine the appropriate dilution to achieve an output within the range of the standard curve (A_{450} <1.2).

A.1.4.5 Median tissue culture infectious dose ($TCID_{50}$) from p24 output

SupT1 cells were seeded at 1 x 10^4 cells per well in a 96-well plate. The remaining wells were filled with PBS to minimise plate dehydration. 24 hours later, plates were centrifuged at 130

x *g* to pellet cells and all media was removed. For subtype B molecular clone p81A-4, cells were resuspended in 200 μl of 4-fold serial virus dilutions, ranging from 4^{-2} to 4^{-8}. For subtype C isolate FV5, cells were resuspended in 200 μl of 8-fold serial virus dilutions, ranging from 2^{-4} to 2^{-22}.

Twenty-four hours post-infection, plates were centrifuged at 130 x *g* and all media removed. Cells were then washed three times with PBS by addition of 300 μl per well followed by centrifugation at 130 x *g* and PBS removal. SupT1 cells were then resuspended in 350 μl media, centrifuged at 130 x *g* and 150 μl stored at -80 °C as day 0 samples. Cells were incubated for 7 days prior to centrifugation to pellet cells and removal of 150 μl media. Both day 0 and 7 samples were analysed by p24 ELISA (see A1.4.4), the former to verify that there was no residual infecting-virus following washing. Day 7 samples were run neat and classed as negative if the A_{450} was less than the absorbance of the negative control + 0.50. The viral titre was calculated as described in equations A1.1 and A1.2 (see A1.4.3).

A1.5 Other standard laboratory procedures

A1.5.1 Bicinchoninic acid (BCA) method for protein quantification

Protein in cell lysates was quantified using the Pierce® bicinchoninic acid (BCA) Protein Assay Kit (Thermo Fisher Scientific Inc, IL, USA). This method combines reduction of Cu^{+2} to Cu^{+1} by protein in an alkaline medium (the biuret reaction) with colourimetric detection of the cuprous cation (Cu^{+1}) using a reagent containing bicinchoninic acid. Briefly, 10 μl of cell lysate was mixed with 186 μl of Reagent A and 4 μl of Reagent B and incubated at 37 °C for 30 min. Absorbance at a wavelength of 562 nm was determined in a BioSpec-mini spectrophotometer (Shimadzu, Kyoto, Japan), blanked with a sample that contained 10 μl RIPA buffer instead of cell lysate. A standard curve, generated by L. Thompson using bovine serum albumin at various concentrations, was used to determine the concentration of protein in samples.

A2 Transcriptional gene silencing of HIV-dependency factors

A2.1 Introduction

This appendix describes initial attempts to induce classical TGS at HDFs and preceded the work reported in Chapter 4. This work was performed at The Scripps Research Institute, La Jolla, CA, USA, in collaboration with K. Morris.

TGS effectors may be processed from longer dsRNA precursors or from siRNAs by the RNAi pathway. In addition, promoter-targeted expressed antisense RNAs (asRNAs) alone have been shown to mediate TGS (Weinberg *et al.* 2006). PTGS of PSIP1/LEDGF/p75 (Llano *et al.* 2006a), HTATSF1 (Miller *et al.* 2009) and DDX3X (Ishaq *et al.* 2008) have all been shown to inhibit HIV-1 replication. Given the advantages of TGS over PTGS, this study investigated the possibility of inducing TGS at these genes with promoter-targeting asRNAs and/or siRNAs in HEK293T cells.

Both asRNA and siRNA target sites were initially selected by K. Morris using the Cold Spring Harbor (CSH) transcriptional regulatory element database (TRED) (Zhao *et al.* 2005; Jiang *et al.* 2007). Although some siRNAs did result in suppressed PSIP1 expression, further investigation into the mode of knock down, including examination of the levels of target region transcripts, the role of Ago proteins, histone modifications at the target region and transcription rate, supported a PTGS, not TGS, mechanism. Alignment of the target site with the genome revealed targeting of the *psip1* 3' UTR, not the promoter, demonstrating that regulatory element databases may be flawed in their prediction of TSSs.

A2.2 Materials and methods

A2.2.1 Design of target sites within HDF promoters

Target sequences were selected by K.Morris, The Scripps Research institute, CA, USA, as follows: promoters for *htatsf1* (Entrez gene ID 27336; promoter ID 43694), *psip1* (Entrez gene ID 11168; promoter ID 123760) and *ddx3x* (Entrez gene ID 1654; promoter ID 43224) were retrieved from the CSH TRED (http://rulai.cshl.edu/cgi-bin/TRED/tred.cgi?process=home) (Zhao *et al.* 2005; Jiang *et al.* 2007). Sequences 700 bp upstream of the designated TSS for these promoters were analysed by siRNA target finder (http://www.ambion.com/techlib/misc/siRNA_finder.html) (Ambion, TX, USA) (Table A2.1).

A2.2.2 Cloning of expressed asRNAs

Generation of asRNA expression cassettes was by a PCR approach. Amplification was carried out with a U6 forward primer (Table 2.2), containing a *Bam*HI restriction site, and a reverse primer that contained sequences complementary to the 3' end of the U6 promoter, the asRNA and a Pol III termination signal (Table A2.1). The DNA plasmid pTZU6+1 (Bertrand *et al.* 1997), which contains the human U6 small nuclear RNA promoter (GenBank® accession number X59362), was used as a template. PCR was performed on a Mastercycler® (BioRad, CA, USA) using Expand High Fidelity^{PLUS} PCR System reagents (Roche, Mannheim, Germany). The PCR reaction included 10 ng template, Expand HiFi^{PLUS} reaction buffer (proprietary information), 1.5 mM $MgCl_2$, 200 μM dNTP mix, 2.5 U of Expand HiFi^{PLUS} Taq polymerase and 10 pmol of each primer in a volume of 50 μl. Thermocycling conditions were 95 °C for 10 min, 30 cycles of denaturation at 95 °C, annealing at 62 °C and extension at 72 °C, each for 30 sec, followed by a final extension at 72 °C for 10 min. Amplification specificity was confirmed by resolution of PCR products on a 2% agarose gel (see A1.2.3).

PCR products were ligated directly into pTZ57R/T (Fermentas, WI, USA), a TA cloning vector. The ligation reaction contained a 3:1 molar ratio of insert to vector (180 fmol insert to 60 fmol vector) and 20 U T4 DNA ligase (New England Biolabs®, Inc., MA, USA) in a volume of 20 μl, which was incubated at room temperature overnight. Ten microlitres of the ligation reaction mixture was used to transform 100 μl of chemically competent *Escherichia coli* DH5α (Invitrogen, CA, USA) (see A1.1.1, A1.1.2 and A1.1.3). Transformed cells were plated on LB agar plates

containing ampicillin (Gibco, BRL, UK), X-gal (Sigma, MO, USA) and IPTG (Roche, Mannheim, Germany) and incubated at 37 °C overnight (see A1.1.1).

Table A2.1: Oligonucleotide primers used to generate asRNAs targeting HDF promoters.

Primer	Sequence (5'-3')	Target location (relative to TSS)
asH1 reverse	AAA**AAA**GAAGCGGAGTTTCCCGGCGGTGTTTCGTCCTTTCCACAA	-542 to -522
asH2 reverse	AAAA**AA**GTTCGAAGCTCGAGCGCCAGGTGTTTCGTCCTTTCCACAA	-471 to -451
asH3 reverse	AAAA**AA**GCTCGAGCGCCACTGAAATGGTGTTTCGTCCTTTCCACAA	-464 to -444
asH4 reverse	AA**AAAA**GACGCGGTTCCTGAGCGGGTGTTTCGTCCTTTCCACAA	-194 to -174
asP1 reverse	**AAAAAA**TCATGATGCTTATACGGTGTTTCGTCCTTTCCACAA	-571 to -551
asP2 reverse	AAA**AAA**TCATGGCTATATTGTTTTGGTGTTTCGTCCTTTCCACAA	-506 to -486
asP3 reverse	AAA**AAA**TGGCTTAAGACCAGCGTCGGTGTTTCGTCCTTTCCACAA	-256 to -236
asP4 reverse	AAAA**AA**GACCAGCGTCTCTGTATGAGGTGTTTCGTCCTTTCCACAA	-247 to -227
asP5 reverse	AAAA**AA**CCTGCATATATACCAGTCAGGTGTTTCGTCCTTTCCACAA	-122 to -102

Sequences complementary to the 3' end of the U6 promoter are underlined and to a Pol III termination signal are shaded grey. Those complementary to the asRNA (and, therefore, the same sequence as the target site) are in bold.

White colonies were screened for the presence of asRNA expression cassettes by culturing in 4 ml LB broth supplemented with ampicillin overnight at 37 °C with shaking (see A1.1.1). Plasmids were prepared using the High Pure Plasmid Isolation Kit (Roche, Mannheim, Germany) (see A1.2.1) and screened for insert presence and orientation by digesting 5 μg with 5 U *Eco*RI and 5 U of either *Hind*III or *Bam*HI (Fermentas, WI, USA) for I hour at 37 °C in a 20 μl reaction containing Tango Buffer (Fermentas, WI, USA). Digestion reactions were resolved on a 2% agarose gel (see A1.2.3). Clones that exhibited the desired digestion products were sequenced using M13 forward (5' GTA AAA CGA CGG CCA G 3') and reverse (5' CAG GAA ACA GCT ATG AC 3') primers by automated cycle sequencing (Inqaba Biotec, Pretoria, South Africa). Those containing correct asRNA expression cassettes were cultured in 300 ml ampicillin-supplemented LB overnight at 37 °C (see A1.1.1) with shaking prior to bulk plasmid DNA preparation using the QIAGEN Plasmid Maxi Kit (QIAGEN, CA, USA) (see A1.2.2).

A2.2.3 Design and generation of siRNAs

Target sequences were again selected by K. Morris, as described in A2.2.1, focusing on sequences within 200 bp upstream of the TSS, as empirically he had the most success inducing TGS when targeting this region. siRNAs were generated from oligonucleotide templates (Table A2.2) using the *Silencer*® siRNA Construction Kit (Ambion, TX, USA) (see 4.2.3).

Table A2.2: Oligonucleotide templates used to generate HDF promoter-targeting siRNAs.

siRNA	Strand	Oligonucleotide template	Target location (relative to TSS)
siH10	Sense/passenger	AACAGGACAAATTTCTCTGGCCCTGTCTC	-132 to -112
	Antisense/guide	*AAGCCAGAGAAATTTGTCCTG*CCTGTCTC	
siH11	Sense/passenger	AATGGGGCGGACAGGACAAATCCTGTCTC	-123 to -103
	Antisense/guide	*AAATTTGTCCTGTCCGCCCCA*CCTGTCTC	
siH12	Sense/passenger	AAGAAGGAAGGCGTGGGAATACCTGTCTC	-47 to -27
	Antisense/guide	*AATATTCCCACGCCTTCCTTC*CCTGTCTC	
siP10	Sense/passenger	AAACAAGTTTACACGTTGAACCCTGTCTC	-44 to -24
	Antisense/guide	*AAGTTCAACGTGTAAACTTGT*CCTGTCTC	
siP11	Sense/passenger	AAATGAAAACAAGTTTACACGCCTGTCTC	-38 to -18
	Antisense/guide	*AACGTGTAAACTTGTTTTCAT*CCTGTCTC	
siP12	Sense/passenger	AATTCTTTAATGAAAACAAGTCCTGTCTC	-31 to -11
	Antisense/guide	*AAACTTGTTTTCATTAAAGAA*CCTGTCTC	
siD10	Sense/passenger	AAAGGTCTAAGCTCAAGGGATCCTGTCTC	-157 to -137
	Antisense/guide	*AAATCCCTTGAGCTTAGACCT*CCTGTCTC	
siD11	Sense/passenger	AAAATTCCCTGTCGCGGCGACCCTGTCTC	-111 to -91
	Antisense/guide	*AAGTCGCCGCGACAGGGAATT*CCTGTCTC	
siD12	Sense/passenger	AACCTCCCTCTCACACCGCAACCTGTCTC	-94 to -74
	Antisense/guide	*AATTGCGGTGTGAGAGGGAGG*CCTGTCTC	

All siRNA oligonucleotide templates had the CCTGTCTC sequence at their 3' end to allow hybridisation with the T7 promoter primer within the *Silencer*® siRNA Construction Kit (Ambion, TX, USA). Target sequences are in italics. Two nucleotide overhangs are underlined.

A2.2.4 Screening for suppression of target gene expression

To assess whether asRNAs or siRNAs targeting HDF promoters were capable of reducing target mRNA levels, HEK293T cells were seeded at 2 x 10^5 cells per well in a 24-well plate. Twenty-four hours later, cells were transfected with 1 μg plasmid DNA per well (see 2.2.5) or 100 nM HDF-targeting siRNA (see 4.2.4) in triplicate. A negative control siRNA, R149, targeting the *ccr5* promoter (Entrez gene ID 1234) with guide sequence 5' CCG CCA AGA GAG CUU GAU AAA 3' was also used (gift from K. Morris) (Kim *et al.* 2006). Another negative control siRNA targeting the *gapdh* (Entrez gene ID 2597) mRNA, siGAPDH, was prepared using the control oligonucleotide templates (sequences not provided) supplied in the *Silencer*® siRNA Construction Kit (Ambion, TX, USA) (see 4.2.3). Seventy-two hours post-transfection, RNA was extracted (see A1.3.1), DNase-treated and reverse transcribed (see 3.2.12). qPCR on the resulting cDNA was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4) and

primers listed in Table 2.4. Where indicated, target gene expression was reported relative to *gapdh* mRNA (for primers see Table A2.3), rather than *actb*.

A2.2.5 Detection of promoter-associated transcripts

To examine whether transcripts were detectable in various regions of the *htatsf1* and *psip1* promoter, RNA was extracted from HEK293T cells. One microgram of RNA was DNase-treated and reverse transcribed (see 3.2.12) in duplicate. Another 1 μg of RNA per cell line was DNase-treated but not reverse transcribed to allow quantification of undigested genomic DNA (gDNA). An additional control lacking RNA was included per cell line.

qPCR was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4). Primers were designed to amplify regions encompassing the siRNA target sites and a region upstream of the siRNA target sites in the *psip1* promoter (Table A2.3). A coding region within *gapdh* was also amplified (Table A2.3), to which promoter transcript abundance could be compared. Expression of transcripts was calculated as described in 4.2.5 and qPCR products resolved on a 2% agarose gel (A1.2.3).

Table A2.3: Primers used to determine levels of promoter-associated transcripts.

Primer	Sequence (5'-3')	Amplicon length (bp)
psip1 upstream forward	AGCAACTGACATGGAACTGAGA	125
psip1 upstream reverse	TATACCCAGTCAGTTGTCTGCA	
psip1 siRNA target forward	AGAGGTCTTATAATGGAACTAGGA	191
psip1 siRNA target reverse	GCAAGAAAGTAATTTACTAGCAGA	
htatsf1 siRNA target forward	AAAAGACGCGGTTCCTGA	179
htatsf1 siRNA target reverse	GGAGGGAAACGGAAGGAA	
gapdh forward	CCACCCATGGCAAATTCC	69
gapdh reverse	TGGGATTTCCATTGATGACAAG	

A2.2.6 Examination of the role of Ago proteins

HEK293T cells were seeded at 2 x 10^5 cells per well in a 24-well plate. Immediately following administration of cells to culture plates, cells were transfected with 50 nM of siRNAs targeting either Ago1, Ago2 (Hawkins *et al.* 2009) or the HIV-1 LTR (Weinberg *et al.* 2006) (all siRNAs gifts from K. Morris) (Table A2.4), as described in 4.2.4. Twenty-four hours later, cells were again transfected, this time with siP10 or R149 at 100 nM (see 4.2.4) in triplicate. Seventy-two hours later, RNA was extracted (see A1.3.1), DNase-treated and reverse transcribed (see 3.2.12).

qPCR on the resulting cDNA was performed as described in 2.2.7 using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4) and primers given in Table 2.4, 4.2.7 and Table A2.3. Controls and cycling parameters were as described in 2.2.7. Gene expression was reported relative to *gapdh* mRNA.

Table A2.4: siRNAs used to examine the role of Ago proteins in siP10-mediated *psip1* silencing.

siRNA	Strand	Sequence (5'-3')
Anti-HIV LTR	Guide	GUGUUAGAGUGGAGGUUUGUU
	Passenger	CCAACCUCCACUCUAACACC
Anti-Ago1	Guide	CGUUGCCAAUGGGCAGUGCUU
	Passenger	GCACUGCCCAUUGGCAACG
Anti-Ago2	Guide	UUCAGAUGGACUUCCGUGCUU
	Passenger	GCACGGAAGUCCAUCUGAA

A2.2.7 Chromatin immunoprecipitation (ChIP) assay to examine epigenetic changes at siRNA-targeted regions

Costar® 60 cm^2 dishes (Corning, Inc., NY, USA) were seeded with 4 x 10^6 HEK293T cells. Cells were transfected immediately post-seeding with 100 nM siP10 or R149 in duplicate, as described in 4.2.4. Seventy-two hours later, formaldehyde was carefully added to cells to a final concentration of 1% and cells then incubated on a shaking platform for 10 min at room temperature. Crosslinking was stopped by the addition of glycine to a final concentration of 0.125 M and continued shaking for 5 min. Culture media was removed and 3 ml PBS with 0.1% (v/v) phenylmethanesulfonylfluoride (PMSF) (Sigma-Aldrich, MO, USA), a serine protease inhibitor, added to cells. Cells were harvested by scraping from dishes and centrifugation at 130 x *g* for 4 min. Cell pellets were lysed in 600 μl ChIP lysis buffer (50 mM HEPES, 140 mM NaCl, 1% Triton X-100, 0.1% sodium deoxycholate) with 0.1% (v/v) PMSF and incubated on ice for 10 min. Samples were centrifuged at 2,000 x *g* for 5 min at 4 °C to pellet nuclei. The pellet was again resuspended in 600 μl ChIP lysis buffer with 0.1% (v/v) PMSF and incubated on ice for 10 min. Samples were sonicated on ice to produce chromatin of an average length of 200 bp: 6 intervals of 25 sec pulses and 2 min rests on power setting 3 using a S-50 sonicator (Branson Ultrasonics Corp, CT, USA). Sonicated samples were centrifuged at 8,000 x *g* for 5 min at 4 °C and the supernatant of each sample transferred to a clean tube. Background was reduced by pre-clearing chromatin with 25 μl protein A sepharose beads blocked with salmon sperm (Invitrogen, CA, USA) and incubation on a gently rotating RM-2M Intelli-mixer (ELMI Ltd., Riga, Latvia) at room temperature for 15 min. Beads were removed by centrifugation at 16,000 x *g* and the chromatin-containing supernatant

was distributed between three tubes. The volume in each tube was adjusted to 1 ml with ChIP lysis buffer with 0.1% (v/v) PMSF. A tube of each sample received 1 μg of either anti-dimethylated histone 3 lysine 9 (H3K9me2) (Upstate #07-441; Millipore, MA, USA) or anti-trimethylated histone 3 lysine 27 (H3K27me3) (Upstate #07-449) antibody or no antibody and gently rotated on a RM-2M Intelli-mixer (ELMI Ltd., Riga, Latvia) at 4 °C overnight.

To samples 20 μl of magnetic protein G Dynabeads® blocked with salmon sperm (Invitrogen, CA, USA) was added prior to incubation on a gently rotating RM-2M Intelli-mixer (ELMI Ltd., Riga, Latvia) at room temperature for 15 min. Beads were separated using a magnet and supernatant from no antibody controls transferred to a clean tube and stored as 'input' samples. Supernatant from samples subject to immunoprecipitation was discarded.

Beads were washed with 600 μl ChIP lysis buffer, 600 μl ChIP lysis high salt buffer (50 mM HEPES, 500 mM NaCl, 1% Triton X-100 and 0.1% sodium deoxycholate) and 600 μl ChIP wash buffer (10 mM Tris.HCl, 250 mM LiCl, 1 mM EDTA, 0.5% sodium deoxycholate and 0.5% NP-40). For each wash, beads were resuspended and samples rotated for 3 min prior to bead separation with a magnet. Beads were finally resuspended in 100 μl elution buffer (50 mM Tris.HCl, 1% SDS and 10 mM EDTA) and incubated at 65 °C for 10 min to elute antibody-protein-DNA complexes. Beads were separated and the supernatant transferred to clean tubes. Elution was repeated and supernatant combined with that recovered from the first elution.

Samples, including input samples, were mixed with 10 μg RNase A (Invitrogen, CA, USA) and 20 μl of 5 M NaCl and incubated at 65 °C for 4 hours. To the combined eluates, 10 μl 0.5 M EDTA, 20 μl of 1 M Tris.HCl, pH 6.9, and 20 μg proteinase K (Roche, Mannheim, Germany) was added and incubated at 45 °C for 1 hour. DNA was recovered by addition of 250 μl phenol:chloroform:IAA in a 25:24:1 ratio (Ambion, TX, USA), centrifugation at 16,000 x *g* for 10 min at 4 °C, transfer of the aqueous layer to 300 μl ethanol, 100 μl 3 M sodium acetate and 3 μl glycogen (Ambion, TX, USA) and storage at -20 °C overnight.

DNA precipitate was recovered by centrifugation of samples at 16,000 x *g* for 10 min at 4 °C. Pellets were washed with 0.5 ml 75% ethanol, allowed to air-dry for 5 min and resuspended in 30 μl nuclease-free water. DNA was analysed by qPCR using primers spanning the siRNA target site and an upstream region (Table A2.3) using the SensiMix™ *Lite* Kit (Quantace Ltd, London, UK) (see A1.3.4). Total DNA isolated from HEK293T cells (1 ng, 100 pg, 10 pg and 1 pg) was also analysed by qPCR and used to construct a standard curve for each primer set using C_p medians, which were calculated by the Lightcycler® V.2 software (Roche, Mannheim, Germany). From the standard curve, the amount of DNA in each sample was calculated and then expressed as a

function of a radix of 10 (arbitrary units). No antibody values were subtracted as background and the remaining values expressed as a fraction of the input samples.

A2.2.8 Nuclear run-on assay to determine targeted HDF transcription rates

Transcription rates in siRNA-treated HEK293T cells were assessed as described in 3.2.15 and 4.2.8.

In the qPCR analysis of *psip1* transcription rates, previously used primers specific to the p75 isoform were used (Table 2.4), as well as primers that amplify a region common to both the p52 and p75 isoforms: p52 & p75 forward 5' GTG ACA ACA GCA ACA GCA TCT G 3'; and, p52 & p75 reverse 5' TGA TGT CAC TCT CTG AAG GAC AG 3'. Transcription rates are reported relative to that of *gapdh* using primers given in Table A2.3.

A2.2.9 Alignment of siRNA target sites to the genome

siRNA target sites were aligned to all human genomic and transcript sequences using nucleotide BLAST (blastn). The current standard reference assembly of the human genome, hg19/Genome Reference Consortium human genome build 37 (GRCh37) (May 2010), was used in alignments. No mismatches between genome sequence and siRNA target sites were permitted.

A2.2.10 Statistical analyses

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 $p < 0.05$. Data was analysed using unpaired, two-tailed t-tests, one-way ANOVAs, followed by a Dunnett's multiple comparison post-test, and two-way ANOVAs, followed by Bonferroni post-tests, where appropriate, calculated using GraphPad Prism 4.0c (GraphPad Software, Inc., CA, USA).

A2.3 Results

A2.3.1 asRNAs do not inhibit target gene expression

Host proteins known to be involved in HIV-1 replication were targeted for TGS. As asRNAs alone are sufficient to induce TGS (Weinberg *et al.* 2006; Turner *et al.* 2009), U6-driven asRNA expression cassettes targeting the *psip1* and *htatsf1* promoters were constructed. These were designed to express asRNAs with perfect sequence complementarity to regions up to 700 bp upstream of the TSS (Figure A2.1), which was retrieved from the CSH TRED by K. Morris.

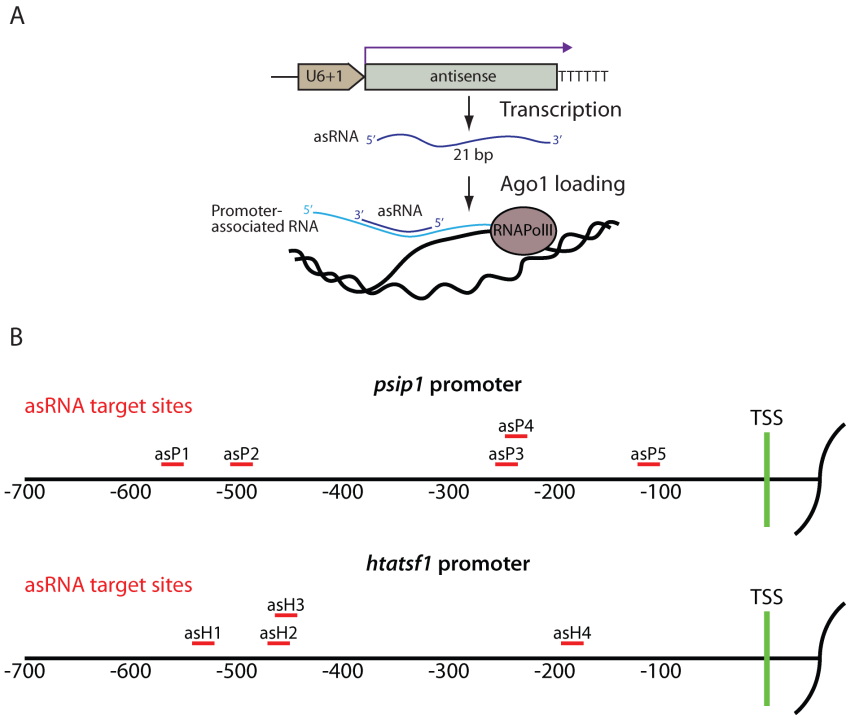


Figure A2.1: Promoter-targeting with expressed asRNAs. A. Schematic of the asRNA expression cassettes and their mode of targeting. **B.** The *psip1* and *htatsf1* promoters relative to the TSSs retrieved from the CSH TRED. Target sites selected from suggestions by the siRNA target finder (Ambion, TX, USA) are shown.

The asRNA expression plasmids were screened for their ability to suppress target gene expression in HEK293T cells. The asRNAs had no effect with a significant decrease in target expression only observed in cells receiving shRNA expression plasmids known to mediate effective PTGS (Figure A2.2).

A number of reasons may account for the lack of gene silencing with the asRNAs, including poor expression, a failure to be recruited into the RITS complex and/or inappropriate targeting within the gene promoter. To exclude the latter possibility, more target sites required examination. However, the time-consuming generation of asRNA expression plasmids is not

conducive to large-scale screening. Therefore, further screening for TGS was carried out using *in vitro* transcribed siRNAs.

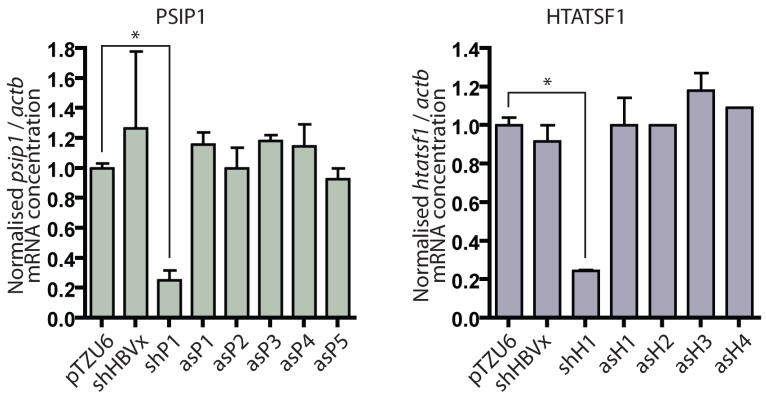


Figure A2.2: asRNAs do not suppress expression of target HDFs. asRNAs were screened for suppression of target expression. HEK293T cells were transfected with asRNA expression constructs, or a shRNA expression construct known to mediate target silencing as a positive control, in triplicate. qRT-PCR was performed on total RNA extracted 72 hours post-transfection. *, $p < 0.05$, one-way ANOVA with Dunnett post-tests relative to pTZU6.

A2.3.2 siRNAs inhibit *psip1* expression

Additional target sites were selected by K. Morris within 200 bp upstream of the CSH TRED TSS following an observation that this region is most conducive to TGS induction (personal communication). Six siRNAs were constructed and screened for their ability to suppress either *psip1* or *htatsf1* expression.

htatsf1 mRNA levels were unaltered in cells treated with promoter-targeting siRNAs compared with those receiving a control siRNA, R149, which targets the *ccr5* promoter (Figure A2.3). In contrast, all *psip1* promoter-targeting siRNAs resulted in a significant decrease in *psip1* expression, with siP10 having the greatest effect (0.39% of R149 control). Given that TGS induction requires promoter-associated transcripts (Han *et al.* 2007; Gonzalez *et al.* 2008; Schwartz *et al.* 2008; Hawkins *et al.* 2009; Tan *et al.* 2009b; Yue *et al.* 2010), whether the discord in *psip1* and *htatsf1* suppression related to levels of promoter-associated transcripts was investigated.

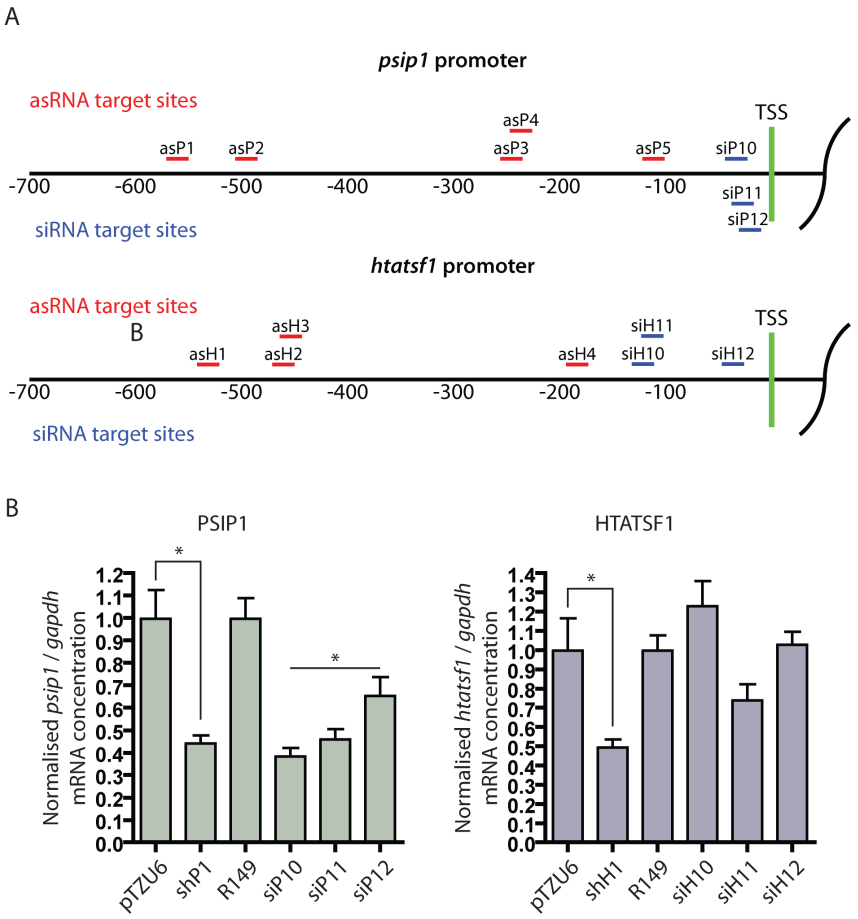


Figure A2.3: *psip1* promoter-targeted siRNAs can suppress target expression. A. The *psip1* and *htatsf1* promoters relative to the TSSs retrieved from the CSH TRED. siRNA target sites within 200 bp of the TSS were selected by the siRNA target factor (Ambion, TX, USA). **B.** Screen of siRNAs for suppression of target expression. HEK293T cells were transfected with siRNAs, or a shRNA expression construct known to mediate target silencing as a positive control, in triplicate. qRT-PCR was performed on total RNA extracted 72 hours post-transfection. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to R149/pTZU6.

A2.3.3 Transcripts overlap siRNA target sites

To quantify the levels of promoter-associated transcripts, primers were designed to amplify regions spanning the siRNA target sites (Figure A2.4A). Levels of transcripts from a region >700 bp upstream of the *psip1* TSS were also examined.

Transcripts were detectable by qRT-PCR from all regions (Figure A2.4B), although those transcribed from the *htatsf1* siRNA target site were >100-fold less abundant than those from either the *psip1* siRNA target site or *psip1* upstream region. Indeed, *htatsf1* siRNA target site amplicons were not sufficiently abundant to visualise on an agarose gel (Figure A2.4C). The low abundance of transcripts at the *htatsf1* siRNA target site may account for the lack of gene silencing.

Surprisingly, the levels of transcripts originating from both the *psip1* siRNA target site and upstream regions were similar to *gapdh* mRNA (Figure A2.4B), suggesting a high level of

transcription across the *psip1* promoter. The lack of 5' or 3' RACE data means, however, that it is unknown whether these transcripts are coding (with an extended 5' UTR) or non-coding. If coding, the gene silencing induced with siP10 may not be TGS, particularly as TGS and PTGS cannot be distinguished by qRT-PCR on total RNA. Thus, additional evidence for TGS as the mechanism of siP10-mediated *psip1* silencing was sought.

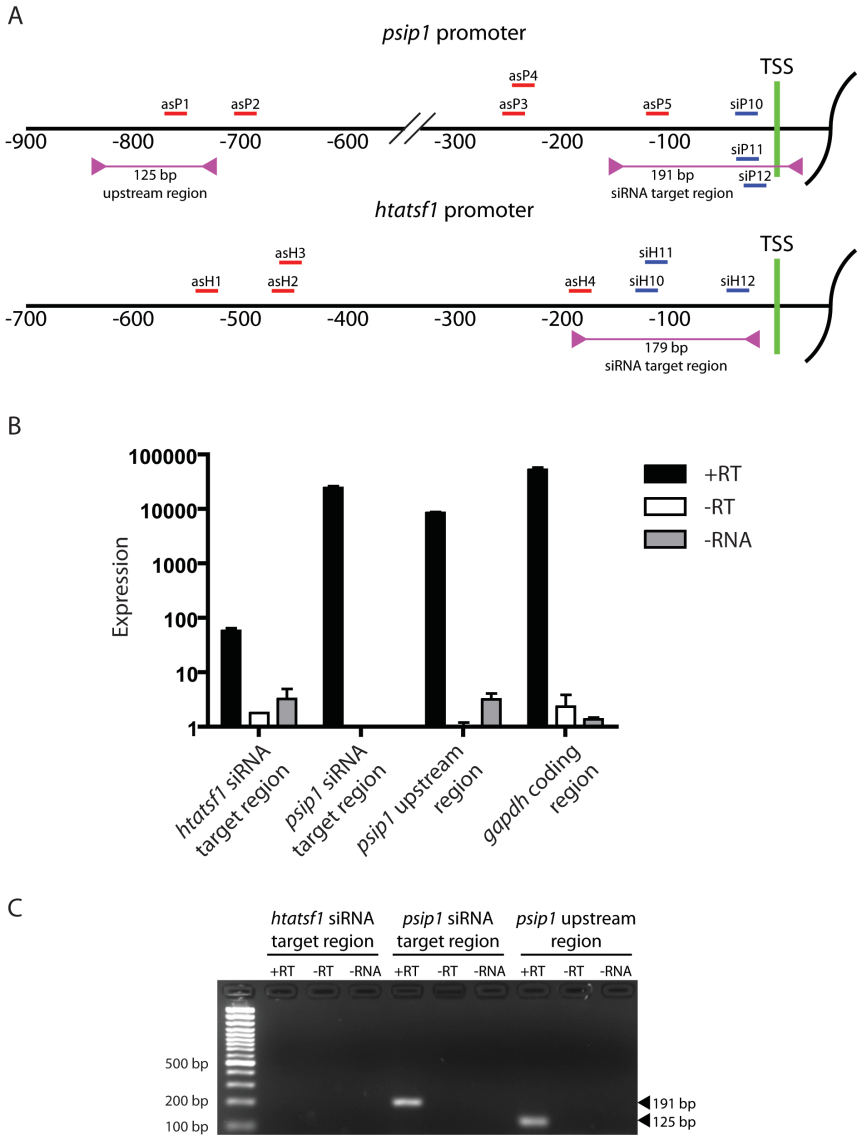


Figure A2.4: Transcription is detectable at siRNA target regions. **A.** *psip1* and *htatsf1* promoters showing the regions analysed for transcription. Primers were designed to amplify transcripts present at siRNA target regions and, for the *psip1* promoter, upstream. **B.** qRT-PCR was performed on total RNA extracted from HEK293T in duplicate. Primers specific to a region in the *gapdh* mRNA coding region were also included. **C.** Following qPCR, samples were resolved on an agarose gel.

A2.3.4 Ago2 suppression attenuates *psip1* silencing

The relative roles of Ago1 and Ago2 in siP10-mediated *psip1* silencing were investigated. HEK293T cells were first treated with siRNAs targeting either Ago1 or 2 or the HIV-1 LTR as a negative control. Twenty-four hours later, cells were again transfected, this time with the control siRNA, R149, or siP10. A slight attenuation of *psip1* silencing was observed in cells where levels of Ago2 were decreased but not in cells treated with the Ago1-targeting siRNA (Figure A2.5A), although this may be a result of the low-level of Ago1 mRNA suppression attained (Figure A2.5B). However, Ago mRNA suppression was quantified in RNA extracted 96 hours post-transfection and thus is probably less pronounced than at an earlier time point, when TGS induction would be expected to take place. In addition, the Ago2 mRNA suppression was also moderate 96 hours post-transfection, yet sufficient to have an effect on siP10-mediated silencing. Furthermore, Ago1 mRNA levels are increased when Ago2 is suppressed, which may indicate a mechanism to compensate for the loss of Ago2 function. However, despite this increased Ago1, there is still attenuation of the siP10-mediated silencing.

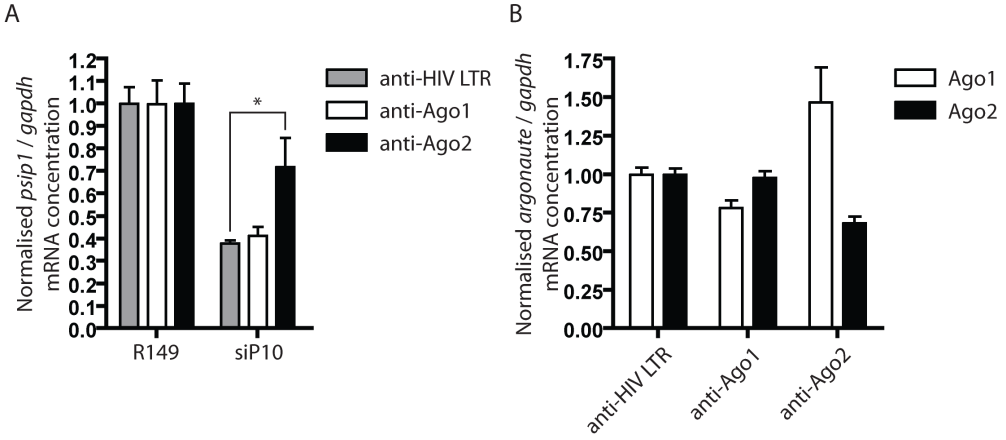


Figure A2.5: Ago2 suppression alleviates siP10-mediated *psip1* silencing. **A.** Both HEK293T cells were transfected with siRNAs targeting either Ago1 or Ago2 or a siRNA targeting the HIV-1 LTR as a control. Twenty-four hours later, cells were again transfected with promoter-targeting or control siRNAs in triplicate. *psip1* mRNA levels were assessed relative to *gapdh* by qRT-PCR on RNA extracted 96 hours post-primary transfection. **B.** As (A), but showing levels of Argonaute mRNA levels in cells transfected with R149. *, p <0.05, two-way ANOVA with Bonferroni post-tests.

Therefore, this data suggests that Ago2, and not Ago1, is necessary for siP10-mediated silencing of *psip1*. To further investigate whether siP10-mediated silencing is via classical TGS, the siRNA target site was examined for epigenetic changes.

A2.3.5 siP10 does not induce significant epigenetic changes at the siRNA target site

ChIP is a powerful technique for the examination of the interaction between specific proteins and DNA in the cell. Levels of H3K27me3 and H3K9me2 were assessed at the siP10 target

region and upstream as a control (Figure A2.4A). These histone modifications are associated with compact chromatin and, thus, transcriptionally-silent genes.

Despite a trend toward increased H3K27me3 at the siRNA target region in cells treated with siP10 compared to the R149-treated control, this was not statistically significant. Further, no enrichment for H3K9me2 at the target site was observed. The lack of significant silent-state histone modification enrichment at the target site does not support a classical TGS mechanism.

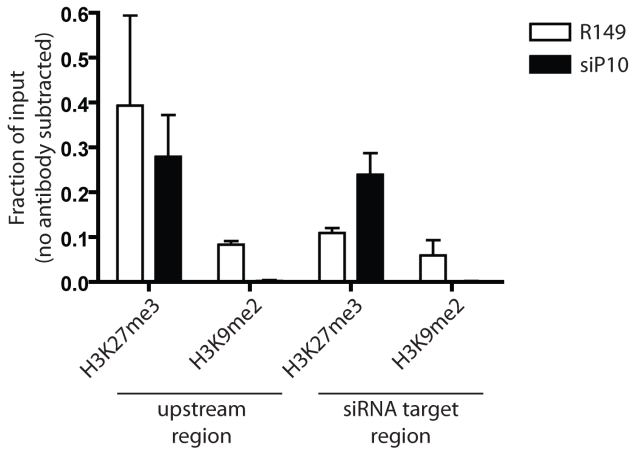


Figure A2.6: siRNA-mediated silencing of *psip1* is not accompanied by epigenetic changes at the target region. HEK293T cells transfected with siRNAs were analysed by ChIP 72 hours later in duplicate. Both the siRNA target region and an upstream region were examined for changes in the level of H3K27me3 and H3K9me2, two histone modifications associated with silent-state chromatin. Two-way ANOVA with Bonferroni post-tests.

A2.3.6 siP10 does not alter the rate of *psip1* transcription

A nuclear run-on experiment was performed to discern *psip1* transcription rates. In HEK293T cells treated with siP10, transcription of *psip1* was unaltered, whether quantified using primers that amplify a region common to the p52 and p75 isoforms, or primers specific to the p75 isoform (Figure A2.7). This result strongly implicates siP10 functioning via a PTGS, not a TGS, mechanism. This conclusion garners further support from the previous data showing a high abundance of transcripts originating from the target region, a role of Ago2, but not Ago1, in silencing and no accompanying epigenetic changes at the target site. Therefore, to verify that the target sites were within the *psip1* promoter, an alignment with the human genome was performed.

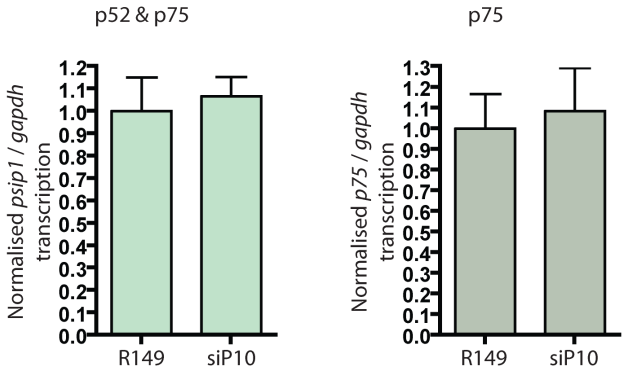


Figure A2.7: *psip1* transcription rates are unaltered by siRNA-mediated silencing. A nuclear run-on assay was performed on samples 72 hours post-transfection with promoter-targeting siRNAs in triplicate. Levels of *psip1* transcription were assessed using primers specific to the p75 isoform and common to both the p52 & p75 isoforms. Unpaired, two-tailed t-tests.

All siRNAs designed to target the *psip1* promoter were in fact found to be targeting a region within the RefSeq *psip1* 3' UTR, according to the latest human genome build, hg19/GRCh37. It is unsurprising, therefore, that the mode of siP10-mediated silencing was PTGS. The TSS listed in the CSH TRED, from which promoter sequences were retrieved, is also located within the 3'UTR, which demonstrates that TSSs listed in the CSH TRED may not be at the 5' end of a gene. Thus, promoter sequences retrieved from the CSH TRED, in the absence of alignments to verify their location, are an unreliable data source from which to design TGS target sites.

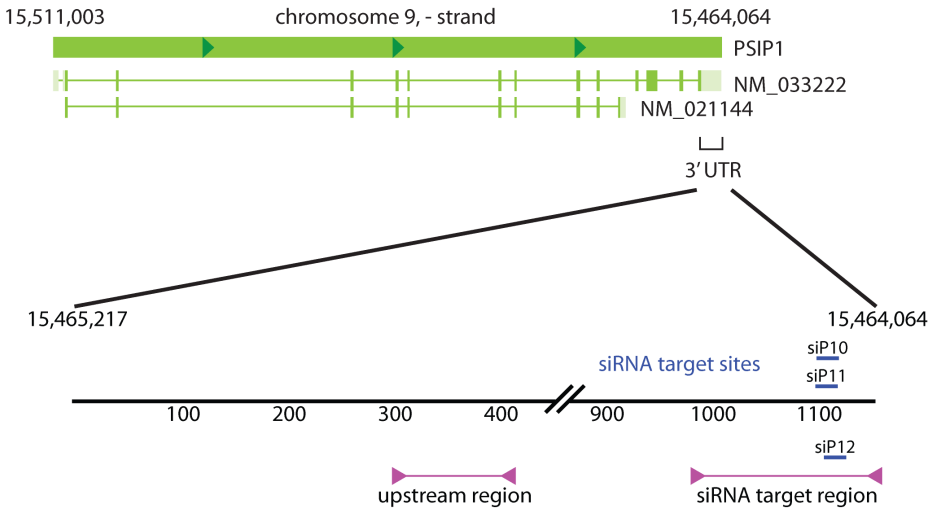


Figure A2.8: siRNA target sites do not align to the *psip1* promoter but the 3'UTR. siRNA target sites were aligned to the current standard reference assembly sequence of the human genome, build hg19/GRCh37, and found to map to the 3'UTR of *psip1*, not the promoter.

A2.3.7 siRNAs do not induce TGS of *ddx3x*

Similar problems were encountered in attempting to induce TGS of *ddx3x*: K. Morris designed siRNA target sites within 200 bp of the TSS retrieved from the CSH TRED. qRT-PCR

demonstrated significant mRNA knock down (Figure A2.9A), although this was not accompanied by a change in the *ddx3x* transcription rate (Figure A2.9B). Alignment with the human genome placed the siRNA target sites within the RefSeq *ddx3x* 5' UTR (Figure A2.9C) and, therefore, the CSH TRED within the 5' UTR also. Thus, the data suggests that the siRNAs are suppressing *ddx3x* mRNA expression via a PTGS mechanism.

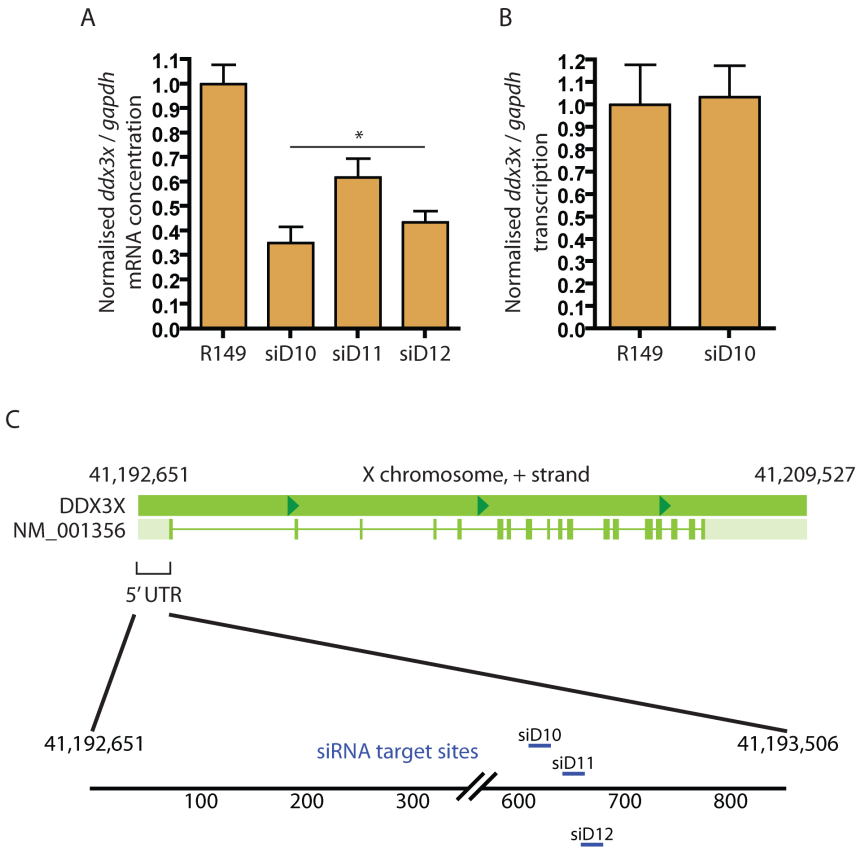


Figure A2.9: siRNAs targeting the CSH TRED *ddx3x* promoter do not induce TGS and align to the 5'UTR. **A.** siRNAs designed to target the *ddx3x* promoter based on the TSS retrieved from the CSH TRED database suppress *ddx3x* mRNA levels. HEK293T cells were transfected with siRNAs in triplicate and qRT-PCR performed on total RNA extracted 72 hours later. **B.** A nuclear run-on assay was performed 72 hours post-transfection with siRNAs. **C.** siRNA target sites were aligned to the current standard reference assembly sequence of the human genome, build hg19/GRCh37, and found to map to the 5'UTR of *ddx3x*. *, p <0.05, one-way ANOVA with Dunnett post-tests relative to R149 and unpaired, two-tailed t-test.

Together, the data demonstrates that PTGS, and not TGS, is the mechanism of *psip1* and *ddx3x* siRNA-mediated silencing. This follows inadvertently targeting regions within the coding transcript, and not the promoter, demonstrating that the CSH TRED promoter sequences require an alignment to verify their location prior to use in TGS target site design.

A2.4 Discussion

The approach used in this study was to design asRNA/siRNA target sites within a HDF promoter in the hope of inducing TGS. Unfortunately, TGS was not induced at either the *htatsf1*, *psip1* or *ddx3x* genes: no asRNAs or siRNAs reduced *htatsf1* mRNA levels (Figures A2.2 and A2.3) whilst siRNAs that were successful at reducing *psip1* and *dd3x* mRNA levels functioned by a PTGS mechanism (Figures A2.7 and A2.9). For *psip1*, this was supported by several lines of evidence; high levels of target region-associated transcripts (Figure A2.4), a role of Ago2 in silencing (Figure A2.5), a lack of silent-state histone modification enrichment at the target site (Figure A2.6) and, most convincingly, unaltered transcription rates (Figure A2.7). This was a result of the sequences inadvertently used in the design of siRNA target sites aligning to coding transcripts rather than promoters (Figure A2.8).

The CSH TRED lists several TSS per gene, and it is unknown on what basis K. Morris selected a TSS, particularly as the quality (where lower numbers reflect a higher quality level, as determined by the reliability of the supporting evidence) of the *psip1* TSS retrieved was only '3.2: refseq' versus another of the same quality and one of higher, ie ' 3.1: refseq, predicted', the latter correlating with the 5' end of the RefSeq gene. Nevertheless, it seems peculiar that this database annotates a region within the *psip1* 3' UTR as a TSS. The CSH TRED claims to annotate human, mouse and rat promoters through an automated pipeline that extracts known promoters from databases, including Genbank, the Eukaryotic Promoter Database (EPD) and the Database of Transcription Start Sites (DBTSS), and employs a promoter-finding program, Human First Exon Finder (FirstEF), combined with mRNA/EST information, cross-species comparisons and hand curation (Zhao *et al.* 2005). However, despite an additional modality being added in 2007 (Jiang *et al.* 2007), the underlying CSH TRED content has not been reappraised since its assembly in 2005. Since then, deep sequencing CAGE technologies, such as those used in the FANTOM project, have increased depth and accuracy of promoter annotation. Consultation of the FANTOM genome browser revealed no predicted promoter within the 3' UTR of *psip1* (http://fantom.gsc.riken.jp/4/gev/gbrowse/hg18/).

Given that the TSS retrieved was actually in the 3' UTR of the *psip1* gene, the asRNA target sites were also within the 3' UTR. It is surprising, therefore, that no gene silencing was observed with the asRNA expression plasmids. Although it is possible that none of the target sites were appropriate, given that the same algorithm designed the siRNA target sites, which all proved effective, it seems unlikely.

asRNAs have been shown to mediate TGS (Weinberg *et al.* 2006) and can even recruit Ago1 to the target site (Turner *et al.* 2009). However, it is unknown whether asRNAs are

incorporated as efficiently into Ago2-containing RISC as Ago1-containing RITS. One study has shown that incorporation of siRNAs into Ago1 versus Ago2 is influenced by the degree of complementarity between strands, with perfectly-matched duplexes being preferentially incorporated into Ago2 and bulged duplexes into Ago1 (Gonzalez *et al.* 2008). Moreover, an influence of duplex thermodynamics on the associated Ago protein has precedence in *Drosophila* (Tomari *et al.* 2007). One can envisage then that a single-stranded asRNA, with no complementary base-pairing, may be inefficiently incorporated into Ago2 complexes.

Of the three TSS listed on the CSH TRED for *ddx3x*, K. Morris did retrieve the TSS ranked of the highest quality (ie '2: known'), which corresponds to the promoter annotated by the FANTOM project. Thus, although targeting downstream of the extreme 5' end of the RefSeq gene, it is possible that the siRNAs were indeed targeting the *ddx3x* promoter. Nevertheless, the nuclear run-on assay showed no change in transcription rates supporting a PTGS mechanism. Unlike *S. cerevisiase* (David *et al*. 2006), transcription of human genes rarely commences from a defined TSS but rather over a region (Juven-Gershon *et al.* 2008), with the size of this region varying considerably between genes (Balwierz *et al.* 2009). The TSS can also vary between tissues and conditions and, as such, the most active TSS in HEK239T cells could differ from that described by the FANTOM4 project, which focused on a differentiating macrophage-like cancer cell line, THP-1 (Kawaji *et al.* 2009). It is possible, therefore, that whilst the siRNAs were targeting the most defined promoter region, this same region could be part of a coding transcript in HEK293T cells. Indeed, of the two other TSS for *ddx3x* given in the CSH TRED, one does correlate with the 5' end of the RefSeq gene (Promoter ID 43226; 764 bp upstream of targeted Promoter ID 43224) and the other is further upstream still (promoter ID 43225 – 543 bp upstream of Promoter ID 43226 and the 5' end of the RefSeq gene). If either of these were the most active TSS in HEK239T cells, the siRNAs would be targeting a coding transcript. The primary TSS in HEK293T cells could be interrogated by 5' RACE.

In summary, before designing siRNA target sites, 5' RACE data for the cell line to be used should be obtained. Alternatively, promoters should be retrieved from recently compiled data sources, such as those available through the FANTOM project, that use deep sequencing CAGE identification of TSSs combined with transcription factor binding sites, predicted by algorithms and experimentally determined by ChIP-Seq, to annotate promoters. Although TGS was not successfully induced, this study demonstrates that the experiments performed were sufficient to discriminate PTGS from TGS. Such interrogation of the observed suppression of gene expression is advised before a verdict of classical TGS is reached.