Optimization of anti-HIV RNAs for use in combination therapy to achieve a functional HIV cure

Ryan Goguen

Faculty of Medicine
Department of Microbiology and Immunology
McGill University, Montréal, Québec, Canada
April 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

Table of Contents

List of Figures

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter VI

List of Tables

Chapter I

Chapter II

Chapter III

Abstract

Current drug therapies against human immunodeficiency virus (HIV) infection can prevent the clinical progression to AIDS but requires continuous administration and is not able to eliminate the infection. An alternative long-term treatment option, which could eliminate the need for combination antiretroviral therapy (cART) and its daily administration, is gene therapy using lentiviruses expressing anti-HIV RNA molecules. As is the case with cART, a combination of at least two antiviral RNA molecules will most likely be necessary to avoid the emergence of resistance and viral rebound.

Anti-HIV gene therapy has advanced to clinical trials in a few instances, but the effectiveness in controlling the infection in the absence of cART is limited despite being safe and well tolerated in people living with HIV (PLWH). The overall goal of this project was to optimize different anti-HIV RNAs to maximize inhibitory effects while avoiding the occurrence of cytotoxicity. Two different classes of RNAs were optimized and used as antivirals. One of the RNA classes we investigated were short hairpin (sh)RNAs which bind to their viral RNA target and direct the RNA-induced silencing complex (RISC) to cleave it. The other RNA class evaluated in this project are aptamers, which bind and inactivate a target viral protein or RNA through their three-dimensional structure. We optimized these two RNA classes by evaluating their expression from the RNA polymerase (Pol) III promoters H1, 7SK and U6. We also explored various molecular designs such as the inclusion of flanking RNA secondary structures as well as combining shRNAs and aptamers into a single transcript.

In this project, we show that anti-HIV shRNAs are more potent when expressed by the promoters U6 and 7SK compared to the H1 promoter due to a higher expression level. Although we have

demonstrated that cytotoxicity exists when certain shRNAs are expressed from the U6 and 7SK promoters, we have discovered that replacing the loop of shRNAs with an aptamer to generate aptamer-shRNA chimeras effectively eliminates this toxicity. We were successful in achieving the goal of maximizing antiviral effects in the absence of cytotoxicity. The project led to the identification of several aptamer-shRNA chimeras that could strongly restrict viral replication without negatively impacting cell growth, with maximal inhibitory effects observed using the 7SK promoter. Overall, our work in optimizing the expression of anti-HIV RNAs is critical to designing the most effective combination gene therapies and has led to the discovery of a method to prevent shRNA-mediated cytotoxicity.

Résumé

Les thérapies antirétrovirales actuelles contre l'infection par le VIH peuvent empêcher la progression clinique vers le SIDA mais nécessitent une administration permanente et ne sont pas capables d'éliminer l'infection. Une autre option de traitement à long terme qui pourrait éliminer le besoin d'une thérapie antirétrovirale combinée (ARVc) ainsi que son administration quotidienne serait une thérapie génique utilisant des lentivirus pour exprimer des molécules d'ARN anti-VIH. Tout comme les thérapies ARVc utilisées aujourd'hui, une combinaison d'au moins deux molécules d'ARN antivirales seront certainement nécessaires pour éviter l'émergence de résistance et le rebond viral.

La thérapie génique anti-VIH a fait l'objet d'essais cliniques dans quelques cas, mais l'efficacité du contrôle de l'infection en l'absence de thérapie ARVc est limitée bien qu'elle soit sûre et bien tolérée par les personnes vivant avec le VIH (PVVIH). L'objectif global de ce projet était d'optimiser différents ARN anti-VIH afin de maximiser les effets inhibiteurs tout en évitant l'apparition de cytotoxicité. Deux classes différentes d'ARN ont été optimisées et utilisées comme antiviraux. L'une des classes d'ARN que nous avons utilisée est représentée par des petits ARN en épingle à cheveux (shRNA) qui se lient à un ARN viral pour diriger le complexe de mise en silence par l'ARN (RISC) vers sa cible, où il va cliver l'ARN viral cible. L'autre classe d'ARN utilisée dans ce projet regroupe des aptamères qui se lient et inactivent une protéine ou un ARN viral cible grâce à leur structure tridimensionnelle. Nous avons optimisé ces deux classes d'ARN en évaluant leur expression à partir des promoteurs H1, 7SK et U6 qui utilisent l'ARN polymérase (Pol) III. Nous avons également exploré des conceptions moléculaires diverses telles que l'inclusion de

structures secondaires d'ARN flanquantes ainsi que la combinaison de shARN et d'aptamères en un seul transcrit.

Dans ce projet, nous montrons que les shARN anti-VIH sont de meilleurs inhibiteurs lorsqu'ils sont exprimés par les promoteurs U6 et 7SK de l'ARN Pol III par rapport au promoteur H1 en raison d'un niveau d'expression plus élevé. Bien que nous ayons démontré qu'il existe de la cytotoxicité lorsque certains shARN sont exprimés à partir du promoteur U6 et 7SK, nous avons découvert que le remplacement de la boucle des shARN par un aptamère pour générer des chimères aptamère-shARN permet d'éliminer cette toxicité. Nous avons réussi à atteindre notre objectif de maximiser les effets antiviraux en l'absence de cytotoxicité. Le projet a conduit à l'identification de plusieurs chimères aptamère-shARN qui inhibent fortement la réplication virale sans impact négatif sur la croissance cellulaire, avec des effets inhibiteurs maximaux observés lors de l'expression à partir du promoteur 7SK. Dans l'ensemble, nos travaux d'optimisation de l'expression des ARN anti-VIH sont essentiels à la conception de thérapies géniques combinées efficaces et ont permis la découverte d'une méthode pour prévenir la cytotoxicité générée par les shARN.

Acknowledgements

I would like to begin by acknowledging my supervisor, Dr. Anne Gatignol, as well as Dr. Robert Scarborough. Both of you served as mentors during my PhD training and I could not have completed this degree without your guidance. You both created a comfortable learning environment to not only develop my scientific skills, but to refine my writing, presentation and critical thinking skills. Thank you for all your feedback over the years, the numerous edits for each publication, poster presentation and oral presentation, as well as our conversations relating to the direction to pursue the research project. All of this has allowed me to become a competent and productive scientist. Anne has dedicated her entire career to our understanding of the molecular biology of HIV with the goal of developing new treatment options to increase the quality of life of infected individuals and I am honored to be her final PhD student contributing to this goal.

I would also like to acknowledge my advisory committee members, Dr. Chen Liang and Dr. Pascale Legault, for their guidance over the years as you were both extremely helpful in structuring my research priorities at year long intervals. Additionally, I would like to thank Christian Young for his help with flow cytometry and cell sorting, which made up a significant bulk of the results of this thesis. Furthermore, the technical help that I received from Aïcha Daher was indispensable in allowing me to be successful in the design and execution of numerous experiments ranging from RNA work to cloning to cell culture. To past lab members Sergio AL, Roman R, and Olivier DC, each of you have not only helped me professionally but have been a source of friendship through which many laughs have been shared.

I acknowledge the Fonds de Recherche du Québec – Santé (FRQS) for financially supporting me with a generous scholarship and for the financial support provided by the Canadian Institutes of Health Research (CIHR) to my supervisor.

Finally, I would like to acknowledge my family for their love and support. To Amanda, who listened to my nightly stories of experiments succeeding or failing and is my biggest supporter in life. To my daughter Lucy, who makes each day exciting as you discover the world. To my parents, who have always encouraged me to chase my dreams. Thank you all for being part of this journey and a part of my life.

List of Abbreviations

AAV	Adeno-associated virus
Ago	Argonaute
AIDS	Acquired immunodeficiency syndrome
APOBEC3G	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide like 3G
ARV	AIDS associated retroviruses
AZT	Zidovudine
BACH2	BTB and CNC homolog 2
BCL11B	B cell leukemia 11b
BETis	Bromodomain and extra-terminal motif protein inhibitors
bNABs	Broadly neutralizing antibodies
Brd4	Bromodomain-containing protein 4
CA	Capsid
Cas9	CRISPR associated protein 9
cART	Combination antiretroviral therapy
CCR5	C-C chemokine receptor type 5
CD	Cluster of differentiation
CDK9	Cyclin-dependent kinase 9
cDNA	Complementary DNA
CNS	Central nervous system
COVID-19	Coronavirus disease 2019
CPM	Counts per minute
CRISPR	Clustered regularly interspaced short palindromic repeats
CRM1	Chromosomal maintenance 1
CTL	Cytotoxic T lymphocyte
CXCR4	C-X-C chemokine receptor type 4
CycT1	Cyclin T1
DDX3	DEAD-box helicase 3

DGCR8	DiGeorge syndrome critical region gene 8
DIS	Dimerization initiation site
DNMT	DNA methyltransferase
dNTP	Deoxynucleotide triphosphate
DSB	Double-stranded break
dsRBD	dsRNA-binding domain
dsRNA	Double stranded RNA
eIF2	Eukaryotic translation initiation factor 2
ESCRT	Endosomal sorting complexes required for transport
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FDA	Food and drug administration
GALT	Gut-associated lymphoid tissues
GFP	Green fluorescent protein
Gp	Glycoprotein
gRNA	guide RNA
GVHD	Graft versus host disease
HAART	Highly active antiretroviral therapy
HDAC	Histone deacetylase
HDR	Homology-directed repair
HDV	Hepatitis delta virus
HEK	Human embryonic kidney
HEXIM1	Hexamethylene bisacetamide inducible proteins 1
HH	Hammerhead
HIV	Human immunodeficiency virus
HMT	Histone methyltransferase
Hp	Hairpin
HSC	Hematopoietic stem cell
HSV	Herpes simplex virus
HTLV	Human T-cell leukemia virus

ICAM-1	Intercellular adhesion molecule 1
IFN	Interferon
IL	Interleukin
Imp7	Importin 7
IN	Integrase
Indel	Insertion or deletion
LANL	Los Alamos National Laboratory
LAV	Lymphadenopathy associated virus
LDLR	Low-density lipoprotein receptor
LEDGF	Lens-epithelium-derived growth factor
lncRNA	Long non coding RNA
LRA	Latency reversing agent
LTNP	Long-term nonprogressor
LTR	Long terminal repeat
LV	Lentivirus
MA	Matrix
MCS	Multiple cloning site
Medipal	Merlin, Dicer, protein kinase R [PKR] activator [PACT] liaison
MHC-I	Major histocompatibility complex class I
MID	Middle domain
miRNA	MicroRNA
MKL2	Myocardin-like 2
mTOR	Mammalian target of rapamycin
MuLV	Murine leukemia virus
NC	Nucleocapsid
Nef	Negative regulatory factor
NF-κB	Nuclear factor kappa B
NHEJ	Non-homologous end joining
NNRTI	Nonnucleoside reverse transcriptase inhibitor
NRON	Non-coding repressor of NFAT

NRTI	Nucleoside reverse transcriptase inhibitor
ORF	Open reading frame
PAP	Poly(A) polymerase
PAS	Polyadenylation site
PAZ	Piwi/Argonaute/Zwille
P-bodies	Processing bodies
PBS	Primer binding site
PEI	Polyethylenimine
PEP	Post-exposure prophylaxis
PIC	Preintegration complex
piRNA	Piwi RNA
PIWI	P-element induced wimpy testis
PKR	Protein kinase R
Pol	Polymerase
PPT	Polypurine tract
PR	Protease
PrEP	Pre-exposure prophylaxis
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
P-TEFb	Positive transcription elongation factor b
PTGS	Post-transcriptional gene silencing
Rev	Regulator of expression of virion
RIG-I	Retinoic acid inducible gene I
RISC	RNA induced silencing complex
RITS	RNA-induced transcriptional silencing
RNAi	RNA interference
RRE	Rev response element
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RTC	Reverse transcription complex

RT-PCR	Reverse transcription PCR
SAMHD1	Sterile alpha motif and HD domain 1
SARS-COV-2	Severe acute respiratory syndrome coronavirus 2
SCID	Severe combined immunodeficiency
SCT	Stem cell transplant
SELEX	Systematic evolution of ligands by exponential enrichment
SEM	Standard error mean
SERINC3	Serine incorporator 3
SERINC5	Serine incorporator 5
SF3B1	Splicing factor 3B subunit 1
SHAPE	Selective 2′-hydroxyl acylation analyzed by primer extension
shRNA	Short hairpin RNA
SIN	Self-inactivating
siRNA	Small interfering RNA
SIV	Simian immunodeficiency virus
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein complex
SOFA	Specific on/off adaptor
Ss	Splice sites
SSC	Saline sodium citrate
STAT5B	Signal transducer and activator of transcription 5B
TALEN	Transcription activator-like effector nuclease
TALE	TAL effector
TAR	Trans-activating response element
TAT	Trans-activator of transcription
TDN	Trans-dominant negative
TGS	Transcriptional gene silencing
TLR	Toll like receptor
TRBP	TAR RNA binding protein
TRIM	Tripartite motif-containing protein

tRNA	Transfer RNA
TRN-SR2	Importin α/importin β heterodimer and transportin SR2
U1i	U1 interference
UPF1	Upframeshift protein 1
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U
VSV-G	Vesicular stomatitis virus G protein
WAGO	Worm-specific Argonautes
WST	Water soluble tetrazolium salt
ZFN	Zinc finger nuclease

Contribution to original knowledge

The work presented in this thesis elucidates novel findings in the field of anti-HIV gene therapy, which contribute to maximizing antiviral effects while avoiding cytotoxicity. Overall, this thesis presents a collection of work considered original scholarship along with important contributions to knowledge. The presented information shows that both the expression strategy and molecular design of anti-HIV RNAs have large effects on their inhibitory capabilities and cytotoxic potential. Specifically, the thesis shows:

Chapter II

- The U6 and 7SK promoters express more potent anti-HIV shRNAs than the H1 promoter due to a higher transcriptional activity.
- Cytotoxicity is frequently elicited when shRNAs are expressed from the U6 and 7SK promoters
- The U6 promoter is more accurate than the H1 and 7SK promoter in using the +1 transcriptional start site but this has no effect on guide strand identity as Dicer cleavage sites are measured from the first base pair of the duplex.
- There exist inconsistencies between these results and similar investigations reported in the literature. We therefore recommend independently optimizing any other anti-HIV shRNAs.

Chapter III

- shRNA-mediated cytotoxicity is alleviated by replacing the loop sequences of shRNAs with large aptamers to generate aptamer-shRNA chimeras.

- Cells expressing aptamer-shRNA chimeras contain lower levels of shRNA guide strands compared to when canonical shRNAs are expressed.
- Current data suggests that replacing the loop sequences of shRNAs with large aptamers may compromise the efficiency of shRNA processing, causing an alleviation of cytotoxicity.

Chapter IV

- Selecting anti-HIV shRNAs previously reported in the literature based on viral target site conservation allows for the identification of highly potent shRNAs.
- The nucleotide identity of anti-HIV shRNAs is also a determinant towards cytotoxic potential as one shRNA was not cytotoxic when expressed from the U6 and 7SK promoters.
- Aptamer-shRNA chimeras are less potent as antivirals compared to canonical shRNAs but their lack of cytotoxicity makes them desirable as therapeutics.

Manuscripts not included in this thesis for which the candidate made significant contributions before or during the course of his thesis work include:

Goguen, R.P., Chen, M.J., Dunkley, O.R.S., Gatignol, A. & Scarborough, R.J. 2023. Gene therapy to cure HIV infection. *Virologie*, In press.

Goguen, R.P., Gatignol, A. & Scarborough, R.J. Cloning and Detection of Aptamer-Ribozyme Conjugations. 2021. *Methods Mol Biol* **2167**, 253-267.

Scarborough, R.J., Goguen, R.P. & Gatignol, A. 2019. A second patient cured of HIV infection: hopes and limitations. *Virologie (Montrouge)* **23**, 1-4.

Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. 2019. *Mol Ther Nucleic Acids* **18**, 815-830.

Contribution of authors

I hereby disclose the author contribution for each chapter

Chapter I is a literature review of HIV-1, its molecular dynamics, and potential cure strategies.

Chapter II is an original paper on the expression strategies of anti-HIV shRNAs: “Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK and H1 promoters”. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021). AG and RJS conceived the study. RJS and RPG designed the experiments, analyzed the data and wrote the manuscript. AG revised the manuscript. RPG conducted most of the experiments. ODC, CMGM, AD, SPAL and MJC assisted in conducting experiments. All authors read and approved of the manuscript.

Chapter III is an original paper exploring the aptamer-shRNA chimera molecular design and the mechanism leading to the alleviation of shRNA-mediated cytotoxicity: “Aptamer-shRNA chimeras as a molecular design to prevent shRNA-mediated cytotoxicity”. RPG, AG and RJS conceived the study. RJS and RPG designed the experiments and analyzed the data. RPG wrote the manuscript. RJS and AG revised the manuscript. RPG conducted most of the experiments. AD and MJC assisted in conducting experiments. All authors read and approved of the manuscript.

Chapter IV is an original paper evaluating the expression strategy and molecular design of previously published anti-HIV shRNAs: “Anti-HIV aptamer-shRNA chimeras expressed from the U6 and 7SK promoters strongly inhibit HIV replication in lymphocytic cells”. RPG, AG and RJS conceived the study. RJS, RPG and CMGM designed the experiments and analyzed the data. RPG wrote the manuscript. RJS and AG revised the manuscript. RPG, CMGM and MJC conducted

most of the experiments. AD assisted in conducting experiments. All authors read and approved of the manuscript.

Chapter I

Introduction

1.1 History of HIV/AIDS

1.1.1 Lentivirus family and the emergence of HIV

Lentiviruses constitute a genus of retroviridae and the molecular hallmark of their infection is their ability to integrate viral complementary DNA (cDNA) into the chromosome of the host cell. The retroviridae family is divided into the orthoretroviridae and spumaretrovirinae subfamilies, and these subfamilies are in turn divided into seven genera[1]. Lentivirus is a genus of the orthoretrovirinae subfamily, along with alpharetrovirus, betaretrovirus, gammaretrovirus, deltaretrovirus and epsilonretrovirus. These pathogens have existed for millions of years and have contributed to shaping the evolution of humans and many other vertebrates. For example, genetic analyses show that the placental barrier evolved through co-opting of ancient retroviral envelope genes[2]. As much as 8% of the human genome consists of ancient retroviruses generated from integration events within germline cells[3]. Whereas five genera of the orthoretrovirinae subfamily are potentially oncogenic, lentiviruses are not oncogenic but can cause serious chronic diseases in their hosts. Of contemporary interest is the human immunodeficiency virus (HIV) which was identified as the lentivirus causing acquired immunodeficiency syndrome (AIDS)[4].

Lentiviruses can be further classified based on their natural host. Simian immunodeficiency viruses (SIVs) consist of a group of species that infect African non-human primates. Both HIV-1 and HIV-2 arose from cross-species transmission of SIVs. The ancestry of HIV-1 has been tracked to an SIV found in chimpanzees (SIV_{cpz})[5]. The jump into the human population occurred from four independent cross-species infection events which resulted in the HIV-1 groups M-P[6] (Figure 1.1). Groups M and N arose from two distinct direct jumps of SIV_{cpz} to humans while groups O and P crossed the interspecies barrier to gorillas (SIV_{gor}) before jumping to humans[7] (Figure 1.1). Although each of the four groups can lead to the development of AIDS, group M is the principal

cause of the current AIDS pandemic[8]. Groups N and P consist of a few cases isolated in Cameroon while group O is also present in other West African countries[7]. In the case of HIV-2, it originated from an SIV ancestor within sooty mangabey monkeys (SIV_{smm})[9] (Figure 1.1). Two successful zoonotic events led to the occurrence of HIV-2 A and B found in West Africa, which can cause AIDS and are transmissible, while other cross-species infections have led to non-transmissible dead-end infections[10].

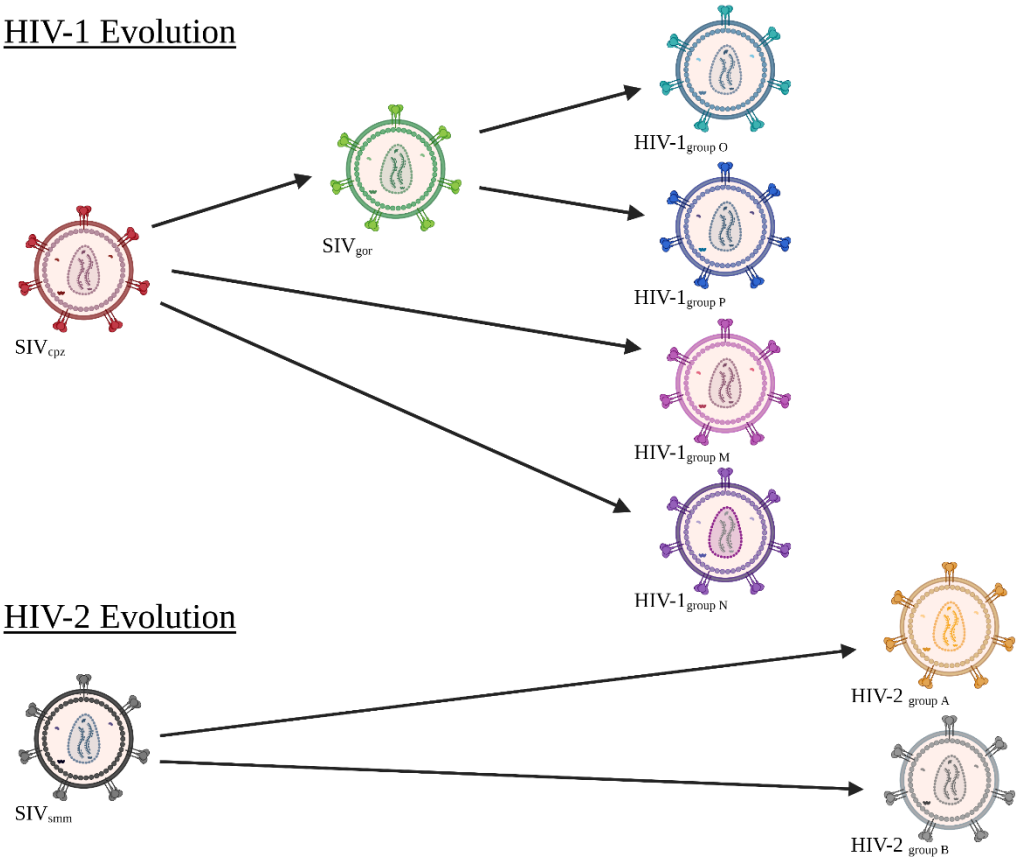

Figure 1.1 Evolution of HIV-1 & HIV-2. HIV-1 arose from four different interspecies infection events origination from SIV_{cpz} or SIV_{gor}, while HIV-2 emerged from cross-species infections originating from SIV_{smm}.

1.1.2 HIV pandemic origins

Viral pathogens have plagued humankind throughout history. Many of these have led to pandemics, which are defined as diseases that have spread on an international scale with multiple countries or continents being affected. Notable viruses that have led to a significant burden of disease include smallpox, various influenza viruses, which includes the H1N1 strain of 1918 known as the "Spanish flu", and severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) responsible for the currently ongoing Coronavirus Disease 2019 (COVID-19) pandemic. Various other large viral outbreaks such as the 2013-2016 West African Ebola virus epidemic have threatened to become widespread, but their transmission was controlled through public health interventions. Other isolated outbreaks of Ebola virus have since occurred across central Africa, prompting further public health interventions to limit their spread. The 2009 H1N1 flu pandemic had largely spread around the world, but the mortality rate was limited and is now considered as an endemic seasonal flu. The 2013-16 Zika virus pandemic has reached all tropical countries, had consequences for new-born babies from infected mothers, but has currently waned. The HIV/AIDS global pandemic is the longest lasting pandemic, with a duration of over 40 years since the first documentation of AIDS in 1981[11]. With no end in sight to the HIV/AIDS pandemic, it is important to recognize the clinical progress that has been made since the virus was discovered and the areas of study that are still needed to end this pandemic.

The interspecies jump of HIV-1 group M into the human population likely occurred in Kinshasa around 1920 where it circulated within brothels among railway and sex workers[12]. Signs of HIV-1 circulating among the population was only noticed in the 1980s when physicians in the United States were documenting clusters of mysterious illnesses that progressed to immune system deficiencies within homosexual men[11,13]. The disease was initially named the "gay plague" and

negative views of the homosexual community discouraged funding and research from being allocated towards discovering the cause of the disease. Identification of the causative agent of AIDS was initially made in 1983 by scientists in Dr. Luc Montagnier's lab, where a new retrovirus was isolated from the lymph nodes of patients affected by an immune disorder and was named lymphadenopathy associated virus (LAV)[14]. The same year, Dr. Robert Gallo's lab considered the virus causing AIDS to be similar to the previously identified Human T-cell Leukemia virus (HTLV) and called it as HTLV-III[15] , but a year later, his group recognized HTLV-III as a different virus causing AIDS, but kept the initial name[16]. Similar retroviruses were also independently isolated in 1984 by Dr. Jay Levy's lab and named AIDS associated retroviruses (ARV)[17]. After long argumentations, the final name of the retrovirus was decided in 1986 to be HIV by the International Committee on the Taxonomy of Viruses[4].

1.1.3 Current state of the HIV-1 pandemic and efforts to limit spread

The HIV/AIDS global pandemic, mainly caused by HIV-1, has resulted in an estimated 79.3 million infected individuals and 38.4 million people were estimated to be living with HIV as of 2021[18]. Only 85% of PLWH knew their HIV status in 2021, which serves as a significant factor to the 1.5 million new infections occurring annually. The global distribution of HIV is not homogenous, Eastern and Southern Africa have the highest number of PLWH which accounts for over 50% of the global HIV infections[18]. The primary contributors towards the high incidence of HIV in these areas include migration patterns and the lack of appropriate medical infrastructure[19,20]. The HIV/AIDS pandemic also disproportionately affects certain ethnicities which have unequal access to health care and often live in poverty. For example, black women in the U.S. are disproportionally affected by HIV-1 where gendered power imbalances as well as economic disadvantages are strong social determinants affecting the pandemic in this population[21].

Access to HIV care has also been affected by the COVID-19 pandemic as public transportation has been restricted and financial hardships are common. These challenges have led to treatment interruption in some individuals[22,23]. Although protocols to mail antiretroviral drugs have been developed in certain countries, this can lead to unwanted exposure of HIV status where some individuals conceal their status for fear of stigmatization from family members[24].

While current combination antiretroviral therapy (cART) can decrease viremia to undetectable levels in infected individuals, only 75% of PLWH had access to such treatment in 2021. Public health efforts to control the virus with current therapies focus on increasing the proportion of infected individuals who know of their HIV positive status as well as increasing the proportion of people who can access cART. The goal is to implement a 90-90-90 plan (90% HIV diagnosed, 90% on therapy and 90% suppressed) where annual new infections would decrease sufficiently for the pandemic to die out over time and would prevent illness, death as well as health care cost[25]. A new obstacle in reaching the 90-90-90 goal has arisen from the COVID-19 pandemic, where lockdowns have disrupted HIV testing and caused a decline in diagnoses[26]. Although 90-90-90 is an important goal to reach given the current therapies that we have available, the development of innovative treatment strategies are desperately needed to offer a cure for HIV. Such cure strategies that have been suggested by scientists include “shock and kill”, “block and lock”, immune therapies, as well as gene therapy[27,28].

1.2 Human immunodeficiency virus

1.2.1 HIV-1 genome and viral proteins

The genome of HIV-1 consists of a full-length 9-10 kb positive-strand RNA molecule, with a single virion containing two copies of this positive sense ssRNA genome. The genome includes various regulatory elements such as the long terminal repeats (LTRs), the trans-activating response

element (TAR) and the rev response element (RRE)[29]. Multiple genes are present in the HIV-1 genome which are expressed from the full-length or alternatively spliced genomic RNA to produce envelope, structural, enzymatic, accessory and regulatory proteins. The structural and enzymatic proteins come from the Gag and Gag-Pol polyproteins, which are alternatively translated from the full-length RNA by a ribosomal frameshift. The occurrence of a ribosomal frameshift during translation of Gag leads to the readthrough of a stop codon to generate the Gag-Pol polyprotein. The Gag polyprotein includes the matrix (MA), capsid (CA) and nucleocapsid (NC) proteins, which are generated after viral protease (PR) cleavage[30]. The Gag-Pol polyprotein includes the aforementioned proteins as well as the viral enzymes PR, reverse transcriptase (RT) and integrase (IN) (Figure 1.2A). The envelope glycoprotein 160 (gp160) is generated by the *env* gene and is processed by host cellular protease furin into transmembrane (gp41) and surface (gp120) subunits[31]. Additional regulatory and accessory proteins assist in viral replication such as the trans-activator of transcription (Tat), the regulator of expression of virion proteins (Rev), viral protein R (Vpr), viral protein U (Vpu), viral infectivity factor (Vif) and the negative regulatory factor (Nef)[32] (Figure 1.2A).

Mature HIV-1 particles possess an outer envelope acquired through budding from host cell membrane which is lined by an inner viral MA shell[33] (Figure 1.2B). Envelope proteins are present in the viral membrane as trimers to allow for entry into host cell where gp120 interacts with surface cell receptors and gp41 is responsible for membrane fusion[29]. The viral core lies within the viral membrane and is formed from the assembly of capsid molecules, which in turn contains NC coated viral genomic RNA as well as the viral proteins RT, IN, Vif and Vpr[34] (Figure 1.2B). The functions of each viral protein is expanded upon in Table 1.1.

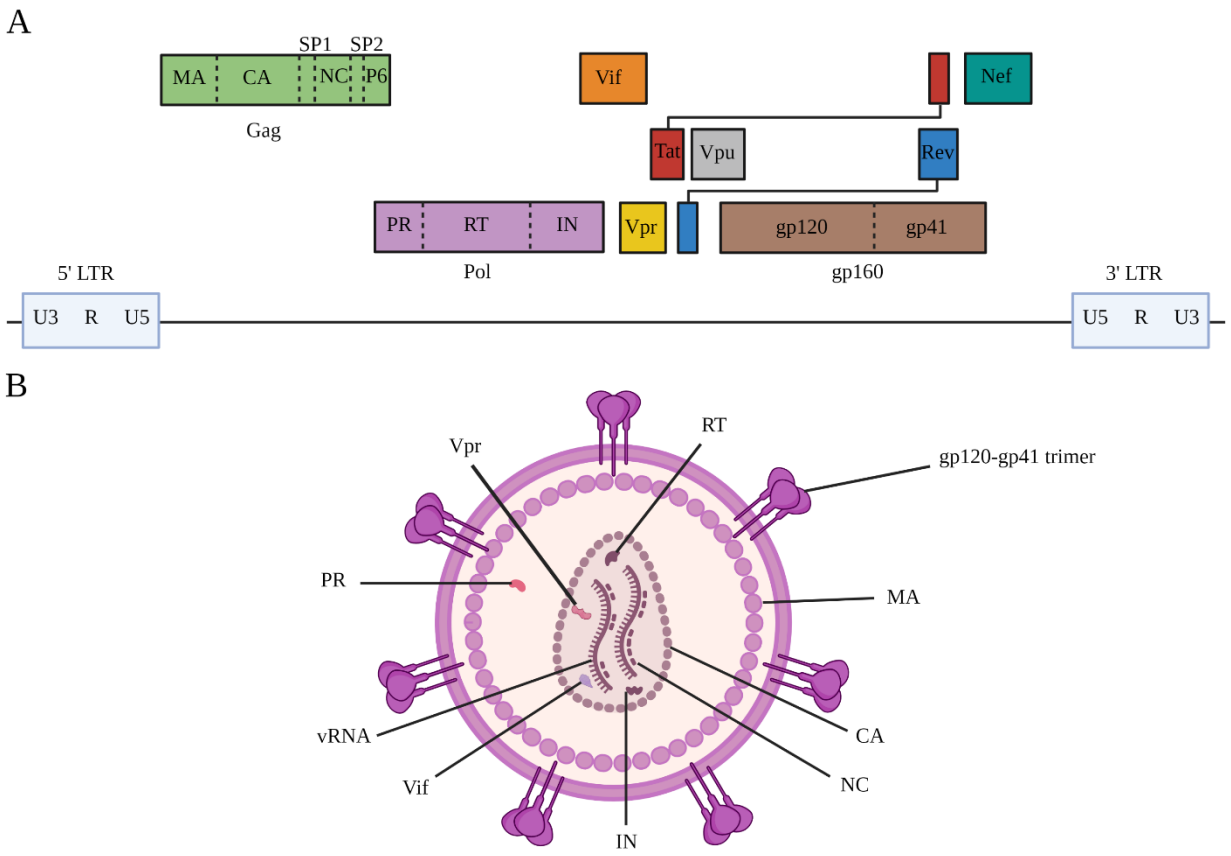

Figure 1.2 Structure of HIV-1 virion and viral genome. A) Alternative splicing of viral transcripts as well as various open reading frames produce a total of 16 viral proteins. B) The viral envelope is made up of a cell membrane acquired during budding with the gp120-gp41 trimer anchored within. MA coats the inner surface of the membrane which encapsulates the viral core. The viral core is made up of CA which protects the two strands of viral RNA bound by NC. The viral proteins PR, IN, RT, Vpr and Vif are also present within mature virions. Adapted from Scarborough R.J, 2015[35].

Table 1.1 Function of viral proteins.

Viral protein	Role in replication step	Protein function
Matrix (MA)	Assembly and budding	Targets Gag to the plasma membrane for virion assembly
Capsid (CA)	Assembly	Forms the virion core
Nucleocapsid (NC)	Assembly and reverse transcription	Coats and stabilizes the viral RNA genome, acts as a chaperone during RT
P6	Budding	Promotes budding by interacting with the ESCRT machinery
Protease (PR)	Maturation	Proteolytically cleaves the Gag and Gag-Pol polyproteins
Reverse Transcriptase (RT)	Reverse transcription	Converts viral RNA into cDNA
Integrase (IN)	Integration	Inserts viral cDNA into host chromosome
Vif	Reverse transcription	Inhibits the deamination of cytidine to uridine function of host APOBEC3G
Vpr	Integration and viral enhancer of infectivity	Assists in the nuclear entry of viral cDNA and maintains the cell at the G2/M phase
Tat	Viral transcription	Greatly enhances the rate of proviral transcription
Rev	Export of viral transcripts	Exports unspliced and incompletely spliced viral transcripts from nucleus to cytoplasm
Vpu	Budding	Induces degradation of cellular CD4 and tetherin
Gp120	Entry	Binds to the CD4 primary receptor and CCR5 or CXCR4 secondary receptor of target cells
Gp41	Entry	Mediates fusion of viral & cellular membrane
Nef	Viral enhancer of infectivity	Downregulates cell surface CD4 & MHC-I, promotes cell activation, counteracts host restriction factors

1.2.2 Overview of the HIV-1 replication cycle

HIV-1 predominantly infects surface cluster of differentiation 4 $(CD4)^+$ T lymphocytes ($CD4^+$ T cells) and macrophages but also dendritic cells[36]. The replication cycle is centralized around the ability of the virus to integrate cDNA into cellular chromosomes. Entry begins with interaction between gp120 and the CD4 receptor present on a cell followed by subsequent binding to a coreceptor, typically C-C chemokine receptor type 5 (CCR5) or C-X-C chemokine receptor type 4 (CXCR4). Once the virus has entered the host cell, the viral core will migrate towards the nucleus while reverse transcription and uncoating begin which will result in dsDNA from the viral RNA and the formation of the preintegration complex (PIC) (Figure 1.3). Once the viral core has localized within the nucleus, reverse transcription and uncoating steps are finalized to release the completed viral dsDNA. This cDNA is then processed by viral IN for integration within the host chromosomal DNA to generate the provirus[29]. Integrated provirus will subsequently begin viral transcription which is largely enhanced by the action of Tat (Figure 1.3). The full length 9-10 kb RNA is synthesized and various different transcripts can be generated from this full length RNA by alternative splicing[37]. The multiple open reading frames (ORFs) of these viral transcripts will then be translated into viral proteins necessary for replication, assembly and budding or will be packaged into newly formed viral particles. Viral particles containing the RNA genome will bud from the host cell to acquire an outer envelope and will mature through internal cleavage of polyproteins by viral PR.

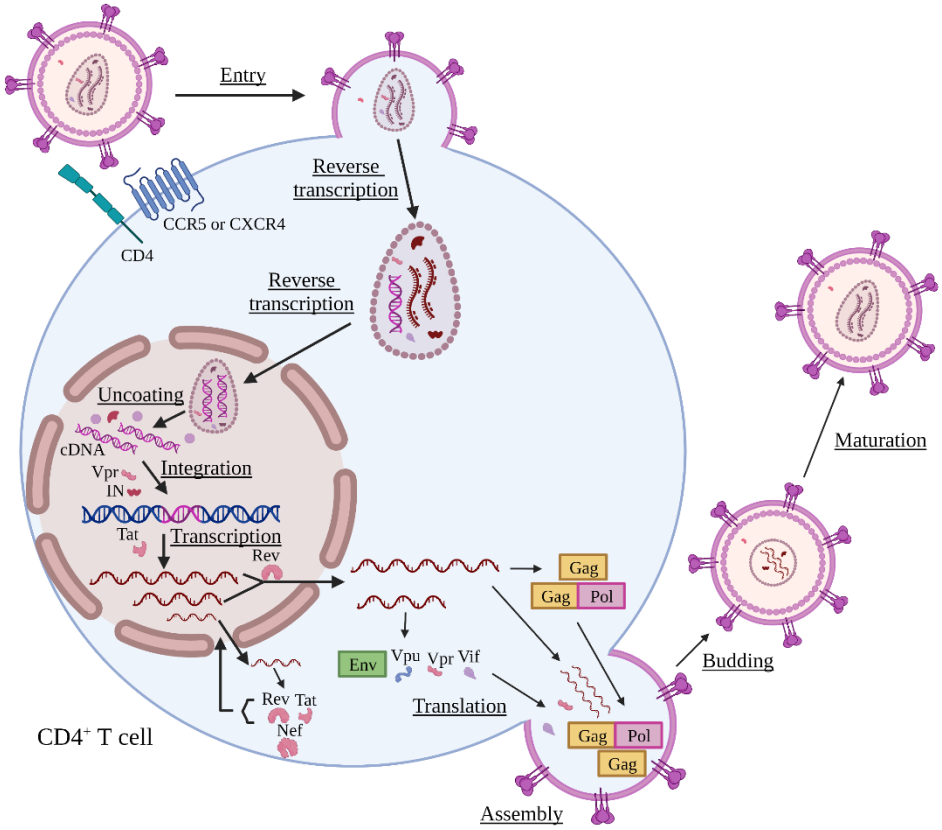

Figure 1.3 HIV-1 replication cycle. HIV-1 enters lymphocytes through its attachment to CD4 and to CCR5 or CXCR4. The viral genomic RNA is reverse transcribed into cDNA within the core that will form the PIC, which migrates to the nucleus. Inside the nucleus the IN will mediate the integration of the HIV-1 cDNA into the host chromosome to form a provirus. Proviral transcription and translation generates the various proteins needed for virion assembly which buds from the host cell membrane. Adapted from Arhel, N.J., 2020[38].

1.2.2.1 Virus binding and entry

Viral entry into $CD4^+$ T cells, macrophages and dendritic cells begins with nonspecific attachment of the virus to an encountered cell. This initial virus-cell attachment is mediated by gp120 interactions with surface molecules such as proteoglycans or integrins and facilitates the subsequent gp120-CD4 interaction by reducing the proximity between the two molecules[39,40]. HIV-1 entry is dependent upon binding of viral envelope proteins to the CD4 receptor and CCR5

or CXCR4 co-receptor. Once the initial interaction of gp120 and CD4 takes place, conformational changes occur within gp120 which lead to stronger contact with CD4 along with binding to the CCR5 or CXCR4 coreceptor[40]. Coreceptor preference is dependent upon the stage of HIV-1 infection, where CCR5 is primarily used during acute infection but CXCR4 becomes more utilized as the infection progresses[41]. Binding of gp120 to a coreceptor elicits a conformational change in gp41 to expose a fusion peptide which inserts into the host cell membrane. This event brings the amino-terminus of gp41 close to the cell membrane while the carboxy-terminus remains close to the viral membrane[39]. The action of bringing together the two membranes then forces the formation of a fusion pore through which the viral core can enter the host cell. Following entry, uncoating of the viral core was classically thought to occur in the cytoplasm either at the start or the completion of reverse transcription[34]. However, recent evidence suggests that the conversion of viral RNA to cDNA is initiated inside an intact viral core within the cytoplasm, where reverse transcription and uncoating steps are finalized in the nucleus following nuclear import[38,42].

1.2.2.2 HIV-1 reverse transcription

The HIV-1 RT is responsible for reverse transcription of viral RNA into cDNA, a complex process involving multiple steps. RT consists of a heterodimer composed of two subunits, known as p66 and p51[43]. Both subunits are generated by proteolytic cleavage of Gag-Pol polyprotein by viral PR. Maturation of viral RT involves the transition from a homodimer to a heterodimer protein through folding rearrangements which cause one subunit to be proteolytically labile[44]. While both subunits include a polymerase domain, p66 also contains an RNase H domain which cleaves RNA that is part of an RNA/DNA duplex[45].

Reverse transcription begins through the formation of the reverse transcription complex (RTC) within the viral core and continues during and after uncoating of the viral core[42]. This

complex is made up of the viral RNA genome as well as the viral proteins RT, NC, IN, CA and Vpr[46]. Reverse transcription in the case of HIV-1 is primed by cell-derived transfer RNA lysine 3 ($tRNA^{lys3}$), which is captured within the viral particles during budding[47]. This tRNA is complementary and binds to the primer binding site (PBS) located at the 5' end of the viral RNA. This binding leads to the synthesis of a short 100-150 base DNA fragment which includes the R and U5 region of the 5' LTR. The 5' end of the viral genome that was used as template is then degraded by the RNase-H domain of RT[43]. The R region present within the formed DNA fragment allows for the first strand transfer to occur where it relocates to the 3' end of the viral RNA to bind to the complementary R region present there[47]. The DNA fragment then primes further DNA synthesis from its 3' end where the remaining viral genome is used as template. Simultaneously, the RNase H domain of RT digests this template RNA strand which results in a single negative strand DNA molecule. However, there are two critical GC rich areas where degradation of the template does not occur and these act as primers for synthesis of the secondary positive DNA strand. These areas that resists RNase H degradation are known as the polypurine tracts (PPT) and are located at the center as well as the 3' end of the viral RNA genome[47]. Synthesis of the positive DNA strand begins from the 3' ends of both the central and 3' PPT. The DNA fragment generated at the 3' PPT includes the U3, R, U5 and PBS sequences. The PBS sequence present within this positive DNA fragment synthesized at the 3' PPT allows for a second strand transfer to occur where it relocates to the 5' end of the negative DNA strand to anneal through the complementary PBS sequence present there. From this point, synthesis of the positive DNA strand continues at the 3' end of the PBS and goes through the central PPT to displace the 5' end of the positive DNA fragment generated at the central PPT which produces a DNA flap of approximately 100 nucleotides[48,49]. The overall resulting product of reverse transcription is thus a continuous minus

DNA strand and a discontinuous plus DNA strand containing a flap at the center. The flap is eventually removed, and the plus DNA strand is made continuous prior to integration by cellular enzymes.

1.2.2.3 HIV-1 integration

Once reverse transcription is complete, the resulting viral cDNA associated to the elements of the RTC will form the PIC. Like RT, IN is encoded within the *pol* gene and is generated by proteolytic cleavage of Gag-Pol polyprotein by viral PR[50]. Once the PIC has been localized to the nuclear membrane, transport across this membrane must occur for the viral cDNA to be delivered into the nucleus. Nuclear membrane transport has been shown to be dependent on interactions between viral elements present in the PIC and host molecules present on the nuclear surface such as importin 7 (Imp7), the importin α/importin β heterodimer and transportin SR2 (TRN-SR2)[51-53]. The viral proteins MA, Vpr, IN and CA as well as the central flap present within the viral cDNA are all elements that contribute to ensuring the PIC is translocated across the nucleus[54,55].

After nuclear translocation, two nucleotides from the 3’ ends of each DNA strand of the viral cDNA are removed by IN, creating reactive hydroxyl groups which can mediate nucleophilic attacks on both strands of host DNA. The 5’ phosphate ends resulting from the nucleophilic attack are then covalently joined to the ends of the viral cDNA[56]. As the two sites of joining are separated by 5 nucleotides, there is a resulting 5 base pair duplication on each side of the integrated provirus following repair by the cellular DNA repair machinery[57]. The lens-epithelium-derived growth factor (LEDGF) is critical in binding IN to the host chromatin[58]. LEDGF is also implicated in targeting the integration site of the viral cDNA towards active transcriptional units. Targeting of active transcriptional units by LEDGF is mediated by recognition of highly spliced regions of the host genome[59].

1.2.2.4 Transcription and posttranscriptional control

Integration of the provirus into transcriptionally active regions of host chromatin is important to ensure that initial viral transcription occurs. Proviral transcription is driven by features of the 5' LTR such as the RNA polymerase (Pol) II promoter, transcription factor binding sites and enhancer segments. These elements are located within the U3 region of the LTR. Notable constituents of the 5' LTR that facilitate transcription are nuclear factor kappa B (NF-κB) binding sites, the TATA box and SP1 binding sites located within the promoter region[60-62]. Although these features stimulate the initial synthesis of transcripts, efficient rates of viral transcription requires interaction between the TAR element, synthesized from the 5' end of the R region, and the viral protein Tat[63]. Early viral transcripts undergo double cellular splicing and contain the genes coding for Rev, Nef and Tat[64]. Tat and Rev both contribute to increasing the number of viral transcripts in the cytoplasm required for replication.

The mechanism of transcription elongation by Tat is critical, as in the absence of Tat the RNA Pol II stalls at the 5' end of the viral transcript. For efficient function, Tat requires the recruitment of cellular factors in the pre-initiation complex around the TATA box and a super elongation complex to pursue transcription elongation[65-67]. Binding of Tat to the TAR RNA motif, an RNA hairpin present at the 5' end of all viral transcripts, leads to recruitment of the positive transcription elongation factor b (P-TEFb) taken away from the 7SK small nuclear RNA (snRNA) complex where it is normally stored in the cell[68-70]. Tat binds to P-TEFb composed of cyclin T1 (CycT1) and cyclin-dependent kinase 9 (CDK9), where Tat is bound directly to CycT1 but not to CDK9[69,71]. Transcription elongation by P-TEFb occurs by mediating the recruitment of transcription factors towards the promoter to form the super elongation complex. Following this, the RNA Pol II is hyperphosphorylated by CDK9 to activate elongation and greatly increases the

rate of transcription[68,72]. This process establishes a positive feedback loop as the hyperphosphorylation of RNA Pol II leads to an increase in Tat being produced which in turn can further recruit P-TEFb and leads to even higher rates of transcription.

Viral transcription will lead to full-length genomic viral RNA that will either remain unspliced or will give rise to alternatively spliced RNA species[73]. In the early phase of viral replication, most of the genomic RNA is doubly spliced and exported from the nucleus to the cytoplasm, where it is translated into Tat, Rev and Nef. Tat and Rev proteins go to the nucleus due to their nuclear localization signal. While Tat increases transcription, Rev will mediate the nuclear export of unspliced and singly spliced HIV-1 RNA. Indeed, although the cellular RNA export machinery is used to export completely spliced viral RNA, the incompletely and unspliced RNA cannot use this pathway to reach the cytoplasm. The viral protein Rev binds to the RRE RNA present within the *env* gene of unspliced and incompletely spliced RNAs and is able to bring the RNA species out of the nucleus through its nuclear export signal[74]. Rev dependent nuclear export occurs through numerous interactions with cellular export molecules such as upframeshift protein 1 (UPF1), chromosomal maintenance 1 (CRM1), DEAD-box helicase 3 (DDX3) and p62[75,76]. During the late phase of viral replication, the unspliced RNA can be used for packaging to form new virions or can be translated into Gag and Gag-Pol polyproteins while the singly spliced viral transcripts code for Vif, Vpr, Vpu and Env[77].

1.2.2.5 Virus assembly, release and maturation

Unspliced RNA destined to be packaged as genomic viral RNA or for translation into Gag and Gag-Pol is essential to form new virions. Gag and Gag-Pol polyproteins contain the structural components necessary for virus assembly. The composition of Gag includes MA, CA, NC, P6 along with the spacer peptides SP1 and SP2[31]. Synthesis of Gag-Pol instead of Gag occurs through

ribosomal frameshifting at the 3' end of the *gag* gene which results in readthrough of the *gag* stop codon[78]. The organization of the Gag and Gag-Pol polyproteins is such that the N terminals contain the MA domain, which is responsible for recruiting the polyproteins to the cell membrane. In the Gag polyprotein the NC domain, present at the C terminal, binds to the psi (Ψ) packaging signal located in the 5' UTR of unspliced genomic viral RNA to ensure dimerization and packaging of the viral RNA in budding particles. The Ψ packaging signal is made up of four stem-loop structures (SL1-SL4) and the dimerization initiation site (DIS) present within SL1 is critical for the dimerization process[79]. Zinc finger motifs in NC possess binding specificity for the genomic viral RNA. The $tRNA^{Lys3}$ used to prime reverse transcription is also captured during this dimerization event. Viral CA proteins assemble to form capsids that protect the HIV-1 core through interactions with cyclophilin A and transitions from an immature spherical shape to a mature conical capsid[80]. Another important viral component to form a complete viral particle is the envelope protein encoded by the *env* gene. This gene codes for the protein gp160 which is glycosylated by host Golgi apparatus and eventually cleaved by cellular protease Furin into gp120 and gp41[81]. The gp120-gp41 product becomes localized to the cell membrane through its trafficking of the secretory pathway where it is recruited by the MA domain of Gag to be incorporated into the lipid bilayer as a heterotrimer gp120-gp41, forming a complete viral envelope protein[82].

Budding from the host plasma membrane occurs by the virus co-opting elements of the endosomal sorting complexes required for transport (ESCRT) machinery. The P6 domain of Gag is critical for this process as it recruits various cellular elements of the ESCRT machinery[83]. Once an immature virion buds from the host cell membrane, viral PR mediates maturation by proteolytic cleavage of the Gag and Gag-Pol polyproteins. PR consists of a homodimer and it auto-catalyzes its own release from two Gag-Pol polyproteins[84]. It then goes on to act on specific protease

cleavage sites present between each domain of Gag and Gag-Pol to release each individual protein present within the polyproteins.

1.2.3 HIV accessory proteins

The HIV proteins Vif, Vpr, Vpu and Nef consist of a group of viral proteins which do not possess enzymatic activities but are indispensable for viral replication by either counteracting host restriction factors or downregulating specific cell surface receptors[85]. The viral protein Vif is necessary to induce ubiquitination of the cellular restriction factor Apolipoprotein B mRNA-editing enzyme catalytic polypeptide like 3G (APOBEC3G) to cause its proteosomal degradation[86]. APOBEC3G restricts viral replication by being packaged into nascent viral particles where it induces hypermutations in the viral cDNA during reverse transcription in a newly infected cell by deamination of cytosine residues which results in uracil residues[87]. The expression of Vif is therefore an evolutionary response to counteract the actions of APOBEC3G. Vpr is a multifunctional protein whose complete role in viral infection is not fully understood, but various functions have been attributed to this protein. These include acting as regulator of reverse transcription and nuclear import[88], promoting the degradation of APOBEC3G[89], repressing the actions of an unidentified macrophage restriction factor[90], as well as serving as an adaptor of various cellular proteins[91]. Regardless of its overall role, Vpr has been shown to be necessary for viral fitness as defects in the protein negatively affect disease progression[85]. The accessory protein Vpu is unique to HIV-1 and counteracts the host restriction factor known as tetherin which is present in the cell plasma membrane where it blocks virus release[87]. Binding of Vpu to the transmembrane domain of tetherin leads to proteosomal degradation of the restriction factor[92]. It is important to note that while Vpu is only found in HIV-1, HIV-2 Env protein can also counteract host tetherin[93]. Another important function of HIV-1 Vpu is the downregulation of the cell surface

molecules CD4 and intercellular adhesion molecule 1 (ICAM-1) by ubiquitin-mediated proteolysis which aids in the evasion of the immune response[87]. While Vpu is present in HIV-1, but not HIV-2, Vpx is an accessory protein unique to HIV-2. This protein counteracts the cellular restriction factor sterile alpha motif and HD domain 1 (SAMHD1) found in non-dividing myeloid cells[94]. This restriction factor possesses exonuclease activity to restrict viral replication by cleaving the viral RNA genome as well as being able to reduce dNTP pool concentrations[95,96]. As with other accessory proteins, HIV-2 Vpx causes the ubiquitination of the restriction factor target which leads to proteasomal degradation[97]. Nef is yet another multifunctional accessory protein which promotes HIV infection through actions that include down-regulation of cell-surface CD4 and major histocompatibility complex class I (MHC-I) molecules[98,99], as well as inhibiting the incorporation of serine incorporator 3 & 5 (SERINC3 & SERINC5) into new virions[100].

1.3 Clinical aspects of HIV-1 infection

1.3.1 Acute infection with the HIV-1 founder virus and progression to AIDS

The most common routes of HIV-1 transmission are by intravenous injection, mostly by drug users, and by sexual intercourse, but vertical transmission can occur during labor or from breastfeeding[101]. While the transmission probability for heterosexual coital acts is extremely low, heterosexual transmission is the largest contributor to new HIV-1 infections. Transmission rates for all routes are variable based on factors such as the presence of genital ulcers, male circumcision and high viral loads during specific stages of HIV-1 infection[101]. Noticeably, viral load can be effectively decreased by administration of ART, which renders the risk of HIV-1 transmission close to zero if the viral load is suppressed to undetectable levels[102-104]. Consequently, administration of both pre-exposure prophylaxis (PrEP) and post-exposure prophylaxis (PEP) have proven to be highly effective methods to prevent HIV-1 infection[105,106].

HIV-1 transmission consists of a bottleneck where most of the diverse viral population present within the blood of an infected individual cannot go on to infect another individual at the transmission site. This bottleneck occurs in the presence of multiple selection pressures which cause only specific HIV-1 strains, called founder viruses, to be fit in establishing a primary infection[107]. Selection pressures are present in the transmitter as well as the recipient and include physical barriers at the mucosa of infection sites, the interferon (IFN) response from the innate immune system, and neutralizing antibodies present in the transmission fluid[108,109]. Positive selection pressures also exist where the recruitment of dendritic cells at the infection site aids in the establishment of infection as these cells facilitate viral transmission to T cells[110]. These pressures lead to the selection of founder viruses which are highly efficient in crossing mucosal barriers by being resistant to the host IFN response. Founder viruses also contain a high level of envelope protein incorporated in their membrane, they possess capabilities for antibody neutralization and their main tropism is for the CCR5 coreceptor[107,108].

Regardless of the site of transmission, surveilling $CD4^+$ T cells are initially infected with the assistance of local dendritic cells[111]. The virus is then transported to nearby lymph nodes as well as to gut-associated lymphoid tissues (GALT) where the abundance of $CD4^+$ T cells present at these anatomical locations leads to systemic infection[112]. Migration of the virus to these regions causes robust killing and depletion of $CD4^+$ T cells along with a rapid expansion of HIV-1 which can often cause flu-like symptoms in individuals during this early stage of the infection. The first 7 to 21 days are known as the eclipse phase where the virus cannot be detected in the plasma. Initial detection of virus is only possible by nucleic acid amplification methods as other assays to detect viral load require higher concentrations of HIV-1 RNA[112]. After the eclipse phase, there is a sharp rise in plasma levels of viral RNA which coincides with a sharp decrease in plasma levels

of circulating $CD4^+$ T cells[113]. The final defining feature of the acute infection is the appearance of antibodies and amplification of antiviral $CD8^+$ T cells which concurs with a decline in the peak of viremia[114]. Despite the ability of $CD8^+$ T cells to be able to kill infected $CD4^+$ T cells and suppress viral replication, $CD4^+$ T cell counts will gradually decline while viremia will increase to ultimately cause the development of AIDS in the absence of treatment.

Clinical manifestations in the first 2-4 weeks of HIV-1 infection are initially nonspecific and mild. Common symptoms during acute infection are fever, weight loss, headache, lymphadenopathy, fatigue, pharyngitis, rash and myalgia as well as arthralgia[115]. Skin conditions including seborrheic dermatitis and folliculitis are exceedingly common as an initial symptom and can worsen during progression towards AIDS. These variable symptoms often disappear once the adaptative immune system is activated, although failure of the immune system to completely clear the virus will eventually end the asymptomatic stage. An increase in recurrent viral and bacterial infections are common when $CD4^+$ T cell counts drop to 200-500 cells/μL. Individuals with $CD4^+$ counts of 50-200 cells/μL are considered to have AIDS and are susceptible to opportunistic infections as well as to the development of various cancers which eventually leads to death[115,116].

It is worth noting that some individuals remain asymptomatic for a considerable portion of their lives after acute HIV-1 infection. These individuals are known as long-term nonprogressors (LTNPs) and do not require administration of cART to maintain high $CD4^+$ counts. While rare, estimates on the rates of occurrence of LTNPs is difficult to establish as there is no standardized definition of LTNPs as well as it being possible to lose one's LTNP status[117]. On the other hand, a subset of LTNPs known as elite controllers present a greater viral control in the absence of treatment where viral loads often remain at undetectable levels. Such individuals are extremely rare with one study identifying a total of 0.55% elite controllers from a sample size of 4586

individuals[118]. Some elite controllers have achieved such strong viral control that their viral reservoirs are transcriptionally silent and contain no intact functional virus, demonstrating a natural cure in these cases[119,120]. Unfortunately, not all elite controllers are protected against clinical progression. A minority of these individuals have been found to progress to AIDS with extremely low CD4 cell counts even though viremia was maintained at extremely low levels[117].

1.3.2 Establishment of latency and the viral reservoir

The hallmark process of integration of viral cDNA during HIV-1 infection can lead to latently infected cells which harbor integrated virus possessing extremely low rates of transcription. It is the inability of the immune system to completely clear the virus over the course of the infection which allows for the establishment of these latent viruses which form the viral reservoir. The consequence is such that undetectable levels of viremia either by strong immunological control or effective cART treatment will not be maintained if the immune system collapses or treatment is interrupted. Activated $CD4^+$ T cells produce infectious HIV-1 particles whereas resting $CD4^+$ T cells harbor the transcriptionally silent integrated provirus[121]. Since infection of resting $CD4^+$ T cells is generally inefficient and does not result in integration of viral cDNA, latently infected cells are likely the outcome of activated $CD4^+$ T cells which become infected and then revert to a resting state[122]. The cellular reservoir of latent provirus includes naïve $CD4^+$ T cells[123], stem cell-like memory $CD4^+$ T cells[124], central memory $CD4^+$ T cells[125] and effector memory $CD4^+$ T cells[126], with central memory $CD4^+$ T cells containing the majority of viral genomes[125]. As the cells that form the reservoir possess a half-life of approximately 44 months, viral persistence is therefore a consequence of the maintenance of resting $CD4^+$ T cells and the reservoir is not expected to be eliminated during the lifetime of an infected individual[127]. Additionally, recent evidence suggests that the reservoir is dynamic and is maintained by clonal

expansion of cells infected with provirus as identical proviral integration sites have been identified in different cells, a result of genomic duplication during mitosis from cellular proliferation[128]. This maintenance of $CD4^+$ T lymphocytes is driven either by integration near genes that control cell division[129], by homeostatic proliferation from host cytokines[130], or by antigenic stimulation from co-infections with other viruses such as cytomegalovirus and Epstein-Barr virus[131]. Studies have elucidated a potential link for preferential integration of provirus in genes that control cellular proliferation, such as BTB and CNC homolog 2 (BACH2), signal transducer and activator of transcription 5B (STAT5B) and myocardin-like 2 (MKL2), as well as within oncogenes, resulting in an upregulation of clonal expansion in these infected cells[128,129,132]. On the other hand, basal levels of homeostatic proliferation of resting $CD4^+$ T cells are driven by interleukin 7 (IL-7) and this also plays a role towards the persistence of the viral reservoir[130]. Compared to the small contribution of homeostatic proliferation towards the expansion of latently infected cells, antigen stimulation of reservoir cells causes a massive proliferation of these activated cells. The antigenic specificity of latently infected $CD4^+$ T cells has been shown to be made up of HIV-1 specific cells as well as other more common antigens derived from cytomegalovirus, Epstein-Barr virus and influenza virus[131,133,134].

Although $CD4^+$ T cells are the primary targets of HIV-1 and make up the bulk of the viral reservoir, myeloid cells can also be infected by HIV-1 and contribute to the establishment of the reservoir[135]. This subset of the cellular reservoir becomes increasingly important during disease progression as $CD4^+$ T cells are depleted and infected macrophages become the primary cell type comprising the reservoir[136]. However, as the depletion of $CD4^+$ T cells does not typically occur in individuals taking cART, the relevance of myeloid cells towards the HIV-1 reservoir in this context is poorly understood. In cART-naïve HIV-1 positive individuals, macrophages quickly become

infected to contribute to the establishment of the viral reservoir and through these infected cells the virus infiltrates the central nervous system (CNS) where an anatomical reservoir is formed[137]. Maintenance of the infection within the brain has several consequences such as structural changes[138] and neuronal injury[139]. The contribution of monocytes and macrophages towards HIV-1 infection in HIV-1 positive individuals taking cART is unclear. Studies that seek to detect the presence of HIV-1 in these cells are often not able to detect integrated viral DNA in the monocytes and macrophages of all donor samples[140,141]. Overall, monocyte and macrophage reservoirs likely persist in some HIV-1 positive individuals taking cART and differences in their relevance towards global infection may be influenced by treatment effectiveness[135].

The establishment of viral reservoirs is multifactorial, where factors that impede viral transcription such as nucleosomal structure, epigenetic events within the HIV-1 promoter and specific transcriptional inhibitors encourage latency rather than a productive infection of a cell[142]. A particular threshold of transcription typically must be reached for complete viral replication to take place. This threshold is determined by the synthesis of Tat, where a sufficient basal transcription state leading to its synthesis causes recruitment of P-TEFb and an increase in the elongation efficiency of the initiated RNA polymerase during HIV transcription which is necessary to form new viral particles[142]. However, recent evidence has suggested that latency may be the result of a series of defects in transcription elongation, polyadenylation and multiple splicing rather than solely being a consequence of defects in transcription elongation[143]. Certain cellular factors have been found to promote latency in resting $CD4^+$ T cells, such as the long non coding RNA (lncRNA) non-coding repressor of NFAT (NRON) which induces proteasomal degradation of Tat[144] and B cell leukemia 11b (BCL11B) which represses the transactivation of Tat upon the viral LTR[145]. The viral integration site affects the basal transcription state as insertion of viral cDNA in

a repressive chromatin region will hinder transcription[146,147]. Interestingly, insertion of the provirus in either orientation relative to the host gene can impair transcription. In the case of identical orientations, the transcriptional machinery from neighboring genes can lead to steric hindrance and the removal of transcription factors from the HIV-1 promoter[142]. Conversely in the case of opposite orientations, collisions between the RNA Pol II mediating viral transcription and the transcription machinery of neighboring genes can occur during elongation which results in truncated transcripts from premature termination[142]. A failure to recruit sufficient NF-κB to the promoter can also lead to transcription rates not reaching the threshold needed for Tat synthesis. Specific levels of NF-κB needed to induce Tat expression is dependent on the genomic location of the provirus[148]. Post-translational modifications of histones as well as epigenetic modifications of both the surrounding chromatin and the proviral promoter contribute to maintaining latency since these events can decrease the accessibility of transcription factors due to chromatin condensation[142]. Deacetylation and methylation events of histones negatively affect transcription rates and are mediated by histone deacetylases (HDACs) and histone methyltransferases (HMTs), respectively[149]. Acetylation of lysine residues of histones weakens the electrostatic interactions between histones and DNA to make the chromatin more accessible while amino acid residues of histones which are methylated serve as attachment regions for chromatin remodeling agents[142]. Both of these processes have been shown to be tightly regulated within latently infected cells where histones are both hypoacetylated and hypermethylated[150-152]. DNA methylation consists of another factor which interferes with transcription and is mediated by DNA methyltransferases (DNMTs). Methylation is concentrated on CpG islands where transcription is impaired from steric hindrance which prevents the binding of transcription factors or by causing the eventual recruitment of chromatin remodeling agents[142]. As to the relevance of DNA methylation in the maintenance of

HIV-1 latency, two CpG islands flanking the transcriptional start site of the viral promoter have been found to be hypermethylated in latently infected cells and demethylation agents cause viral reactivation in these same cells[153].

1.3.3 Current antiretroviral strategies

Various antiretroviral drugs have been designed to inhibit HIV-1 replication by targeting specific steps of the viral replication cycle. The first antiretroviral drug approved to treat HIV-1 infection was a nucleoside reverse transcriptase inhibitor (NRTI) called zidovudine (AZT). Although this drug was effective at inhibiting viral replication, it was also associated with unwanted side effects. AZT and other NRTIs function as chain terminators where they compete with other deoxynucleotide triphosphates (dNTPs) present within the cell for binding to the catalytic site of HIV-1 RT[154]. The next antiretroviral drugs to be developed were nonnucleoside reverse transcriptase inhibitors (NNRTI). NNRTIs can inhibit the action of HIV-1 RT by binding to sites other than the catalytic site[155]. While the two types of reverse transcriptase inhibitors could initially control viremia within infected individuals, this effect was not sustainable. It therefore became evident in the late 1990s that a combination of drugs was required to prevent resistance[156,157]. The development of an entirely new class of antiretrovirals known as protease inhibitors, which inhibit viral PR through binding to the active site formed from the two PR subunits and acting as uncleavable peptides[155], was therefore used along with the previously developed NRTIs and NNRTIs. Combinations of these drugs were named highly active antiretroviral therapy (HAART) or cART and a few other drug classes have since been developed, such as fusion inhibitors, CCR5 antagonists, integrase inhibitors, post-attachment inhibitors and attachment inhibitors[158]. The most recent research efforts have led to the development of long-acting antiretrovirals such as carbotegravir which can be administered bi-monthly and are thus

attractive treatment options to avoid pill fatigue from daily cART administration[159]. A list of antiretroviral drugs approved by the US Food and Drug Administration (FDA) to treat HIV-1 infection is shown in Table 1.2.

Modern cART regimens typically include three drugs consisting of two NRTIs along with either an IN inhibitor, an NNRTI or a PR inhibitor[160]. Despite the advancements made in HIV-1 treatment, the modern era has seen a rise in resistant strains. This has led to the need to screen newly infected individuals for resistance markers by viral genotyping to ensure a personalized drug cocktail is administered to maintain low levels of viremia[161]. The emergence of resistance even in the presence of cART is a consequence of the high mutation rate of the virus. A substantial source of resistant HIV-1 strains occur in lower-income areas where viral genotyping and viral load monitoring is often unavailable, meaning that an infected individual will be unaware whether an escape mutant arises and a switch to a new antiretroviral combination will therefore never occur[161]. As we have learned through vaccine inequality during the COVID-19 pandemic, the emergence of novel viral strains is a global issue, as regardless of the geographical location in which they arise, these strains travel the globe to infect individuals indiscriminately of their socio-economic background.

Administration of effective cART to reach low levels of viremia is not only important to prevent an infected individual from progressing clinically to AIDS but is also paramount to avoid transmission. Studies have shown that undetectable levels of HIV-1 RNA in HIV-1 positive partners eliminate the risk for sexual transmission in heterosexual couples[104,162]. Data from these studies has pushed the narrative of undetectable equals untransmittable (U=U). While cART requires continuous administration of drugs, there exists a minority of individuals who do not show viral rebound after treatment interruption[163]. Post-treatment controllers typically begin cART at

extremely early stages of HIV-1 infection, and their ability to maintain viremia at low levels after cessation of therapy is thought to be due to poor establishment of the viral reservoir[164,165]. It is also possible to prevent HIV-1 infection through administration of PrEP or PEP to uninfected individuals engaging in behaviors that pose a risk of HIV-1 infection such as sharing injection equipment[166]. Typical PrEP regimens consist of oral antiretroviral drugs taken 2-24 hours prior to exposure followed by daily dosing for 2 days while PEP should be initiated immediately after exposure[160].

Table 1.2 List of antiretrovirals recommended to treat HIV-1 infection in the United States.

Drug Class	Generic drug name	FDA approval year
NRTI	Abacavir	1998
	Emtricitabine	2003
	Lamivudine	1995
	Tenofovir disoproxil fumerate	2001
	Zidovudine	1987
NNRTI	Doravirine	2018
	Efavirenz	1998
	Etravirine	2008
	Nevirapine	1996
	Nevirapine XR	2011
	Rilpivirine	2011
Protease inhibitors	Atazanavir	2003
	Darunavir	2006
	Fosamprenavir	2003
	Ritonavir	1996
	Saquinavir	1995
	Tipranavir	2005
Fusion inhibitors	Enfuvirtide	2003
CCR5 antagonists	Maraviroc	2007
Integrase inhibitors	Carbotegravir	2021
	Dolutegravir	2013
	Raltegravir	2007
	Raltegravir (ISENTRESS HD)	2017
Post-attachment inhibitors	Ibalizumab-uiyk	2018
Attachment inhibitors	Fostemsavir	2020

1.3.4 HIV-1 cure strategies

While cART has been successful in increasing the lifespan and quality of life of PLWH, these drugs are not curative and require chronic administration as treatment interruption results in viral rebound from the HIV-1 reservoir. The principal barrier to achieving an HIV-1 cure is the presence of this reservoir as these viral sanctuaries are extremely effective at allowing evasion of the immune response or to therapies targeting the virus[167]. The removal of all traces of the virus, known as a sterilizing cure, is therefore extremely difficult to achieve due to this characteristic of the latent virus. However, long term remission of viremia in the absence of cART, known as a functional cure, allows for a control of HIV-1 infection to prevent disease progression without the need to eliminate the viral reservoir[65]. Natural functional cures have been documented in at least two instances where infected individuals were able to naturally suppress viral loads so successfully that although integrated proviral DNA could be detected, intact functional virus was absent[119,120]. Additionally, functional cures have been observed in post-treatment controllers who began cART extremely early in infection[163,168,169], however this control of viral infection in the absence of cART has been shown to be lost in one case[170].

Curative strategies have been possible through hematopoietic stem cell (HSC) transplants from HIV-1 resistant donors. Such transplants have resulted in long-term viral remission without the need for daily medication, as well as the absence of detectable virus in bodily fluids or tissues. This procedure is only considered in HIV-1 positive individuals that are also suffering from hematological malignancies and involves an HSC transplant from HIV-1 resistant donors who possess a homozygous 32-bp deletion in the CCR5 gene (CCR5Δ32/Δ32)[171]. Mutations in this cellular gene impair viral entry since CCR5 is one of the coreceptors of HIV-1. This limitation on viral entry can confer immunity to HIV-1 positive patients receiving the transplant and the

CCR5Δ32/Δ32 mutation is not known to be associated with any negative consequences[172,173]. Successful remission of HIV-1 infection in the absence of cART resulting from an HSC transplant has occurred in three instances[171,174,175]. In the first case, Timothy Brown (also known as the "Berlin Patient") received an HSC transplant from a CCR5Δ32/Δ32 donor in 2007, although a second transplant was needed because of a relapse in leukemia[176]. Interruption of cART occurred immediately after the transplant and virus was undetectable in his tissues or bodily fluids up until his death in September 2020 caused by a rebound in his leukemia. In the second case, Adam Castillejo (also known as the "London Patient") received an HSC transplant from a resistant donor in 2016 and the virus has since been undetectable even while discontinuing cART 16 months after the transplant[177]. The "Düsseldorf patient" represents a third case of long-term viral remission from an HSC transplant that was confirmed recently[175]. This patient received an HSC transplant from a CCR5Δ32/Δ32 resistant donor in 2013 and viral rebound has not been detected since cART interruption in 2018, although traces of the viral reservoir could be detected by in depth analyses. Recently, it has been announced that a fourth instance of HIV-1 remission after HSC transplant from a CCR5Δ32/Δ32 resistant donor has occurred in the "City of Hope" patient[178]. An additional instance of an HIV-1 cure has been successful in a woman after a CCR5Δ32/Δ32 cord blood and CD34-selected haploidentical stem cell transplant (haplo-cord SCT)[179]. This patient achieved a 100% CCR5Δ32/Δ32 chimerism following the procedure and has been reported to be in HIV-1 remission for 14 months without cART. While these cases are encouraging, HSC transplant for all globally infected individuals is currently unfeasible due to a substantial risk of graft versus host disease (GVHD) and the fact that only a finite number of individuals harboring the CCR5Δ32/Δ32 genotype exist[180,181]. Although haplo-cord SCT does not possess as severe inherent risks as HSC transplants, its success in being curative towards HIV-1 infection has only occurred in one instance

so far. Therefore, the development of novel strategies to cure HIV-1 infection are still needed. Current research into cure therapies focus on sterilizing and functional cures through the "shock and kill" strategies, "block and lock" strategies, immunotherapies or gene therapy.

1.3.4.1 "Shock and kill" strategies

The goal of the "shock and kill" strategy is to achieve a sterilizing cure, where all presence of the virus is removed from the body. This could be done by reactivating latently dormant virus in the viral reservoir to elicit their destruction by the host immune response. Reactivation of dormant virus is possible by using latency reversing agents (LRAs) which act through various mechanisms such as changing the epigenetic landscape to promote cellular transcription by inhibiting the histone post-translational modifiers HDACs and HMTs as well as inhibiting the methylation of the CpG islands which flank the HIV-1 promoter[150,182,183]. LRAs can also induce cellular signaling pathways to promote transcription through the activation of transcription factors. PKC agonists consists of LRAs which are able to activate HIV-1 transcription in this manner and three families of PKC agonists have been studied in this context[182]. These LRAs function by activating the PKC signaling pathway which results in the stimulation of the NF-κB pathway and consequently an upregulation of cellular transcription[184]. Other LRAs that can activate transcription factors such as NF-κB to cause activation of viral gene expression include toll like receptors (TLRs) agonists, which activate cellular TLRs to cause a signaling cascade leading to increased production of transcription factors[185], and Akt pathway activators which facilitate nuclear translocation of NF-κB[186]. Bromodomain and extra-terminal motif protein inhibitors (BETis) make up another class of LRAs which function through the induction of cellular pathways. They activate viral transcription by playing on the interactions associated to P-TEFb where the LRA causes the dissociation of P-TEFb from the cellular bromodomain-containing protein 4

(Brd4), thus increasing the accessibility of P-TEFb for binding of Tat and the formation of the super elongation complex[187]. Additionally, BETis can promote P-TEFb towards its active configuration by inducing its dissociation from the 7SK snRNA[188]. Other LRAs inhibit suppressor molecules of the NF-κB signaling pathway or suppressor molecules of the JAK-STAT signaling pathway[182].

A significant issue with "shock and kill" relates to toxicity, as most LRAs are non-specific and therefore upregulate transcription of proviral genes along with cellular genes for which there is typically a tight regulation of expression[189]. Furthermore, concerns have also been raised as to the harmful pathogenic effects which could arise from reactivation of the latent reservoir[190]. LRAs have been shown to effectively reactivate virus in latently infected cells both *in vitro* and *in vivo*[191-193]. However, reduction in the size of the HIV-1 reservoir does not occur despite this reactivation, emphasizing that the "shock" part of the therapy is effective but the "kill" part is not[194]. The main obstacle preventing "shock and kill" strategies from effectively eliminating virus in the body most likely relates to the complexity and heterogeneity of the viral reservoir. This complexity is demonstrated by the reservoir being spread out across many different anatomical sites such as the brain[195], GALT[196], adipose tissue[197], etc. The reservoir is also comprised of various cellular subsets of the immune system which includes resting $CD4^+$ T cells, regulatory T cells, macrophages and astrocytes[198,199]. It is therefore probable that LRAs fail to reactivate dormant virus in all cell types or are not able to reach each anatomic location of the reservoir.

Recent advancements in the "shock and kill" strategy have been focused on enhancing the clearance of reactivated reservoir cells to successfully reduce the reservoir size as well as developing gene specific transcriptional activators to decrease toxicity. The most common suggestion to boost the killing of reactivated cells is the introduction of broadly neutralizing

antibodies (bNABs) or the use of vaccines to increase the response of cytotoxic T cells towards HIV-1 infected cells[200,201]. Gene specific activation of transcription has been possible by exploiting the properties of clustered regularly interspaced short palindromic repeats (CRISPR) and the CRISPR associated protein 9 (Cas9). In this case, a deficient Cas9 is fused to a transcription activator and is used in combination with a guide RNA (gRNA) to target key features within the proviral DNA sequence that influence viral transcription[202]. Usage of such technology therefore hones LRAs uniquely to the provirus and avoids toxicity originating from the non-specific transcriptional activation of cellular genes.

1.3.4.2 "Block and lock" strategies

The goal of "block and lock" strategies is to achieve a functional cure, where the virus is still present in the body but is unable to reactivate and cause disease even in the absence of cART. The method used to accomplish this can be considered as the reverse of "shock and kill" strategies. Whereas "shock and kill" sets out to reactivate dormant virus to eliminate the reservoir, "block and lock" seeks to promote latency to reach a state termed "deep latency" where the latent provirus is disabled and cannot be reactivated[203,204]. Just as epigenetic changes can promote activation of transcription, deacetylation and methylation of histones can encourage latency. One method to achieve transcriptional silencing of the reservoir would be to induce transcriptional gene silencing (TGS) by using dicer-substrate RNAs complementary to DNA segments of HIV-1 provirus[205]. Targeting of DNA instead of RNA typically elicits the recruitment of Argonaute 1 (Ago1), but possibly can cause recruitment of Ago2, which will form the RNA-induced transcriptional silencing (RITS) complex[206-208]. Once the RITS complex hybridizes to a DNA region by its gRNA, Ago1 or Ago2 can recruit HDACs and HMTs to adjacent nucleosomes which catalyzes epigenetic changes leading to transcriptional repression[209-211]. Target sites of the provirus that have led to

successful gene silencing include important regulatory elements present within the proviral LTR like the NF-κB binding sites[206,212,213]. In fact, the effect of transcriptional repression from RITS has been shown to be so potent that it can resist activation of latent provirus by LRAs[212].

Another strategy to promote transcriptional silencing is by using compounds that can interfere with Tat functionality. Such Tat inhibitors prevent the transcriptional positive-feedback loop elicited by Tat. Didehydro-cortistatin A inhibits Tat transactivation by binding to TAR, making it unavailable for binding by Tat[214]. This compound has been shown to be effective to prevent activation of the viral reservoir *in vitro* and could delay viral rebound by 9 days *in vivo*[215,216]. Inhibition of Tat is also possible through a compound called sudemycin which impedes the functionality of the cellular protein splicing factor 3B subunit 1 (SF3B1)[217]. As SF3B1 is required for Tat-mediated HIV-1 transcription, its inhibition will in turn interfere with Tat functionality. Sudemycin also inhibits RNA splicing, which is a necessary step for Tat production, and thus interferes with Tat production. Administration of sudemycin represses transcription in latently infected cells and reactivation does not occur for at least 72 hours in the presence of LRAs[217]. Another compound able to inhibit the activity of Tat is HT1, which contains domains of hexamethylene bisacetamide inducible proteins 1 (HEXIM1) as well as domains of Tat and acts as trans-dominant negative (TDN) peptide. HT1 competes with Tat for binding to P-TEFb, making the complex unavailable to promote viral transcription by the formation of the super elongation complex[218]. Similarly, a TDN molecule of CycT1 has been developed called CycT1-U7 which has been shown *in vitro* to decrease transactivation of the viral promoter as well as leading to proteasomal degradation of Tat[219]. Two additional molecules that inhibit the actions of Tat include a Tat mutant called Tatnullbasic which leads to a reduction of the recruitment of the RNA Pol II onto the viral promoter and a protein called triptolide which causes proteasomal degradation of

Tat[220,221]. Efforts have also been made to impair the transcriptional activation of Tat by chemical reactions resulting in cycloaddition to the TAR RNA. Such a process has been shown to abrogate the interaction between Tat and TAR *in vitro*[222].

Although large efforts have been made to promote HIV-1 latency by targeting elements that disrupt Tat functionality, it is also possible to promote latency by interfering with cellular factors to impair specific stages of transcription. Spirolactone can inhibit transcription initiation with negative consequences on viral transcription[223]. Transcription elongation can be inhibited by several molecules which include CDK9 inhibitors, curaxin 100 and Q308 to promote viral latency[224-226]. Finally, interfering with broad cellular signaling such as the NF-κB or the mammalian target of rapamycin (mTOR) pathways has been shown to repress viral transcription and promote latency[227,228].

1.3.4.3 Immunotherapies

Over the course of untreated HIV-1 infection, viremia peaks during the acute phase of the infection and subsequently declines after the appearance of HIV-1 specific $CD8^+$ T cells[114]. Although the immune response eventually succumbs to the infection where $CD4^+$ T cells are depleted and viremia rebounds to high levels, the actions of the cytotoxic T lymphocytes (CTLs) can clearly control the infection for a period of time. Therapies seeking to assist the actions of CTLs to make them more effective at controlling HIV-1 infection, known as immunotherapies, may offer an approach to achieving a sterilizing cure. In the context of virally suppressed individuals from effective cART, there is virtually no viral replication occurring which makes the already rare and anatomically distributed latently infected cells extremely difficult to be recognized and eliminated by CTLs[229]. Immunotherapies may therefore be most effective when used in the context of “shock and kill” as priming the immune system to be reactive towards HIV-

1 could facilitate clearance of the virus by CTLs after activation of the reservoir through LRA administration. Methods to increase the ability of CTLs to kill HIV-1 infected cells are comprised of therapeutic vaccine approaches which attempt to elicit the development of bNABs and to cause clonal expansion of HIV-1 specific $CD8^+$ T cells, but passive administration of bNABs can also be used as an immunotherapy[229]. Such passive administration of bNABs can lead to transient viremic control after transfusion of one or more anti-HIV-1 antibodies to HIV-1 infected individuals.

Vaccines against HIV-1 have largely been ineffective in efficacy trials due to various technical challenges. The greatest hurdle is the sequence diversity of the virus and its ability to rapidly mutate under the selective pressure created by HIV-1 epitope recognition of CTLs and neutralizing antibodies after vaccine therapy[230]. The viral reservoir also provides a sanctuary for dormant virus and the glycans associated to the viral Env proteins facilitate evasion of the immune response[231]. The first two attempts at developing an HIV-1 vaccine were called VAX003 and VAX004, both consisted of a recombinant Env protein including two gp120 antigens and neither vaccine prevented HIV-1 infection nor affected viral load in phase 3 efficacy trials[232,233]. Following the results from VAX003 and VAX004, vaccine strategies shifted to using vector-based vaccines to elicit a cellular immune response rather than an antibody response. The STEP and Phambili vaccine trials were initiated using this strategy, however initial results from the STEP trial showed that vaccinees were acquiring HIV-1 at an increased rate which prompted both trials to be terminated early[234,235]. Another vaccine trial using the same adenoviral vector as the STEP and Phambili trials was performed called HVTN 505 which included Env, Gag, Pol and Nef immunogens but this trial was halted due to lack of efficacy[236]. The only vaccine trial to offer some protection so far against HIV-1 infection was the RV144 trial which used a canarypox vector with

two gp120 antigens and demonstrated a 31% efficacy at preventing infection after 3.5 years[230,237]. Follow up trials of RV144 have been conducted in an attempted to boost the immune response over time in vaccinated individuals by introducing additional vaccine components which resulted in increased levels of HIV-1 specific IgG as well as antibodies with features of bNABs[238,239]. Based on the success of RV144, other vaccine trials have used the same strategy of employing a pox-protein approach. One trial called HVTN 100 showed HIV-1 specific immune activation in early phase trials while another called HVTN 702 failed to decrease infection rates and was subsequently halted[230,240]. The most recent attempts in developing therapeutic vaccines against HIV-1 are utilizing polyvalent mosaic antigens to cover the global viral diversity[241], as well as exploring the use of DNA priming as a vaccine option[242]. Based on the success of the COVID-19 mRNA vaccines, the development of HIV-1 specific mRNA vaccines is also being pursued with promising results in animal models[243]. So far, two phase I clinical trials have been launched to evaluate the safety of mRNA HIV-1 vaccines.

Although the goal of vaccines is to elicit a strong activation of HIV-1 specific CTLs and/or the development of bNABs, the passive administration of bNABs through single-cell-based antibody cloning methods has been explored to control HIV-1 infection[229]. Early use of HIV-1 specific neutralizing antibodies demonstrated little success[244,245], however the use of bNABs which can neutralize multiple viral strains has shown promise[246,247]. Specifically, clinical trials using bNABs were able to decrease viral loads in participants but eventually led to the emergence of viral strains able to resist the neutralizing actions of the antibodies. The latest innovations in antibody therapeutics for HIV-1 infection have led to the development of engineered antibodies which can either bind to two separate antigens or recognize a single antigen while simultaneously recognizing and causing the recruitment of effector cells to cause target killing of infected cells[248].

It is important to mention that both vaccine and bNABs administration have been used as the "kill" component of the "shock and kill" strategy, either in animal or human models, with signs of some success such as reductions in total viral DNA reservoir and delaying viral rebound[201,249,250].

1.3.4.4 Gene therapy

Gene therapy strategies are similar to "block and lock" in that they also seek to achieve a functional cure rather than a sterilizing one. While initially proposed in 1988 as intracellular immunization, using gene therapy to control HIV-1 infection is extremely promising based on the "Berlin and London patients", as performing an autologous transplant using patient-derived HSCs that are genetically modified to resist HIV-1 replication could recapitulate these successes while avoiding the risk of GVHD[251,252]. Gene therapy strategies include gene editing technologies to alter the CCR5 gene in patient cells to mimic the procedure used for Timothy Brown and Adam Castillejo[253-256]. However, as HIV-1 possesses tropism for a second coreceptor known as CXCR4, the presence of CXCR4 using viruses in circulation could cause a rebound in viremia following transplantation of CCR5Δ32/Δ32 HSCs. This event occurred with the "Essen patient", where the emergence of CXCR4 using virus prior to the HSC transplant caused a rebound in viremia three weeks after the transplant[257,258]. Since editing of the CCR5 cannot be completely relied upon to halt HIV-1 infection, an arsenal of antiviral genes has also been developed for use in HIV-1 gene therapy. These genes can be integrated into the chromosome of an infected individual's cells using lentiviral vectors. Anti-HIV-1 molecules available for use in gene therapy include proteins as well as various types of non-coding RNA molecules such as ribozymes, aptamers & decoys, short hairpin RNAs (shRNAs) and U1 interference (U1i) RNAs. A comprehensive list of clinical trials performed to generate HIV-1 resistant cells by gene therapy is detailed in Table 1.3.

All current gene editing methods rely on nucleases creating double-stranded breaks (DSBs) by binding to a specific target sequence defined through complexing of a DNA-binding domain to the nucleases. This event is followed by rejoining of the broken ends by non-homologous end joining (NHEJ) or homology-directed repair (HDR) from the cellular DNA repair machinery. NHEJ results in insertions or deletions (indels) to generate knockouts while HDR can create precise edits through introduction of a template during the repair process[259,260]. One class of nucleases used for gene editing are zinc finger nucleases (ZFNs). ZFNs possess a programmable DNA-binding domain comprised of a multitude of zinc finger proteins arranged in a series, and fused to a nuclease domain of the bacterial Fok1 restriction enzyme[261]. Each zinc finger can recognize a 3-base pair sequence and thus an appropriate amount of zinc fingers must be included in the molecule to recognize a given target site length[262]. Since the DNA-binding domain is modifiable, ZFNs can be used to introduce DSBs at specific genomic locations which will be rejoined by NHEJ or HDR[263]. ZFNs have been successfully used to disrupt the expression of CCR5[264-266]. In one clinical trial enrolling 12 HIV-1 positive individuals, gene-modified cells possessed a selective survival advantage when antiretroviral therapy was interrupted following infusion of $CD4^+$ T cells. Unfortunately, viral rebound could not be prevented for any of the participants who underwent treatment interruption[253].

Another class of nucleases available for gene editing are the transcription activator-like effector nucleases (TALENs). In this case, the DNA-binding motif is made up of bacterial proteins called TAL effectors (TALEs) and are again complexed to the bacterial Fok1 endonuclease to generate DSBs. TALEs also consist of a repeat of monomers arranged in a series, however these repeats each recognize a single nucleotide instead of a series of 3 nucleotides as with ZFNs[267,268]. It has been shown that TALENS are able to offer similar gene disruption capabilities compared to

ZFNs, but with much less associated cytotoxicity[269]. As with ZFNs, TALENS have successfully been used to interfere with cellular CCR5 expression[270,271].

The most recent class of nucleases used for gene editing are the CRISPR-Cas nucleases. This nuclease system makes up a part of the immune system in bacteria and archaea to provide resistance to invading exogenous genetic elements such as viruses[272]. Simply put, the system retains DNA sequences from offending genetic elements to generate gRNAs, known as CRISPR RNAs, complementary to the invading DNA. The CRISPR-Cas9, derived from *Streptococcus pyogenes*, is the most commonly used form of this system and specific gRNAs can be incorporated to direct the endonuclease to a particular target site[273-275]. The CRISPR-Cas system has been widely adopted by the scientific community for gene editing as it does not require considerable protein engineering to change the target site as is the case with ZFNs and TALENS. Use of gRNAs directed towards the CCR5 gene have resulted in its inactivation[254,276]. It is worth noting that it is possible to target elements of the provirus to inhibit viral replication and that targeting both LTRs by Cas endonucleases will result in excision of the integrated provirus[277-280].

Many anti-HIV-1 proteins have been designed as TDN mutants modeled after known proteins used in the viral replication cycle. These molecules are deficient in their function contributing to viral replication, but their binding motifs remain unaltered. As binding capabilities of TDN proteins is unaffected, they competitively bind to the binding target of wild type viral proteins where their inability to contribute to specific steps of the viral replication cycle will result in virus inhibition. Various TDN mutants based on Gag along with its proteolytically cleaved products have been designed[281-284]. Another TDN mutant modeled after Rev, called Rev M10, can bind to the RRE of unspliced viral transcripts but lacks export capabilities. It has reached clinical trials, but viral inhibition was not sufficient to pursue its use[285-287]. Other viral proteins that have

had TDN mutants modeled after them to inhibit HIV-1 replication include Tat[288], Vif[289,290] and Nef[291]. Anti-HIV-1 proteins based on restriction factors have also been developed. The restriction factor tripartite motif-containing protein 5 alpha (TRIM5α), which recognizes viral capsid proteins to cause premature uncoating and resulting in an impairment of reverse transcription, has been of particular interest as its isoform found in rhesus macaques (Trim5α_{rh}) induces resistance to HIV-1 infection in this species[292,293]. In an effort to grant this immunity to human cells, chimeric human-rhesus TRIM5α have been designed where only the 13 amino acid motif responsible for HIV-1 restriction from TRIM5α_{rh} is incorporated into the human isoform[294,295]. Furthermore, restriction of HIV-1 replication by downregulation of CCR5 coreceptor has been possible by an intracellular chemokine or "intrakine". This intrakine is specific to CCR5 protein but also possesses an endoplasmic reticulum retention signal which leads to retention of nascent CCR5 in the ER followed by degradation[296,297]. HIV-1 entry can also be impaired through the administration of a peptide called C46 which inhibits fusion of the viral and cellular membranes[298]. This peptide has been used within a clinical trial where it was found to be well tolerated after infusion of $CD4^+$ T cells expressing the inhibitor from a retroviral vector and gene marking could be detected over a 1-year follow-up[299]. Although viral inhibition has been successful using anti-HIV-1 genes coding for proteins, a greater amount of non-coding functional RNAs have been developed and will be discussed in greater detail within the following sections.

Table 1.3 Anti-HIV-1 gene therapy clinical trials.

Antiviral molecule(s)/gene editing method	Cells modified	Status	Reference
CCR5 gene disruption by zinc finger nuclease	$CD4^+$ T cell	NCT02225665, 01543152, 02388594	[253]
	HSCs	NCT02500849	[300]
CCR5 gene disruption by CRISPR/Cas9	HSCs	NCT103164135	-
Trans-dominant negative Rev protein mutant (RevM10)	$CD4^+$ T cell	Completed	[287]
	$CD4^+$ T cell	Completed	[301]
	HSCs	Completed	[302]
Trans-dominant negative Rev protein mutant	$CD4^+$ T cell	Completed	[303]
	HSCs	Completed	[304,305]
Fusion inhibitor (C46)	$CD4^+$ T cell	Completed	[299]
TAR RNA decoy, CCR5-specific ribozyme, tat/rev-specific shRNA	HSCs	NCT01961063, NCT02337985, NCT00569985	[306]
Fusion inhibitor (C46), CCR5-specific shRNA	HSCs & $CD4^+$ T cell	NCT03593187, NCT02390297, NCT01734850	-
TAR RNA decoy, chimeric TRIM5α, CCR5-specific shRNA	HSCs	NCT02797470	-
vpr/tat-specific ribozyme	$CD4^+$ T cell	Completed	[307]
	Syngeneic $CD4^+$ T cell	Completed	[308]
	HSCs	Completed	[309,310]
	HSCs	Completed	[311]
RRE RNA decoy	HSCs	Completed	[312]

1.4 HIV-1 gene therapy using non-coding functional RNAs

1.4.1 RNA interference

RNA interference (RNAi) is a cellular post-transcriptional gene regulation mechanism, and in some organisms serves as a defense mechanism against pathogenic double stranded (ds)RNA. Initially discovered in 1998 by Andrew Fire and Craig Mello, RNAi has since been co-opted by scientists as a therapeutic pathway to treat genetic or infectious diseases by directing the system towards specific genes involved in these pathologies[313-315]. Gene silencing by RNAi can be mediated by various RNA species such as small interfering RNAs (siRNAs), shRNAs and microRNAs (miRNAs)[316]. When these RNAs are present within a cell, one of the two strands will bind to a target mRNA through nucleotide complementarity which leads to either mRNA cleavage, translational repression or targeted degradation.

In mammalian cells, the endogenous gene regulation role of RNAi occurs by the expression of miRNAs and the human genome includes approximately 2300 miRNA genes[317]. In animals, miRNAs are typically not completely complementary to a given target and therefore a single miRNA will regulate the expression of many protein-coding genes[318]. However, a seed region exists between nucleotide positions 2 to 8 in the guide RNA which must be complementary to the mRNA target to allow for binding by a miRNA[319]. Biogenesis of miRNAs begins with transcription of a miRNA gene by the RNA Pol II to form a primary miRNA (pri-miRNA), although miRNAs can also be intron derived where the pri-miRNA is a result of the processing of gene introns[320,321] (Figure 1.4). These pri-miRNAs form a hairpin structure which includes a 5' cap and a polyadenylation signal at the 3' end[322]. The pri-miRNAs are then processed by the Drosha RNase III endonuclease and its cofactor DiGeorge syndrome critical region gene 8 (DGCR8) into precursor miRNAs (pre-miRNAs) of approximately 70 nucleotides consisting of

stem loop structures[323] (Figure 1.4). These pre-miRNAs are then exported to the cytoplasm by interactions with exportin-5 using the Ran-GTP/Ran-GDP mechanism[324]. Once in the cytoplasm, the pre-miRNA will associate with TAR RNA binding protein (TRBP) which is in turn associated to the RNase III endonuclease Dicer[325]. Dicer cleaves the stem loop structure of the pre-miRNA to produce a mature miRNA with two nucleotide 3' overhangs (Figure 1.4). ATP-dependent chaperones then help unwind the miRNA structure to discard the passenger strand and load the guide strand into recruited Ago2 protein, which forms the RNA induced silencing complex (RISC) when associated along with Dicer and TRBP[322]. The guide strand then directs the RISC to specific mRNA targets through hybridization from base pairing where perfect complementarity to the target causes mRNA cleavage by Ago2 but imperfect complementarity leads to translational repression or targeted mRNA degradation in processing bodies (P-bodies) where RNA de-capping and degradation enzymes are concentrated[326] (Figure 1.4).

Therapeutic approaches utilizing the RNAi pathway have employed shRNAs and synthetic siRNAs as these molecules can be engineered to be 100% complementary to a given target to trigger more specific gene silencing than miRNAs. While siRNAs can be introduced directly into the cytoplasm of the cell by endocytosis, shRNAs must first be delivered to the cell in the form of a gene to be transcribed and processed similarly to miRNAs[327]. Therefore, shRNA genes are typically included within plasmids or lentiviral vectors to be introduced into the cell[328]. Transcription of the shRNA gene will synthesize a shRNA consisting of a stem loop structure that is not processed by Drosha and DGCR8 but is instead immediately exported out of the nucleus by interactions with exportin-5[329] (Figure 1.4). Just as with pre-miRNAs, TRBP and Dicer will associate with cytosolic shRNA to be cleaved into 20-25 nucleotide double stranded siRNA with 3' overhangs[330]. Once the siRNA is produced, it is loaded into Ago2 and the RISC is formed.

Because the two strands are perfectly complementary, Ago2 cleaves the passenger strand while the guide strand, which has perfect complementarity to a target mRNA, guides the RISC to its target to be cleaved by Ago2. Synthetic siRNAs entering the cell by endocytosis bypass the cellular processing events required for shRNA biogenesis and instead are directly loaded on Ago2 to mediate hybridization with their mRNA target for cleavage by Ago2[330] (Figure 1.4).

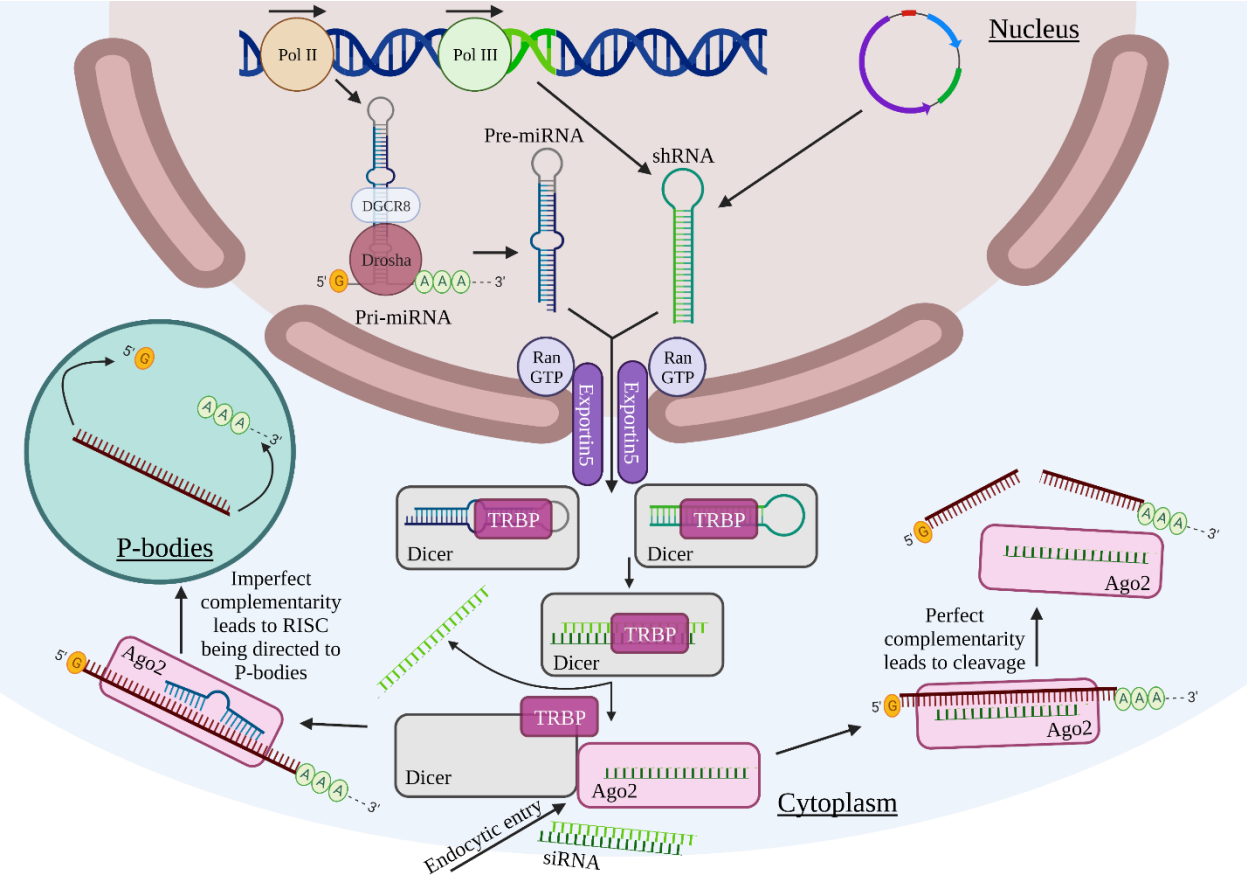

Figure 1.4 The RNAi pathway utilizes miRNAs, shRNAs and siRNAs to inhibit gene expression. The endogenous cellular RNAi pathway utilizes miRNAs for gene regulation. Both shRNAs and siRNAs can be used therapeutically to interfere with the expression of a target gene. Transcription of miRNAs is mediated by RNA Pol II while transcription of shRNAs typically occurs by RNA Pol III. The gene coding for an shRNA can be present in a plasmid which is transfected into the cell or can be integrated in the cellular genome after transduction with a lentiviral vector. Biogenesis of miRNAs initially generates a pri-miRNA which is processed by Drosha and DGCR8 into Pre-miRNA destined for nuclear export while shRNAs are directly exported out of the nucleus. Once in the cytoplasm, both shRNAs and Pre-miRNAs associate with Dicer along with TRBP and the RNA substrates will be cleaved by Dicer into dsRNA products containing two nucleotide 3' overhangs. The complex then recruits Ago2 to form the RISC and

the RNA strands then separate. The guide strand of the shRNA or miRNA is bound to Ago2 while the passenger strand is discarded. Synthetic siRNAs introduced to the cell by endocytosis associate directly with Ago2. Hybridization of the guide strand to the target mRNA can lead to two distinct cellular events. If hybridization is 100% complementary, Ago2 will cleave the target mRNA. If nucleotide mismatches are present between the guide strand and target mRNA, the corresponding RISC will be directed to P-bodies where translational repression or mRNA degradation occurs. Adapted from Singh, S., 2011[331].

1.4.2.1 The structure of RISC: Functions of Dicer, TRBP and Ago2 for shRNA processing

Dicer was originally identified in Drosophila cells where it was characterized as the component of RISC responsible for generating the approximately 22 nucleotide dsRNAs used to direct RNAi to a target mRNA[332]. Although two isoforms of Dicer were identified in Drosophila, mammals possess only a single isoform of the protein. Dicer is part of the RNase III protein family and consists of multiple domains including a helicase domain, two RNase III domains, a dsRNA-binding domain (dsRBD) as well as a Piwi/Argonaute/Zwille (PAZ) domain which is used to recognize dsRNA[333] (Figure 1.5). While dsRBD domains are present, human Dicer requires the assistance of other proteins such as TRBP to be recruited to dsRNA substrates and for proper loading of the guide strand into RISC[334]. Dicer has been shown to possess an L-shape structure, with the PAZ domain on one end of the protein and the helicase domain on the opposite end[335]. The illustration of the overall structure of Dicer served to explain it's "molecular ruler" feature. The spatial separation between the RNase III and PAZ domains determines the cleavage location of the miRNA or shRNA substrate which produces dsRNA products of specific lengths[336]. The Dicer helicase is part of the retinoic acid inducible gene I (RIG-I) family of RNA helicases and is subdivided into three globular domains which undergo conformational changes to clamp down onto dsRNA substrates that are bound by Dicer[337]. Cleavage of these RNA substrates by the RNase III domains of Dicer occurs by metal ion-mediated hydrolysis to result in short dsRNA fragments with two nucleotide overhangs at the 3' end of each strand[338]. Additionally, guide strand selection

from the resulting RNA fragment is mediated by Dicer where the least thermodynamically stable strand (strand with the highest adenine and uracil content) is loaded into the RISC[339].

TRBP is a cellular protein that is vital to the RNAi pathway but was initially identified in the context of HIV-1 infection. This is because TRBP possesses binding specificity towards dsRNA structures and therefore binds to double-stranded regions within the TAR motif present at the 5'end of all HIV-1 RNA transcripts. This binding results in an enhancement of viral expression by relieving a block in translation and consequently amplifying the effect of Tat[340,341]. Two dsRBDs are present in TRBP (dsRBD1 and dsRBD2) and dsRBD2 binds RNA with stronger affinity[342] (Figure 1.5). A third domain (sometimes called half dsRBD or dsRBD type B), is known as the Medipal (Merlin, Dicer, protein kinase R [PKR] activator [PACT] liaison) region, and mediates protein-protein interactions[325,343,344] (Figure 1.5). These domains are present in both isoforms of TRBP. The C4 domain within the Medipal region mediates TRBP binding to Dicer during RNAi[345]. By binding to Dicer, TRBP enhances the cleavage activity through stabilization of the Dicer-dsRNA complex[346]. The exact mechanism through which TRBP facilitates dsRNA processing by Dicer remains unclear but it has been shown that the dsRBD1 and dsRBD2 domains recognize Dicer substrates in a sequence independent fashion[347,348]. In the absence of TRBP, processing of these substrates is impaired and RNAi activity is abolished[345]. Also, proper guide strand selection requires the presence of TRBP as there is an observable defect in the fidelity of strand selection when Dicer is unable to bind to TRBP[349].

Apart from its role in RNAi and HIV-1 replication, TRBP has also been implicated in regulating the cellular interferon response. Indeed, TRBP interferes with the activation of the IFN inducible protein kinase R (PKR). Activated PKR is necessary to phosphorylate the alpha subunit of eukaryotic translation initiation factor 2 (eIF2α) which primes the cell towards an antiviral

response by arresting translation initiation to form stress granules[350,351]. Since dsRNAs act as activators of PKR by inducing its phosphorylation, their binding and sequestration away from PKR by TRBP inhibits PKR activation[352]. Direct binding between TRBP and PKR has also been shown to inhibit PKR phosphorylation and function[353]. In fact, studies have shown that cells which express low basal levels of TRBP do not support productive expression of HIV-1 proteins[354]. Activation of PKR is also possible by cellular PACT, but TRBP can directly interact with PACT to once again inhibit the activation of PKR[355,356].

The family of Ago proteins consists of a group of proteins that were first identified in *Arabidopsis thaliana*[357]. They were initially described as being essential developmental factors and are now known to be critical components of the RNAi pathway by being an essential component of the RISC[358]. Ago proteins can be divided into three subfamilies: the Ago, the P-element induced wimpy testis (Piwi), and the worm-specific Argonautes (WAGO) subfamily[359]. The Ago subfamily binds to miRNAs and siRNAs, as well as to Dicer attached to TRBP, forming the RISC which regulates gene expression by post-transcriptional gene silencing (PTGS)[360,361]. The Piwi subfamily binds to Piwi RNAs (piRNAs) to regulate transposon activity[362]. Finally, the WAGO subfamily binds to secondary siRNAs, which are formed from the action of an RNA-directed RNA Pol being recruited to an mRNA by a primary siRNA[363]. Using the secondary siRNAs, WAGO proteins can silence particular sequences to allow for another layer of gene regulation.

Ago proteins consist of four domains known as the N-terminal domain, the PAZ domain, the middle (MID) domain and the PIWI domain[364] (Figure 1.5). The L1 linker separates the N-terminal and PAZ domain while the L2 linker separates the PAZ and MID domain. The MID domain interacts with the 5’ end of the gRNA while the PAZ domain interacts with the 3’ end[365,366]

(Figure 1.5). The endonuclease activity of Ago proteins occurs by the Piwi domain which contains an RNase H motif and is driven by metal ion-mediated hydrolysis similarly to Dicer[367]. Cleavage of an mRNA target after binding of the gRNA occurs between nucleotide position 10 and 11, with the counting reference point being the 5' end of the gRNA[359]. The human genome includes genes that encode eight Ago proteins, four from the Ago subfamily and four from the Piwi subfamily[368]. The Ago1 and Ago2 proteins have been the most studied and characterized. Ago1 has largely been implicated in chromatin remodulation to induce TGS through the formation of RITS[207,208]. On the other hand, Ago2 has mostly been implicated in PTGS by binding to miRNAs and siRNAs, but may be implicated in TGS in some cases[207,369,370].

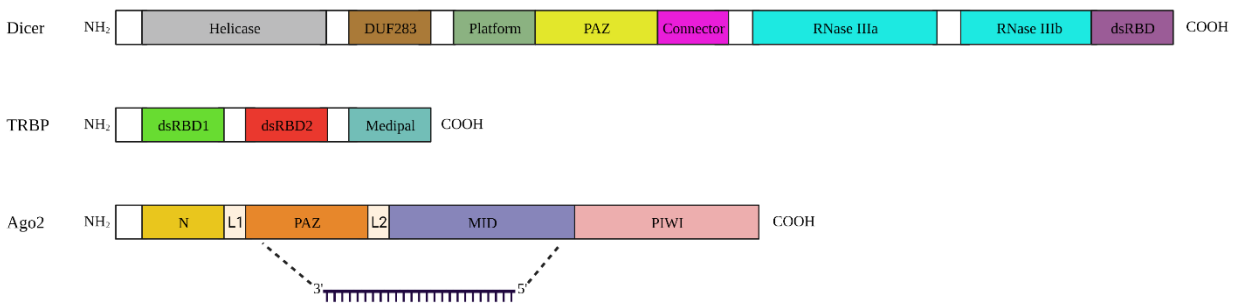

Figure 1.5 RISC components structures. Human Dicer contains a helicase domain, a DUF283 domain of unknown function, a platform-PAZ-connector domain which can recognize dsRNA substrates, two RNase III domains for substrate cleavage and a dsRBD. TRBP is made up of two dsRBDs which bind to RNA as well as the Medipal which mediates protein-protein interactions. Ago2 possesses a N-terminal domain, a PAZ domain, a MID domain and a PIWI domain. Both the PAZ and MID domains are involved in interacting with gRNAs. Figure created with information from Tian, Y., 2014[333], Daniels, S.M., 2012[325], and Schirle, N.T., 2012[371].

1.4.2.2 shRNAs to inhibit viral replication

HIV-1 replication can be inhibited by shRNAs possessing complementarity to specific viral transcripts or to transcripts of cellular genes that are required by the virus to complete its replication cycle (Figure 1.6). The emergence of resistance during treatment of HIV-1 infection is not an issue

unique to cART. Anti-HIV-1 gene therapies that co-opt the RNAi pathway through the introduction of antiviral shRNAs must also contend with this important issue. It has been demonstrated that using only one shRNA during therapy leads to antiviral resistance and administrating multiple shRNAs against different conserved HIV-1 sequences as a combination gene therapy could counteract this problem[372,373]. Resistance is primarily driven by mutations within the target site of shRNAs, as one or two nucleotides mismatches cause the shRNA to be unable to mediate gene silencing[372]. However, mutations outside of the target site which drive the formation of RNA secondary structures surrounding the target site can also confer viral resistance to shRNAs[374]. shRNAs have been designed to target all coding and non-coding regions of the HIV-1 genome[306,328,330,373,375-379].

Targeting cellular elements that are critical for the HIV-1 replication cycle have the benefit that the target site is not under constant evolutionary pressure as is the case with viral targets and therefore cannot mutate in the face of gene therapy. As with gene editing technologies, interfering with cellular CCR5 expression can prevent viral entry and this has been accomplished within various studies using CCR5-specific shRNAs[380-382] (Figure 1.6). Attempts have also been made in designing shRNAs targeting the CXCR4 mRNA[381,383]. While knockdown of the CCR5 gene is not associated with negative consequences, this is not the case for CXCR4. The CXCR4 receptor causes HSCs to localize to the bone marrow and therefore disruption of CXCR4 expression leads to the accumulation of HSCs in the blood, which will interfere with their maturation and differentiation[314,384]. Due to this, gene therapies which interfere with expression of CXCR4 will most certainly be accompanied with severe toxicity *in vivo*.

While it is important to utilize multiple shRNAs and properly select their target sites to ensure strong inhibitory effects without the emergence of resistance, minimizing cytotoxic effects

from the therapy is equally as important. Various mechanisms relating to their antisense functionality and processing by the RNAi machinery can cause cytotoxicity. Off-target effects represent one source of adverse effects as the possibility of tolerance to nucleotide mismatches within the target site can cause the shRNA to bind to unintended cellular mRNAs and interfere with their expression[385]. Cytotoxicity has also been attributed to saturation of the RNAi pathway, where proteins such as exportin-5 and Ago2 that are used for shRNA processing become unavailable for miRNAs biogenesis, resulting in a dysregulation of cellular gene expression[386,387]. Stimulation of the innate immune system through recognition of shRNAs by host RNA sensors is also possible. Activation of TLRs, PKR and RIG-I are all possible in the presence of dsRNAs which ultimately leads to triggering the NF-κB pathway[388,389]. This could upregulate the expression of pro-inflammatory genes to promote cell survival and an anti-viral defense environment[390].

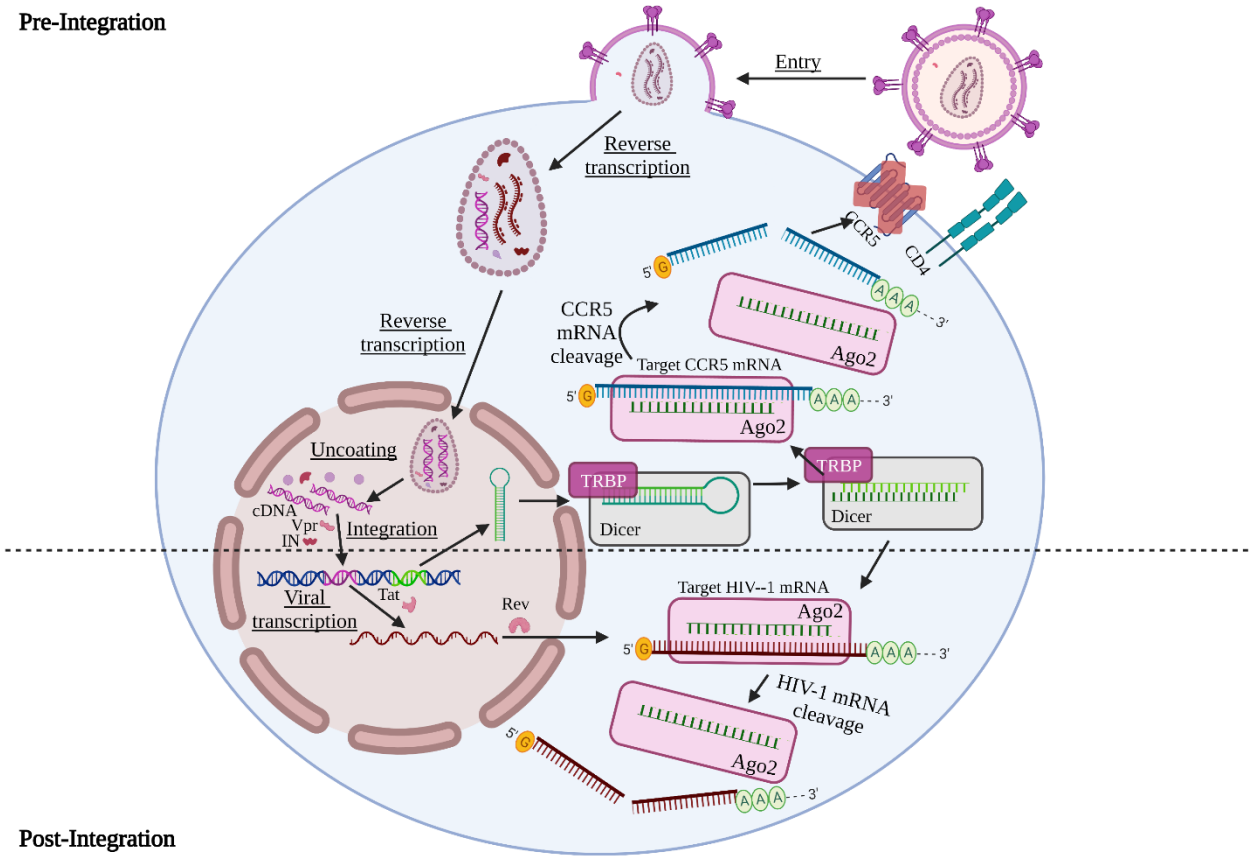

Figure 1.6 Viral replication can be inhibited by shRNAs targeting either cellular CCR5 mRNA or HIV-1 mRNA. Anti-HIV-1 shRNA genes can be integrated into the host genome by a lentiviral vector to act on either pre-integration or post-integration events. CCR5-specific shRNAs cause cleavage of the CCR5 mRNA by Ago2, thus interfering with its expression and inhibiting HIV-1 entry. shRNAs targeting viral transcripts can inhibit viral replication through interfering with the expression of necessary viral elements. A dotted line differentiates pre-integration from post-integration events. Adapted from Goguen, R.P., 2019[314].

1.4.3 Ribozymes to inhibit viral replication

Ribozymes are a type of naturally occurring catalytic RNAs that are necessary for numerous biochemical cellular events such as cleavage of pre-tRNAs to generate functional tRNAs[391], synthesis of proteins by the ribosome[392], self-cleaving satellite RNAs[393] and self-splicing introns[394]. They are also diverse in their structures across different species which forms various ribozyme classes. Since their discovery in 1982, ribozymes have been found to be either

trans or *cis*-acting[35,395]. *Trans*-acting ribozymes include the RNase P ribozymes and ribozymes present within the ribosome, which are responsible for maturation of pre-tRNAs and for peptidyl transferase reactions, respectively[396,397]. *Cis*-acting ribozymes include the self-splicing group I and group II introns, hammerhead (HH) ribozymes, hairpin (Hp) ribozymes and the hepatitis delta virus (HDV) ribozyme[398-401]. The HH, Hp and HDV ribozymes all use the same overall mechanism but are not structurally similar, suggesting that they evolved independently for the same biochemical need[402]. Other self-cleaving ribozymes that have been identified in various organisms are the varkud satellite ribozyme, the glucosamine-6-phosphate synthase ribozyme, the twister ribozyme and the Beta-globin co-transcriptional cleavage ribozyme[403-406].

Exploiting the catalytic capabilities of ribozymes to disrupt gene expression is possible by modifying HH, Hp and HDV ribozymes to act in *trans* instead of in *cis*[35,407] (Figure 1.7). Modifications within stems I and III of HH ribozymes as well as modifications in helices A and B in Hp ribozymes cause these molecules to act in *trans* by binding to specific RNA sequences[408]. Specific on/off adaptors (SOFA) have also been added to the catalytic core of HDV ribozymes to act as a riboswitch where presence of the appropriate RNA sequence turns on the SOFA motif[409,410]. The action of the SOFA motif increases substrate specificity and cleavage efficiency of the ribozyme. The gene silencing mechanism by ribozymes is largely different compared to the one of shRNAs, in that ribozymes mediate the mRNA cleavage themselves and do not use cellular proteins for their catalytic reaction. They therefore have less cytotoxic potential over shRNAs[314]. However, ribozymes have been shown to quickly lose binding specificity to their target when there are mismatches in the target site and they generally offer less potent gene silencing compared to shRNAs[377]. Therapeutic applications have been pursued to use synthetic ribozymes towards viral infections such as HIV-1 and hepatitis C virus[411]. In the context of HIV-1 gene therapy, the RNase

P ribozyme along with the HH, Hp and HDV ribozymes have been engineered to target viral transcripts[35] (Figure 1.7). Such targets include mRNAs containing the regions *vpr/tat*[412], *gag*[377,413], *tat* and *tat/rev*[410], *pol*[414] and the UTR[415]. Targeting of the cellular CCR5 mRNA by a HH ribozyme to block viral entry has also been possible[416] (Figure 1.7). It is worth noting that the intolerance of ribozymes for nucleotide mismatches between the target site and their recognition domain make these molecules susceptible to losing their catalytic activity in the face of viral mutations. Similar to shRNAs, their use in combination gene therapy will be necessary to avoid the emergence of resistance.

Recent applications of ribozymes for HIV-1 gene therapy have been to include their native *cis*-acting structure at the 3' end of other antiviral RNAs. Such molecular design removes the 3' U-tail of variable length that results from transcription by the Pol III promoters, thus producing transcripts with uniform 3' ends[417]. CRISPR guide RNAs, AgoshRNAs and aptamers have all had self-cleaving ribozymes incorporated at their 3' end in an attempt to increase their functionality by generating such uniform products[417-419].

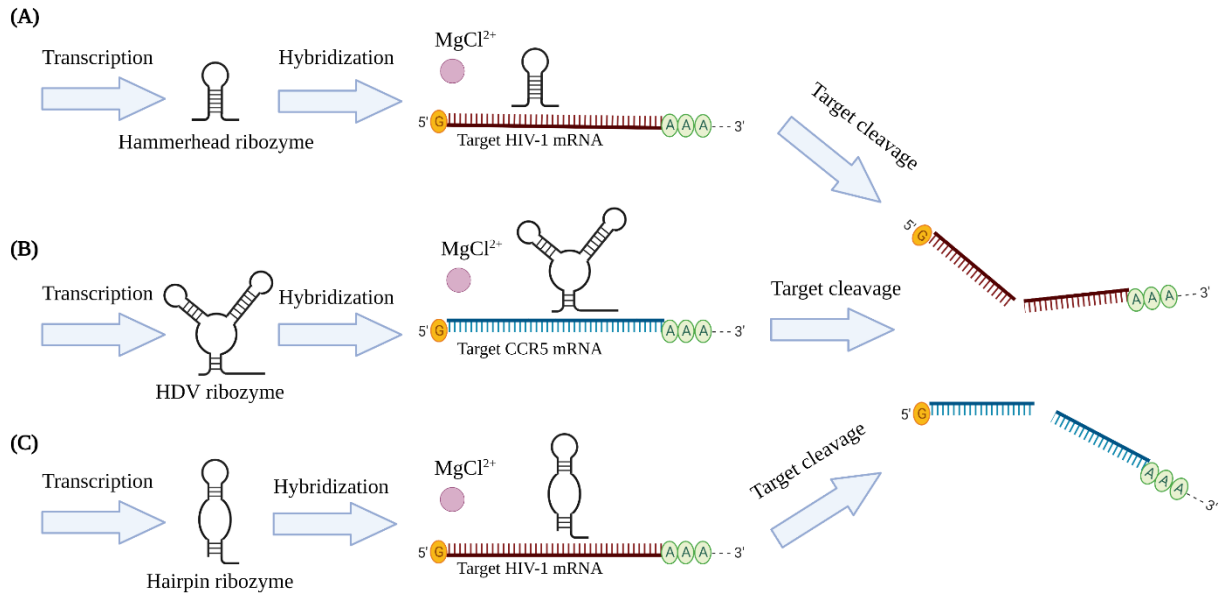

Figure 1.7 Ribozymes can be used to cleave either cellular CCR5 mRNA or HIV-1 mRNA to inhibit viral replication. Various ribozymes including hammerhead **(A)**, HDV **(B)** and hairpin **(C)** ribozymes have been engineered to recognize a target CCR5 mRNA or HIV-1 mRNA in a sequence-specific manner to mediate cleavage. Cleavage of CCR5 mRNAs interferes with expression of the HIV-1 CCR5 coreceptor, effectively blocking viral entry. Cleavage of HIV-1 transcripts inhibits viral replication through impeding the expression of critical viral elements. Adapted from Goguen R.P., 2023. In press.[420]

1.4.4 Aptamers and decoys to inhibit viral replication

Aptamers are RNA molecules which possess binding specificity through their three-dimensional structure instead of by nucleotide complementarity as is the case with shRNAs and ribozymes[314]. Generation of aptamers is possible through a procedure termed systematic evolution of ligands by exponential enrichment (SELEX). This technique was originally developed independently by two different research groups using *in vitro* screens of large libraries of RNA molecules to identify ligands that could bind to T4 DNA polymerase or to organic dyes[421,422]. SELEX can be applied to virtually any molecular target where a large enough sample size of randomized oligonucleotides will lead to a small fraction possessing binding capabilities to the

target. After an initial incubation of the RNA library with the target molecule, non-binding aptamers are discarded while the bound aptamers are amplified by reverse transcription PCR (RT-PCR). RT-PCR is possible through primers binding to the fixed ends included in each RNA molecule of the library[423]. The resulting DNA strands will be transcribed to produce a smaller pool of RNAs with affinity to the target, these are once again incubated with the molecular target to allow for subsequent rounds of enrichment with the goal of identifying aptamers with the greatest affinity for the target[423].

The binding specificity of aptamers have led to their use as biosensors and for bioimaging by incorporating fluorescent dyes or radioisotopes in the aptamer to report the presence and/or location of a target[424-426]. Therapeutic applications are also possible as binding of the aptamer interferes with the functionality of the target which can be used to treat cancers, age-related macular degeneration and infectious diseases[314,427]. Aptamers targeting various HIV-1 proteins and RNA motifs have been designed to inhibit viral replication. Viral targets of aptamers include the HIV-1 protease[428], reverse transcriptase[429], gag polyprotein[430], integrase[431] and UTR[432] (Figure 1.8A-E). Molecular designs of anti-HIV-1 RNAs have now included aptamers as the loop of shRNAs to generate shRNA-aptamer chimeras[431]. In theory, the entire transcript should be exported to the cytoplasm and processed by the RNAi machinery to separate the two antiviral molecules. This would allow for the shRNA component to proceed with the formation of RISC while the aptamer binds on its own to its intended cytoplasmic target such as HIV-1 Gag, protease and RT.

RNA decoy molecules are similar to aptamers, as they can also bind to a target through their three-dimensional structure, but are not developed by SELEX and instead mimic natural RNA structures which are known to bind to viral elements[314]. Both the TAR and RRE RNA motifs have

been mimicked by RNA decoys[433-435] (Figure 1.8F and G). Inhibition of viral replication from these two RNA decoys occurs by competitive binding to their individual targets against the natural HIV-1 TAR and RRE elements. When the TAR decoy is expressed within an infected cell, it will bind to viral Tat and sequester it away from HIV-1 TAR located near the promoter so that the recruitment of P-TEFb does not occur[314] (Figure 1.8F). Tat possesses nucleolar localization properties which requires the TAR decoy to be localized to this region to efficiently inhibit viral replication[436,437]. Therefore, the U16 small nucleolar RNA backbone has been conjugated to the TAR aptamer to allow for its nucleolar localization where it can sequester Tat more effectively[434]. The RRE decoy functions by binding to viral Rev and sequesters it away from the natural RRE found in unspliced and incompletely spliced viral transcripts[435] (Figure 1.8G). This event impairs the Rev-dependent nuclear export of these transcripts and results in an inhibition of HIV-1 replication. The U16 small nucleolar RNA backbone has also been conjugated to the RRE decoy where an increased inhibitory effect was observed compared to the decoy used alone[438].

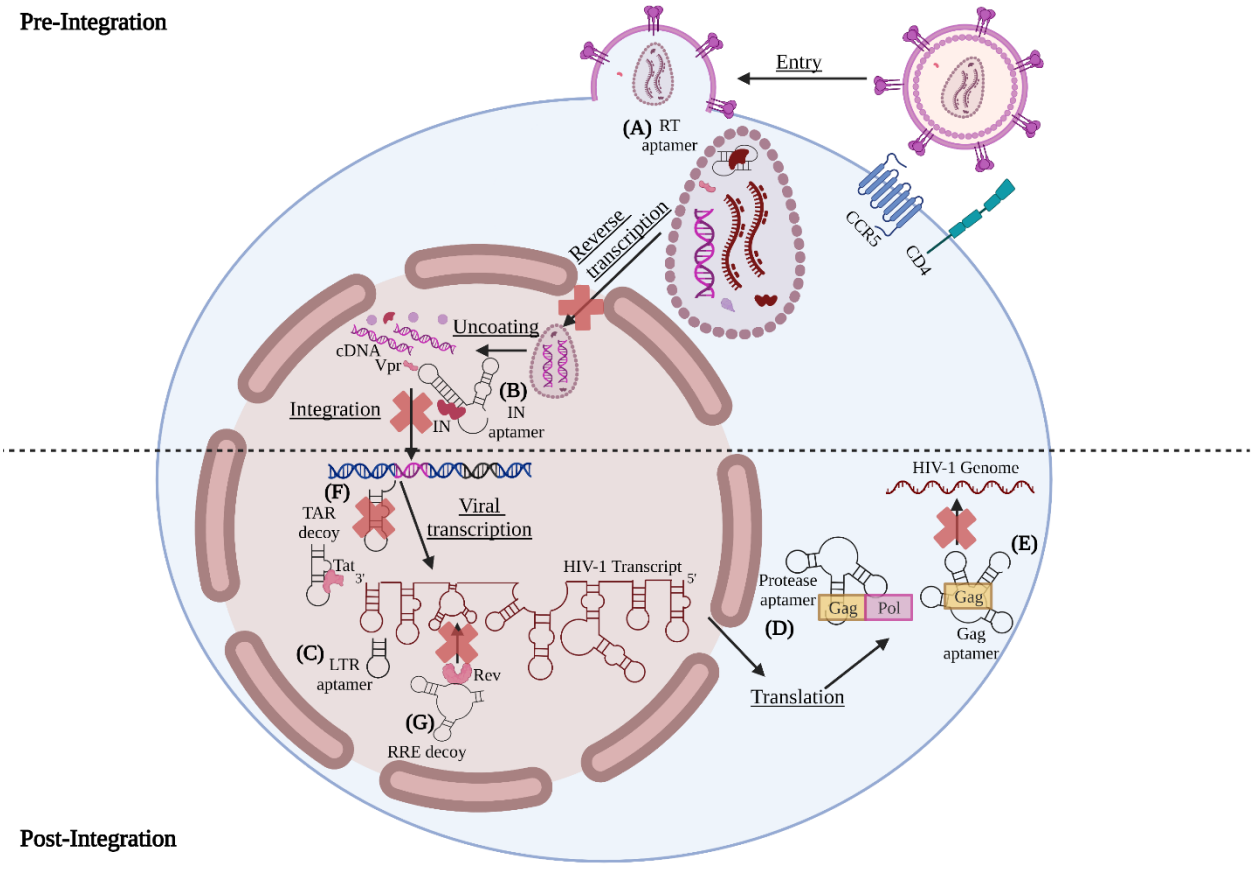

Figure 1.8 Viral replication can be inhibited by aptamers and decoys binding to viral proteins and RNA motifs. Anti-HIV-1 aptamer and decoy genes can be integrated into the host genome by a lentiviral vector to act on either pre-integration or post-integration events. These molecules bind to their target through their three-dimensional structures to inhibit various steps of the viral replication cycle. **(A)** RT aptamers bind to reverse transcriptase to inhibit the conversion of viral RNA to cDNA. **(B)** IN aptamers bind to integrase to inhibit integration of viral cDNA into host cell genomes. **(C)** LTR aptamers bind to the poly (A) domains of HIV-1 transcripts which prevents post-transcriptional processing. **(D)** Protease aptamers inhibit proteolytic cleavage of Gag and Gag-Pol polyproteins by viral protease. **(E)** Gag aptamers interfere with packaging and assembly by competing with viral Gag for binding of genomic viral RNA. **(F)** TAR decoys prevent the enhancement of viral transcription by binding to Tat to sequester it away from viral TAR. **(G)** RRE decoys interfere with nuclear export of viral transcripts by binding to Rev which sequesters it away from viral RRE. A dotted line differentiates pre-integration from post-integration events. Adapted from Goguen, R.P., 2019[314].

1.4.5 U1i RNAs to inhibit viral replication

U1i RNAs consist of reengineered U1 small nuclear RNAs (U1 snRNAs) that are redirected towards a particular target RNA such as HIV-1 RNA. Within their natural context, U1

snRNAs complex with seven Smith proteins as well as with three U1-specific proteins to form a U1 small nuclear ribonucleoprotein complex (snRNP)[439]. These U1 snRNPs interact with the U2, U5 and U4/U6 snRNPs, which each contain a unique snRNA and associated proteins, to assemble into the spliceosome. Spliceosomal assembly is initiated by recruitment of the U1 snRNP to 5' splice sites (5'ss) through RNA-RNA interactions by the 5' end of the U1 snRNA[440]. *Cis* elements within introns are then bound by non-snRNP factors, such as U2AF binding to the polypyrimidine tract and SF1/mBBP binding to the branch point, which allows for recruitment of the U2 snRNP through its recognition of the branch point[439]. The U5 and U4/U6 snRNP preassemble into a U4/U6.U5 tri-snRNP before being recruited towards the U1 and U2 snRNPs-bound site. Once recruitment of the different snRNPs has occurred, various rearrangements of the spliceosomal units take place to result in the association of the U6 snRNA with the U2 snRNA to yield the spliceosome active site which catalyzes the splicing reaction by metal ion-mediated catalysis[441]. Although the U1 snRNP is invaluable in forming the spliceosome, it can also repress the polyadenylation of mRNAs by binding to regions surrounding the polyadenylation site (PAS). This is done by the U1-specific protein known as U1-70K interacting with poly(A) polymerase (PAP) which interferes with 3' end processing of the mRNA[442,443].

U1i RNAs are generated by modifying the 5' recognition domain of the U1 snRNA to be complementary to a target mRNA near its PAS and will result in an inhibition of gene expression through a technique known as U1 interference. Gene silencing occurs as a result of an inhibition in 3' end polyadenylation mediated by the U1-70K protein[440]. Modified U1 snRNAs have also been designed to bind to regions proximal to splice sites in order to correct genetic diseases originating from defects in splicing[444-447]. Familial dysautonomia, Nasu-Hakola disease and retinis pigmentosa are all genetic diseases originating from splice site mutations which result in exon

skipping. For each of these diseases, the splicing disorder caused from the mutations can be rescued by introducing modified U1 snRNAs which are capable of either binding to the mutated ss or bind to downstream region[444-447]. In both instances, spliceosomal components are recruited to the mutated region which significantly reduces exon skipping. Use of U1i RNAs as antiviral molecules has also been explored to interfere with the replication of HIV-1 and HBV[440,448-451]. U1i RNAs targeting HIV-1 transcripts have been designed to bind to the 3' proximal regions of HIV-1 to inhibit polyadenylation[449,450] (Figure 1.9). HIV-1-specific modified U1 snRNAs have also been designed to bind to either 5'ss or downstream of 3'ss to enhance splicing of viral transcripts. This detrimentally alters the expression of HIV-1 transcripts to inhibit viral production and therefore can be considered as another type of U1i[451] (Figure 1.9). U1i RNAs which interfere with polyadenylation were found to be effective inhibitors in viral production experiments but not in viral replication experiments[449]. In contrast, U1i RNAs which promote splicing were effective inhibitors in both HIV-1 production and replication experiments, as well as being more potent inhibitors of HIV-1 production when directly compared to those that inhibit polyadenylation[440]. As is the case of shRNAs and ribozymes, the fact that U1i RNAs recognize their target through nucleotide complementarity makes them prone to becoming ineffective in the event of nucleotide mismatches from viral mutations. Therefore, combination gene therapy will also be necessary when utilizing U1i RNAs to prevent the emergence of resistance. Studies have already acknowledged this problem and have found that using both U1i RNAs and shRNAs to interfere with gene expression confers synergistic inhibitory effects[448,452].

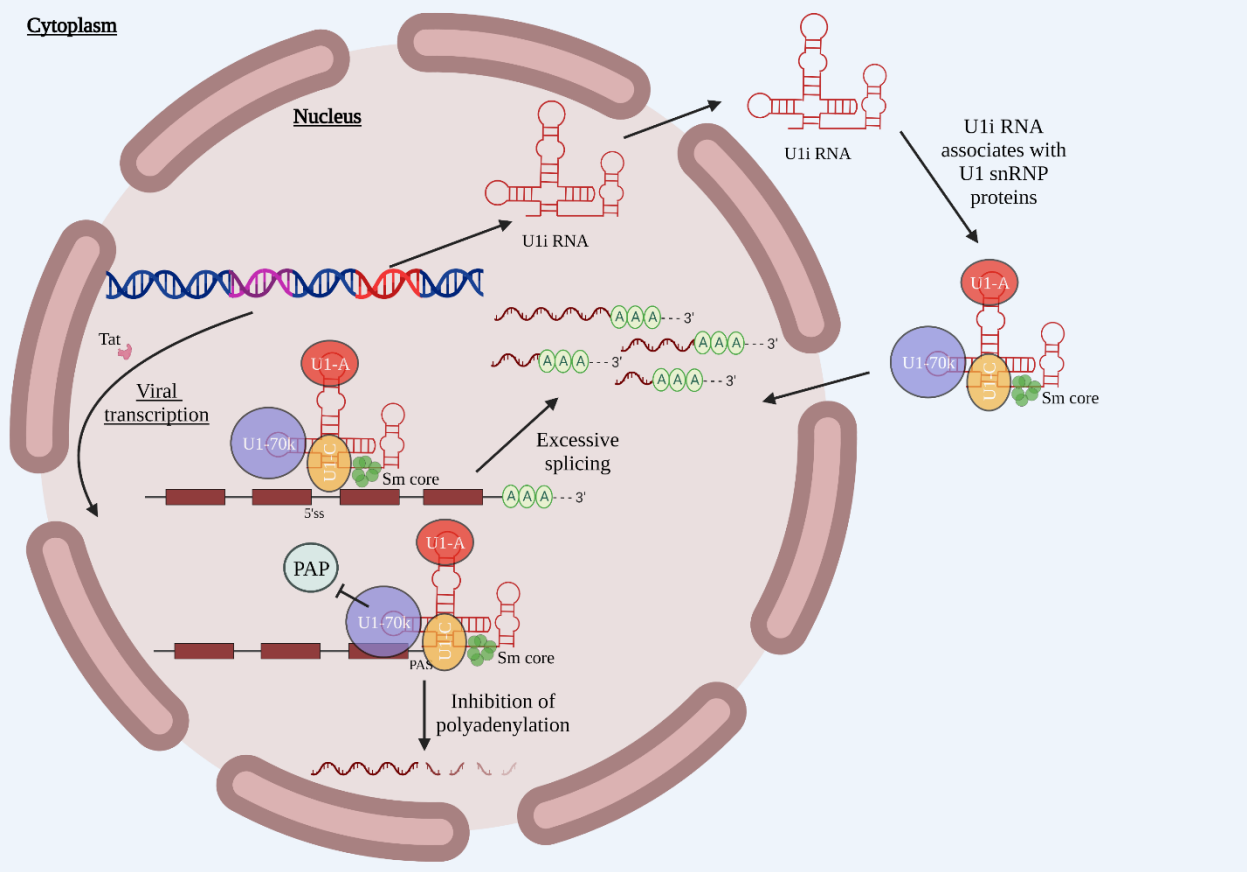

Figure 1.9 Viral replication can be inhibited by U1i RNAs targeting HIV-1 mRNA. Anti-HIV-1 U1i RNA genes can be integrated into the host genome by a lentiviral vector. Binding of the U1i RNA to 5' splice sites or downstream of 3' splice sites causes excessive splicing which leads to the absence of full length viral transcripts. The U1i RNA can also be designed to bind to the regions proximal to the PAS, where interactions between U1-70K and PAP inhibit polyadenylation of the viral transcript, leading to its degradation due to mRNA instability. Adapted from Chen, M.J., 2023[453].

1.5 Gene delivery

Efficient cellular delivery of therapeutic genes is critical to ensure that gene therapy is effective in the treatment of diseases. Gene delivery is generally performed after incorporating the gene of interest into viral vectors and transducing them in cells. The most common viral vectors are derived from retroviruses (including lentiviruses), adenoviruses or adeno-associated viruses (AAV). The first gene therapy clinical trial was for the treatment of adenosine deaminase severe

combined immunodeficiency (SCID) with a retroviral vector[454]. Various gene therapy trials using retroviral vectors, specifically the murine leukemia virus (MuLV)-based gammaretroviral vectors, were subsequently initiated to treat numerous diseases such as X-linked SCID[455], malignant brain tumors[456] and HIV-1 infection[307,310,312]. This boom in clinical progress declined for some time because of the observation of severe adverse effects, which happened when using adenoviral and retroviral vectors. One case resulted in the death of an 18-year-old male due to activation of his immune system in response to the high dose of adenoviral vector administered. This vector was used to deliver a corrective gene for treatment of a mild form of ornithine transcarbamylase deficiency[457]. In another case, the MuLV-based vectors used in early gene therapy trials resulted in uncontrolled clonal expansion of T cells in five patients who participated in clinical trials to treat X-linked SCID[458-460]. The development of leukemia in these patients was shown to be driven by the integration of the retroviral vector near LMO2, a proto-oncogene promoter, which led to its activation and subsequent malignant transformation[459]. Since these untimely events, gene therapy has been regarded more critically and improvements in the selection of effective delivery vectors which minimize cytotoxic as well as oncogenic potential have been prioritized. As a consequence, third generation, very defective adenoviral vectors are used and lentiviruses have replaced MuLV-based retroviral vectors. AAV are also widely used for transient delivery and are not toxic. Importantly, no further severe adverse effects related to viral vectors have been observed in clinical trials and the number of trials has been increasing considerably over the past decade. Gene delivery is now possible with both viral and non-viral vectors which can result in either transient or permanent expression of the delivered nucleic acids. In the context of HIV-1 gene therapy, using delivery vectors to transduce HSCs *ex vivo* with antiviral genes will likely be necessary as it is

difficult to rely on a particular delivery vector to reach all target cells of the virus if the vector is delivered directly *in vivo*.

1.5.1 Non-viral vectors for gene delivery

Non-viral vectors used for gene therapy result in transient expression of the gene. This delivery strategy therefore completely avoids the occurrence of oncogenesis from integration events[461]. Additionally, these vectors offer lower immunogenicity compared to viral vectors[462]. However, non-viral vectors will dilute during cellular proliferation and repeat administration will be necessary to maintain expression of the transgene. Non-viral vectors typically consist of cationic polymers or liposomes which complex with DNA or RNA to form a nanoparticle by interactions with the negative charges present in the phosphate groups of the nucleic acids[463]. This encapsulation protects the nucleic acids from endonucleases, improves their stability during delivery and allows them to enter their target cells. A significant barrier for delivery of DNA is their need to be transported to the nucleus for transcription but this can be circumvented by replacing the DNA with mRNAs[461]. Both synthetic siRNAs and miRNAs can also be encapsulated in polymers or liposomes and delivered to the cells where they will mediate gene silencing through the RISC. Delivery to target tissues and cell types has also been an issue with these vectors as classically they are non-specific. Recent improvement in non-viral vectors have incorporated ligands which possess binding specificity for cellular receptors and therefore can orient the vector towards a target cell or tissue by receptor-mediated endocytosis[464,465].

1.5.2 Viral vectors for gene delivery

Viral vectors can be used to achieve either transient or permanent expression of a delivered gene. This is dependent on whether the vector can integrate into the cellular chromosome or not. Viral vectors are generally more effective at delivering a transgene over non-viral vectors, but they

can be highly immunogenic or result in carcinogenesis[466,467]. These vectors cannot form new virions as they are deficient in their ability to complete an entire replication cycle, but they retain their ability to enter a cell and to inject their viral genome[468]. Incorporating a therapeutic gene within the viral genome therefore allows it to also enter the cell.

1.5.2.1 Non-integrating viral vectors for gene delivery

Several adenoviral based non-integrating vectors have been designed. These recombinant adenoviruses are DNA vectors and allow for broad tropism as over 100 serotypes have been identified which can infect a wide range of cells and tissues[469]. As the first generation of adenoviral vectors elicited a strong immune response, the initial vector design was improved by removing all adenoviral genes from the vector. This gutted vector is then supplied by helper viruses (in *trans*) to provide the missing proteins during vector production[470]. The issue of immunogenicity from adenoviral vectors is compounded by the fact that the resulting transient expression of therapeutic genes necessitates repeated administrations during gene therapy which increases its likelihood of recognition by the adaptive immune system. While eliciting an immune response may in some cases not be harmful to patients, it will nevertheless lead to the vector being rapidly cleared by the immune system and consequently will decrease the effectiveness of the gene therapy. Additionally, antibody cross-reactivity towards the vector due to pre-existing immunity to circulating adenoviral strains is also likely to limit the effectiveness of the therapy[471].

AAV vectors consists of another type of non-integrating viral vector. AAVs are nonpathogenic satellite parvoviruses that are effective vectors for gene delivery as they present very low immunogenicity[472,473]. Broad tropism is also offered due to AAV being composed of multiple serotypes with capsid specificity that allow them to target specific tissues. While wild type AAVs possess the ability to integrate into a specific site in chromosome 19, AAV vectors are

constructed so that the *cis* elements required for this integration event are replaced by the therapeutic gene[469,474]. As is the issue with adenoviral vectors, the absence of an integration event will eventually cause the loss of the delivered gene during proliferation of the target cells and repeated administrations are therefore needed during the therapy[475].

Herpes simplex virus (HSV) has also been explored as an additional type of non-integrating viral vector. Similar to adenoviral vectors, HSV vectors are gutted so as to be replication incompetent and viral elements required for their production are supplied in trans. Transduction using such HSV vectors causes a latent-like infection which leads to long-term expression of the transgene that is included in the vector[476]. Recombinant HSV vectors have found uses in clinical trials for gene therapy against brain tumors as this vector is neurotropic[477,478]. Overall, both non-viral and viral non-integrating vectors may not be suitable for HIV-1 gene therapy as their requirement for frequent administrations will not offer a solution towards the chronic administration of cART.

1.5.2.2 Integrating viral vectors for gene delivery

Integrating viral vectors consists of retroviral RNA vectors. These differ from the previously discussed DNA vectors as their RNA genome must be reverse transcribed by viral proteins to generate the DNA product that will integrate in the host chromosome and will be transcribed to express the therapeutic gene. These vectors are designed so that retroviral genes coding for replication elements are eliminated and replaced with the gene of interest, but the genomic regions required for packaging, reverse transcription and integration located in the LTR are maintained[469]. Retroviral proteins that mediate reverse transcription and integration are instead supplied in *trans* during vector formation[479]. The vesicular stomatitis virus G protein (VSV-G) can be used instead of the original retroviral envelope protein to allow for broad tropism[480]. VSV-G-

pseudotyped lentiviral vectors use the low-density lipoprotein receptor (LDLR) as the primary cell receptor for cell entry[481] (Figure 1.10). Integration of the therapeutic gene in target cells allows for stable expression of the transgene that is not lost during cellular division[482,483]. In the context of HIV-1 gene therapy, this is highly desirable to eliminate the need for chronic administration of therapy as is the case with cART.

Since the early usage of MuLV-based retroviral vectors possessed carcinogenic risks, gene therapies seeking to generate stable expression of a transgene have instead turned towards lentivirus-based retroviral vectors as a delivery tool. Lentiviral vectors possess a low oncogenic potential as they preferentially integrate away from the start site of transcriptional units whereas MuLV-based vectors preferentially integrate in direct proximity of the start site[484]. Indeed, during HIV-1 gene therapies using lentiviral vectors, the integration site occurs in gene rich areas without the development of oncogenesis[485,486]. To further increase their safety, current lentiviral vectors are also typically self-inactivating (SIN). SIN vectors are designed with a deleted segment within the 3' LTR[487] (Figure 1.10). While the 5' LTR is initially intact during the formation of the vector to keep its *cis* acting elements functional, reverse transcription of the viral RNA will result in the 3' LTR deletion being copied into the 5' LTR[487] (Figure 1.10). The presence of this deletion within the 5' LTR abrogates transcription of the integrated provirus and removes the strong enhancer activity of the LTR promoter on nearby genes. SIN vectors are needed for HIV-1 gene therapy as the use of non-SIN vectors may result in lentiviral transcripts being packaged by HIV-1 proteins to create chimeric virions within HIV-1-infected cells[486]. It is important to note that since the proviral transcription is defective in SIN vectors, these lentiviral vectors require the inclusion of an internal promoter to express the therapeutic gene[482].

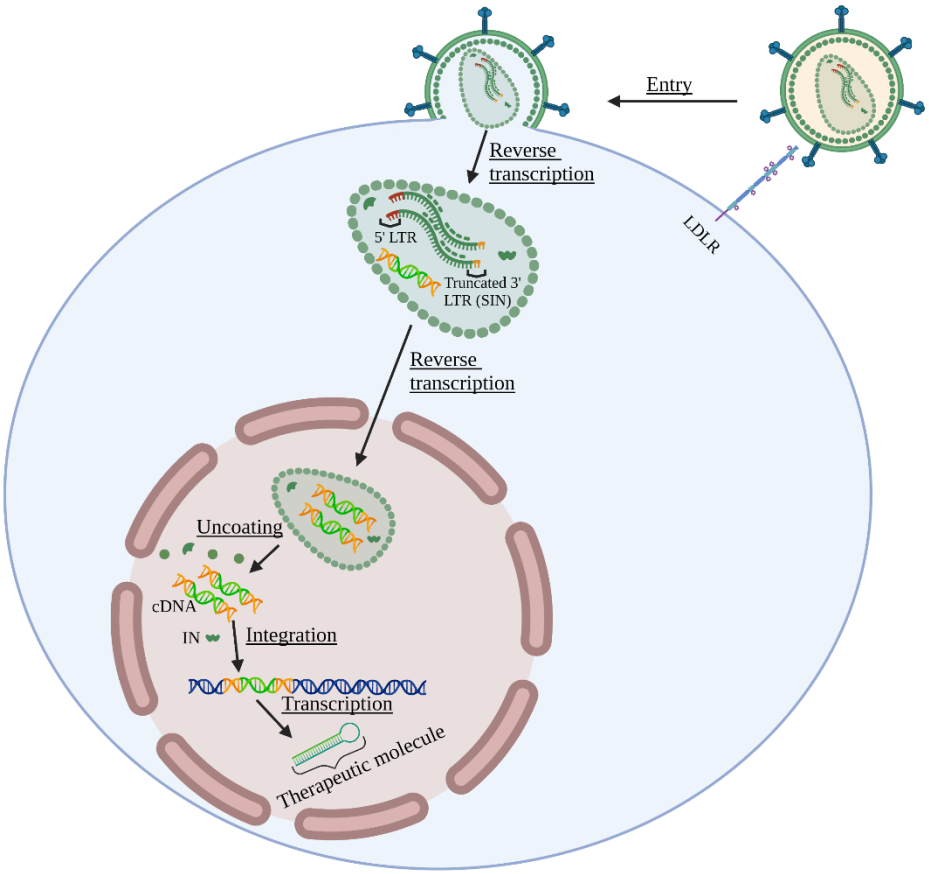

Figure 1.10 Self-inactivating lentiviral vectors mediate transgene delivery into host chromosome. Self-inactivating lentiviral vectors possess elements allowing for reverse transcription and integration, but have a truncated 3' LTR. Reverse transcription causes the deletions of the 3' LTR to be copied into the 5' LTR of the resulting cDNA, thus preventing transcription of the provirus after integration. Self-inactivating lentiviral vectors contain an internal promoter upstream of the transgene to allow for transcription of the therapeutic molecule of interest. Figure created with information from Yu, S.F., 1986[487], and Finkelshtein, D., 2013[481].

1.5.3 Promoters available for expression of anti-HIV-1 RNAs

Eukaryotic cells possess three different types of RNA Pol promoters. The Pol I promoter controls the transcription of most rRNAs while the Pol II promoter is responsible for driving the expression of protein coding genes as well as miRNAs[328]. Pol II promoters have been used in HIV-1 gene therapy to express aptamers as well as antiviral proteins[285,428,430]. The Pol III promoters are unique in that they can be separated into three different types which exclusively transcribe non-

coding RNAs[488]. The type 3 Pol III promoters are typically used in gene therapy since they possess a defined transcriptional start site that does not extend into the promoter sequence, unlike the type 1 and type 2 Pol III promoters[489].

Selection between the Pol II promoters and the type 3 Pol III promoters is typically based on whether the therapeutic RNA is destined for localization in the nucleus, the cytoplasm or whether the transcript must be translated into a functional protein[490]. The Pol II promoters have the advantage of allowing post-transcriptional modifications of a transcript, such as addition of the 5' cap and polyadenylation of the 3' terminus, which is necessary for nuclear export and for translation[491-494]. For this reason, transcripts coding for antiviral proteins as well as aptamers inhibiting cytoplasmic HIV-1 replication events such as protease and gag aptamers are well suited for expression from the Pol II promoters[431]. Although expression from the Pol II promoters allows aptamers to be localized to the cytoplasm, the addition of extraneous RNA sequence have the possibility of disrupting the specific three-dimensional structure of the aptamer and decrease its binding affinity[431]. Transcripts originating from the type 3 Pol III promoters do not undergo post-transcriptional modifications and therefore accumulate in the nucleus where they cannot be translated[495-497]. The exception to this is in the case of shRNAs which, in the absence of post-transcriptional modifications, are directed towards the cytoplasm through a specific nuclear export pathway. Nuclear accumulation is ideal for antiviral molecules which interfere with nuclear HIV-1 replication events such as TAR and RRE decoys. Additionally, therapies which rely on the generation of gRNAs with a specific sequence such a CRISPR-Cas9 gene editing or shRNA gene therapy are typically driven by type 3 Pol III promoters as transcription results in defined 5' and 3' ends[328,498,499].

Optimization of gene expression is absolutely critical in gene therapy. Comparisons between different promoters expressing small RNAs have mostly been done with the type 3 Pol III promoters H1, 7SK and U6[328]. Evaluation of these different promoters has shown that some offer weak antiviral effects in gene therapy and others are associated with cytotoxicity. Complicating the matter, reports of which promoters offer the largest therapeutic effect with the lowest potential for cytotoxicity is not consistent, suggesting that the expressed antiviral molecule is a determinant for promoter choice. Specifically, the U6 promoter has been found to be associated with cytotoxic effects which can be alleviated by using the H1 promoter to drive the expression of therapeutic genes[380,500]. However, there have been other reports showing that the use of the U6 promoter is not associated to any toxicities[501]. In regards to therapeutic effect, the U6 promoter has been found to maximize gene activity in some studies[500-502], while another study showed that the H1 promoter expresses the most potent therapeutic molecules[503]. In addition, one study reported no difference in the activity of a transgene when expressed from the H1, U6 and 7SK promoters[504]. In the face of such conflicting information, it is imperative to properly evaluate each promoter independently for expression when designing a gene therapy against HIV-1. This allows for optimization of expression of anti-HIV-1 RNAs to inhibit viral replication without inducing cytotoxicity. Although clinical trials using gene therapy against HIV-1 infection have been conducted (Table 1.3), their limited effectiveness suggests that adequate optimization is still needed to reach effective therapeutic outcomes by antiviral gene therapy[287,306,311]. Optimization of the expression of individual antiviral molecules is certainly a necessary first step before selecting which molecules to use in parallel. The overall goal of such careful screening is to establish an effective combination gene therapy that controls HIV-1 infection without the emergence of resistance and in the absence of toxicity.

1.6 References.

1. Voisset, C., Weiss, R.A. & Griffiths, D.J. Human RNA "rumor" viruses: the search for novel human retroviruses in chronic disease. *Microbiol Mol Biol Rev* **72**, 157-196 (2008).
2. Chuong, E.B. The placenta goes viral: Retroviruses control gene expression in pregnancy. *PLoS Biol* **16**, e3000028 (2018).
3. Belshaw, R., *et al.* Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci U S A* **101**, 4894-4899 (2004).
4. Coffin, J., *et al.* Human immunodeficiency viruses. *Science* **232**, 697 (1986).
5. Bailes, E., *et al.* Hybrid origin of SIV in chimpanzees. *Science* **300**, 1713 (2003).
6. Sharp, P.M. & Hahn, B.H. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* **1**, a006841 (2011).
7. Hemelaar, J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med* **18**, 182-192 (2012).
8. Sharp, P.M. & Hahn, B.H. The evolution of HIV-1 and the origin of AIDS. *Philos Trans R Soc Lond B Biol Sci* **365**, 2487-2494 (2010).
9. Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H. & Johnson, P.R. An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339**, 389-392 (1989).
10. Visseaux, B., Damond, F., Matheron, S., Descamps, D. & Charpentier, C. Hiv-2 molecular epidemiology. *Infect Genet Evol* **46**, 233-240 (2016).
11. Centers for Disease, C. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb Mortal Wkly Rep* **30**, 305-308 (1981).
12. de Sousa, J.D., Muller, V., Lemey, P. & Vandamme, A.M. High GUD incidence in the early 20 century created a particularly permissive time window for the origin and initial spread of epidemic HIV strains. *PLoS One* **5**, e9936 (2010).
13. Greene, W.C. A history of AIDS: looking back to see ahead. *Eur J Immunol* **37 Suppl 1**, S94-102 (2007).
14. Barre-Sinoussi, F., *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871 (1983).
15. Gallo, R.C., *et al.* Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* **220**, 865-867 (1983).
16. Popovic, M., Sarngadharan, M.G., Read, E. & Gallo, R.C. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* **224**, 497-500 (1984).
17. Levy, J.A., *et al.* Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science* **225**, 840-842 (1984).
18. UNAIDS. Global HIV & AIDS statistics — 2022 fact sheet. https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf. Accessed.
19. Vermund, S.H., Sheldon, E.K. & Sidat, M. Southern Africa: the Highest Priority Region for HIV Prevention and Care Interventions. *Curr HIV/AIDS Rep* **12**, 191-195 (2015).

20. Harrison, A., Colvin, C.J., Kuo, C., Swartz, A. & Lurie, M. Sustained High HIV Incidence in Young Women in Southern Africa: Social, Behavioral, and Structural Factors and Emerging Intervention Approaches. *Curr HIV/AIDS Rep* **12**, 207-215 (2015).
21. Alexander, K.A. Social determinants of HIV/AIDS and intimate partner violence: interrogating the role of race, ethnicity and skin color. *Rev Lat Am Enfermagem* **28**, e3280 (2020).
22. Barbera, L.K., *et al.* HIV and COVID-19: review of clinical course and outcomes. *HIV Res Clin Pract* **22**, 102-118 (2021).
23. Cooley, S.A., Nelson, B., Doyle, J., Rosenow, A. & Ances, B.M. Collateral damage: Impact of SARS-CoV-2 pandemic in people living with HIV. *J Neurovirol* **27**, 168-170 (2021).
24. Sun, S., *et al.* Challenges to HIV Care and Psychological Health During the COVID-19 Pandemic Among People Living with HIV in China. *AIDS Behav* **24**, 2764-2765 (2020).
25. Granich, R., *et al.* Status and methodology of publicly available national HIV care continua and 90-90-90 targets: A systematic review. *PLoS Med* **14**, e1002253 (2017).
26. Guaraldi, G., *et al.* The Impact of COVID-19 on UNAIDS 90-90-90 Targets: Calls for New HIV Care Models. *Open Forum Infect Dis* **8**, ofab283 (2021).
27. Deeks, S.G., *et al.* Research priorities for an HIV cure: International AIDS Society Global Scientific Strategy 2021. *Nature medicine* **27**, 2085-2098 (2021).
28. Dunkley, O.R.S., Scarborough, R.J. & Gatignol, A. 40th year anniversary of HIV discovery: Evolving paradigms for cure strategies. *Virologie*, In press (2022).
29. Engelman, A. & Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol* **10**, 279-290 (2012).
30. Bell, N.M. & Lever, A.M. HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* **21**, 136-144 (2013).
31. Sundquist, W.I. & Krausslich, H.G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* **2**, a006924 (2012).
32. Swanson, C.M. & Malim, M.H. SnapShot: HIV-1 proteins. *Cell* **133**, 742, 742 e741 (2008).
33. Briggs, J.A., Wilk, T., Welker, R., Krausslich, H.G. & Fuller, S.D. Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J* **22**, 1707-1715 (2003).
34. Arhel, N. Revisiting HIV-1 uncoating. *Retrovirology* **7**, 96 (2010).
35. Scarborough, R.J. & Gatignol, A. HIV and Ribozymes. *Adv Exp Med Biol* **848**, 97-116 (2015).
36. Clapham, P.R. & McKnight, A. HIV-1 receptors and cell tropism. *Br Med Bull* **58**, 43-59 (2001).
37. Peterlin, B.M. & Trono, D. Hide, shield and strike back: how HIV-infected cells avoid immune eradication. *Nat Rev Immunol* **3**, 97-107 (2003).
38. Arhel, N.J. Early HIV replication revisited. *Nat Microbiol* **5**, 1065-1066 (2020).
39. Wilen, C.B., Tilton, J.C. & Doms, R.W. HIV: cell binding and entry. *Cold Spring Harb Perspect Med* **2**(2012).
40. Connell, B.J. & Lortat-Jacob, H. Human immunodeficiency virus and heparan sulfate: from attachment to entry inhibition. *Front Immunol* **4**, 385 (2013).
41. Gorry, P.R. & Ancuta, P. Coreceptors and HIV-1 pathogenesis. *Curr HIV/AIDS Rep* **8**, 45-53 (2011).

42. Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V. & Campbell, E.M. Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nat Microbiol* **5**, 1088-1095 (2020).
43. Sarafianos, S.G., *et al.* Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J Mol Biol* **385**, 693-713 (2009).
44. London, R.E. HIV-1 Reverse Transcriptase: A Metamorphic Protein with Three Stable States. *Structure* **27**, 420-426 (2019).
45. Xavier Ruiz, F. & Arnold, E. Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Curr Opin Struct Biol* **61**, 113-123 (2020).
46. Iordanskiy, S. & Bukrinsky, M. Reverse transcription complex: the key player of the early phase of HIV replication. *Future Virol* **2**, 49-64 (2007).
47. Hu, W.S. & Hughes, S.H. HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* **2**(2012).
48. Charneau, P., *et al.* HIV-1 reverse transcription. A termination step at the center of the genome. *J Mol Biol* **241**, 651-662 (1994).
49. De Rijck, J. & Debyser, Z. The central DNA flap of the human immunodeficiency virus type 1 is important for viral replication. *Biochem Biophys Res Commun* **349**, 1100-1110 (2006).
50. Craigie, R. & Bushman, F.D. HIV DNA integration. *Cold Spring Harb Perspect Med* **2**, a006890 (2012).
51. Krishnan, L., *et al.* The requirement for cellular transportin 3 (TNPO3 or TRN-SR2) during infection maps to human immunodeficiency virus type 1 capsid and not integrase. *J Virol* **84**, 397-406 (2010).
52. Christ, F., *et al.* Transportin-SR2 imports HIV into the nucleus. *Curr Biol* **18**, 1192-1202 (2008).
53. Ao, Z., *et al.* Interaction of human immunodeficiency virus type 1 integrase with cellular nuclear import receptor importin 7 and its impact on viral replication. *J Biol Chem* **282**, 13456-13467 (2007).
54. Jayappa, K.D., Ao, Z. & Yao, X. The HIV-1 passage from cytoplasm to nucleus: the process involving a complex exchange between the components of HIV-1 and cellular machinery to access nucleus and successful integration. *Int J Biochem Mol Biol* **3**, 70-85 (2012).
55. Suzuki, Y. & Craigie, R. The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol* **5**, 187-196 (2007).
56. Lusic, M. & Siliciano, R.F. Nuclear landscape of HIV-1 infection and integration. *Nat Rev Microbiol* **15**, 69-82 (2017).
57. Vink, C., *et al.* Analysis of the junctions between human immunodeficiency virus type 1 proviral DNA and human DNA. *J Virol* **64**, 5626-5627 (1990).
58. Lapaillerie, D., *et al.* Modulation of the intrinsic chromatin binding property of HIV-1 integrase by LEDGF/p75. *Nucleic Acids Res* **49**, 11241-11256 (2021).
59. Singh, P.K., *et al.* LEDGF/p75 interacts with mRNA splicing factors and targets HIV-1 integration to highly spliced genes. *Genes Dev* **29**, 2287-2297 (2015).
60. Gaynor, R. Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* **6**, 347-363 (1992).
61. Mbondji-Wonje, C., *et al.* Distinctive variation in the U3R region of the 5' Long Terminal Repeat from diverse HIV-1 strains. *PLoS One* **13**, e0195661 (2018).

62. Pereira, L.A., Bentley, K., Peeters, A., Churchill, M.J. & Deacon, N.J. A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter. *Nucleic Acids Res* **28**, 663-668 (2000).
63. Ne, E., Palstra, R.J. & Mahmoudi, T. Transcription: Insights From the HIV-1 Promoter. *Int Rev Cell Mol Biol* **335**, 191-243 (2018).
64. Wu, Y. & Marsh, J.W. Gene transcription in HIV infection. *Microbes Infect* **5**, 1023-1027 (2003).
65. Delannoy, A., Poirier, M. & Bell, B. Cat and Mouse: HIV Transcription in Latency, Immune Evasion and Cure/Remission Strategies. *Viruses* **11**, 269 (2019).
66. Wilhelm, E. & Bell, B. Selective recognition of viral promoters by host cell transcription complexes: challenges and opportunities to control latency. *Curr Opin Virol* **3**, 380-386 (2013).
67. Mousseau, G. & Valente, S.T. Role of Host Factors on the Regulation of Tat-Mediated HIV-1 Transcription. *Curr Pharm Des* **23**, 4079-4090 (2017).
68. Rice, A.P. The HIV-1 Tat Protein: Mechanism of Action and Target for HIV-1 Cure Strategies. *Curr Pharm Des* **23**, 4098-4102 (2017).
69. Asamitsu, K., Fujinaga, K. & Okamoto, T. HIV Tat/P-TEFb Interaction: A Potential Target for Novel Anti-HIV Therapies. *Molecules* **23**, 933 (2018).
70. Muniz, L., Egloff, S., Ughy, B., Jady, B.E. & Kiss, T. Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. *PLoS Pathog* **6**, e1001152 (2010).
71. Battisti, P.L., *et al.* Additive activity between the trans-activation response RNA-binding protein, TRBP2, and cyclin T1 on HIV type 1 expression and viral production in murine cells. *AIDS Res Hum Retroviruses* **19**, 767-778 (2003).
72. Gatignol, A. Transcription of HIV: Tat and cellular chromatin. *Adv Pharmacol* **55**, 137-159 (2007).
73. Purcell, D.F. & Martin, M.A. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol* **67**, 6365-6378 (1993).
74. Truman, C.T., Jarvelin, A., Davis, I. & Castello, A. HIV Rev-isited. *Open Biol* **10**, 200320 (2020).
75. Ajamian, L., *et al.* HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. *Biomolecules* **5**, 2808-2839 (2015).
76. Kuzembayeva, M., Dilley, K., Sardo, L. & Hu, W.S. Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology* **454-455**, 362-370 (2014).
77. Stoltzfus, C.M. Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. *Adv Virus Res* **74**, 1-40 (2009).
78. Freed, E.O. HIV-1 assembly, release and maturation. *Nat Rev Microbiol* **13**, 484-496 (2015).
79. Mailler, E., *et al.* The Life-Cycle of the HIV-1 Gag-RNA Complex. *Viruses* **8**, 248 (2016).
80. Chen, B. HIV Capsid Assembly, Mechanism, and Structure. *Biochemistry* **55**, 2539-2552 (2016).
81. Checkley, M.A., Luttge, B.G. & Freed, E.O. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J Mol Biol* **410**, 582-608 (2011).
82. Yuan, T., Li, J. & Zhang, M.Y. HIV-1 envelope glycoprotein variable loops are indispensable for envelope structural integrity and virus entry. *PLoS One* **8**, e69789 (2013).

83. Votteler, J. & Sundquist, W.I. Virus budding and the ESCRT pathway. *Cell Host Microbe* **14**, 232-241 (2013).
84. Louis, J.M., Ishima, R., Torchia, D.A. & Weber, I.T. HIV-1 protease: structure, dynamics, and inhibition. *Adv Pharmacol* **55**, 261-298 (2007).
85. Wallet, C., Rohr, O. & Schwartz, C. Evolution of a concept: From accessory protein to key virulence factor, the case of HIV-1 Vpr. *Biochem Pharmacol* **180**, 114128 (2020).
86. Shao, Q., Wang, Y., Hildreth, J.E. & Liu, B. Polyubiquitination of APOBEC3G is essential for its degradation by HIV-1 Vif. *J Virol* **84**, 4840-4844 (2010).
87. Lata, S., Mishra, R. & Banerjea, A.C. Proteasomal Degradation Machinery: Favorite Target of HIV-1 Proteins. *Front Microbiol* **9**, 2738 (2018).
88. Tungaturthi, P.K., Sawaya, B.E., Ayyavoo, V., Murali, R. & Srinivasan, A. HIV-1 Vpr: genetic diversity and functional features from the perspective of structure. *DNA Cell Biol* **23**, 207-222 (2004).
89. Zhou, D., *et al.* The HIV-1 accessory protein Vpr induces the degradation of the anti-HIV-1 agent APOBEC3G through a VprBP-mediated proteasomal pathway. *Virus Res* **195**, 25-34 (2015).
90. Mashiba, M., Collins, D.R., Terry, V.H. & Collins, K.L. Vpr overcomes macrophage-specific restriction of HIV-1 Env expression and virion production. *Cell Host Microbe* **16**, 722-735 (2014).
91. Nodder, S.B. & Gummuluru, S. Illuminating the Role of Vpr in HIV Infection of Myeloid Cells. *Front Immunol* **10**, 1606 (2019).
92. Mangeat, B., *et al.* HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation. *PLoS Pathog* **5**, e1000574 (2009).
93. Hauser, H., *et al.* HIV-1 Vpu and HIV-2 Env counteract BST-2/tetherin by sequestration in a perinuclear compartment. *Retrovirology* **7**, 51 (2010).
94. Bakir, T.M. The role of SAMHD1 expression and its relation to HIV-2 (Vpx) gene production. *Saudi Pharm J* **26**, 903-908 (2018).
95. Ryoo, J., *et al.* The ribonuclease activity of SAMHD1 is required for HIV-1 restriction. *Nature medicine* **20**, 936-941 (2014).
96. Antonucci, J.M., St Gelais, C. & Wu, L. The Dynamic Interplay between HIV-1, SAMHD1, and the Innate Antiviral Response. *Front Immunol* **8**, 1541 (2017).
97. Hrecka, K., *et al.* Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* **474**, 658-661 (2011).
98. Basmaciogullari, S. & Pizzato, M. The activity of Nef on HIV-1 infectivity. *Front Microbiol* **5**, 232 (2014).
99. Staudt, R.P., *et al.* Structure, function, and inhibitor targeting of HIV-1 Nef-effector kinase complexes. *J Biol Chem* **295**, 15158-15171 (2020).
100. Usami, Y., Wu, Y. & Gottlinger, H.G. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature* **526**, 218-223 (2015).
101. Shaw, G.M. & Hunter, E. HIV transmission. *Cold Spring Harb Perspect Med* **2**(2012).
102. Donnell, D., *et al.* Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. *Lancet* **375**, 2092-2098 (2010).
103. Rodger, A.J., *et al.* Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. *Lancet* **393**, 2428-2438 (2019).

104. Rodger, A.J., *et al.* Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy. *JAMA* **316**, 171-181 (2016).
105. Ambrosioni, J., Petit, E., Liegeon, G., Laguno, M. & Miro, J.M. Primary HIV-1 infection in users of pre-exposure prophylaxis. *Lancet HIV* **8**, e166-e174 (2021).
106. DeHaan, E., *et al.* in *PEP to Prevent HIV Infection* (Baltimore (MD), 2021).
107. Keele, B.F., *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* **105**, 7552-7557 (2008).
108. Nijmeijer, B.M. & Geijtenbeek, T.B.H. Negative and Positive Selection Pressure During Sexual Transmission of Transmitted Founder HIV-1. *Front Immunol* **10**, 1599 (2019).
109. Joseph, S.B., Swanstrom, R., Kashuba, A.D. & Cohen, M.S. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nat Rev Microbiol* **13**, 414-425 (2015).
110. Yu, H.J., Reuter, M.A. & McDonald, D. HIV traffics through a specialized, surface-accessible intracellular compartment during trans-infection of T cells by mature dendritic cells. *PLoS Pathog* **4**, e1000134 (2008).
111. Cameron, P.U., *et al.* Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells. *Science* **257**, 383-387 (1992).
112. Cohen, M.S., Shaw, G.M., McMichael, A.J. & Haynes, B.F. Acute HIV-1 Infection. *N Engl J Med* **364**, 1943-1954 (2011).
113. Ribeiro, R.M., *et al.* Estimation of the initial viral growth rate and basic reproductive number during acute HIV-1 infection. *J Virol* **84**, 6096-6102 (2010).
114. Lindback, S., *et al.* Viral dynamics in primary HIV-1 infection. Karolinska Institutet Primary HIV Infection Study Group. *AIDS* **14**, 2283-2291 (2000).
115. Vergis, E.N. & Mellors, J.W. Natural history of HIV-1 infection. *Infect Dis Clin North Am* **14**, 809-825, v-vi (2000).
116. Huang, G., Takeuchi, Y. & Korobeinikov, A. HIV evolution and progression of the infection to AIDS. *J Theor Biol* **307**, 149-159 (2012).
117. Sabin, C.A. & Lundgren, J.D. The natural history of HIV infection. *Curr Opin HIV AIDS* **8**, 311-317 (2013).
118. Okulicz, J.F., *et al.* Clinical outcomes of elite controllers, viremic controllers, and long-term nonprogressors in the US Department of Defense HIV natural history study. *J Infect Dis* **200**, 1714-1723 (2009).
119. Jiang, C., *et al.* Distinct viral reservoirs in individuals with spontaneous control of HIV-1. *Nature* **585**, 261-267 (2020).
120. Turk, G., *et al.* A Possible Sterilizing Cure of HIV-1 Infection Without Stem Cell Transplantation. *Ann Intern Med* **175**, 95-100 (2022).
121. Chun, T.W., *et al.* Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. *Proc Natl Acad Sci U S A* **95**, 8869-8873 (1998).
122. Han, Y., *et al.* Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J Virol* **78**, 6122-6133 (2004).
123. Venanzi Rullo, E., *et al.* Genetic Evidence That Naive T Cells Can Contribute Significantly to the Human Immunodeficiency Virus Intact Reservoir: Time to Re-evaluate Their Role. *Clin Infect Dis* **69**, 2236-2237 (2019).

124. Buzon, M.J., *et al.* HIV-1 persistence in CD4+ T cells with stem cell-like properties. *Nature medicine* **20**, 139-142 (2014).
125. Soriano-Sarabia, N., *et al.* Quantitation of replication-competent HIV-1 in populations of resting CD4+ T cells. *J Virol* **88**, 14070-14077 (2014).
126. Hiener, B., *et al.* Identification of Genetically Intact HIV-1 Proviruses in Specific CD4(+) T Cells from Effectively Treated Participants. *Cell Rep* **21**, 813-822 (2017).
127. Cohn, L.B., Chomont, N. & Deeks, S.G. The Biology of the HIV-1 Latent Reservoir and Implications for Cure Strategies. *Cell Host Microbe* **27**, 519-530 (2020).
128. Maldarelli, F., *et al.* HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* **345**, 179-183 (2014).
129. Wagner, T.A., *et al.* HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science* **345**, 570-573 (2014).
130. Chomont, N., *et al.* HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nature medicine* **15**, 893-900 (2009).
131. Henrich, T.J., *et al.* Human Immunodeficiency Virus Type 1 Persistence Following Systemic Chemotherapy for Malignancy. *J Infect Dis* **216**, 254-262 (2017).
132. Cesana, D., *et al.* HIV-1-mediated insertional activation of STAT5B and BACH2 trigger viral reservoir in T regulatory cells. *Nat Commun* **8**, 498 (2017).
133. Douek, D.C., *et al.* HIV preferentially infects HIV-specific CD4+ T cells. *Nature* **417**, 95-98 (2002).
134. Gantner, P., *et al.* Single-cell TCR sequencing reveals phenotypically diverse clonally expanded cells harboring inducible HIV proviruses during ART. *Nat Commun* **11**, 4089 (2020).
135. Wong, M.E., Jaworowski, A. & Hearps, A.C. The HIV Reservoir in Monocytes and Macrophages. *Front Immunol* **10**, 1435 (2019).
136. Igarashi, T., *et al.* Macrophage are the principal reservoir and sustain high virus loads in rhesus macaques after the depletion of CD4+ T cells by a highly pathogenic simian immunodeficiency virus/HIV type 1 chimera (SHIV): Implications for HIV-1 infections of humans. *Proc Natl Acad Sci U S A* **98**, 658-663 (2001).
137. Valcour, V., *et al.* Central nervous system viral invasion and inflammation during acute HIV infection. *J Infect Dis* **206**, 275-282 (2012).
138. Ragin, A.B., *et al.* Structural brain alterations can be detected early in HIV infection. *Neurology* **79**, 2328-2334 (2012).
139. Kallianpur, K.J., *et al.* HIV DNA in CD14+ reservoirs is associated with regional brain atrophy in patients naive to combination antiretroviral therapy. *AIDS* **28**, 1619-1624 (2014).
140. Gibellini, D., *et al.* HIV-1 DNA load analysis in peripheral blood lymphocytes and monocytes from naive and HAART-treated individuals. *J Infect* **56**, 219-225 (2008).
141. Cribbs, S.K., Lennox, J., Caliendo, A.M., Brown, L.A. & Guidot, D.M. Healthy HIV-1-infected individuals on highly active antiretroviral therapy harbor HIV-1 in their alveolar macrophages. *AIDS Res Hum Retroviruses* **31**, 64-70 (2015).
142. Bouchat, S. & Van Lint, C. [Viral latency of HIV-1]. *Virologie (Montrouge)* **23**, 195-210 (2019).
143. Yukl, S.A., *et al.* HIV latency in isolated patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. *Sci Transl Med* **10**(2018).

144. Li, J., *et al.* Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nat Commun* **7**, 11730 (2016).
145. Cismasiu, V.B., *et al.* BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex. *Virology* **380**, 173-181 (2008).
146. Dahabieh, M.S., Battivelli, E. & Verdin, E. Understanding HIV latency: the road to an HIV cure. *Annu Rev Med* **66**, 407-421 (2015).
147. Ruelas, D.S. & Greene, W.C. An integrated overview of HIV-1 latency. *Cell* **155**, 519-529 (2013).
148. Miller-Jensen, K., *et al.* Chromatin accessibility at the HIV LTR promoter sets a threshold for NF-kappaB mediated viral gene expression. *Integr Biol (Camb)* **4**, 661-671 (2012).
149. Shukla, A., Ramirez, N.P. & D'Orso, I. HIV-1 Proviral Transcription and Latency in the New Era. *Viruses* **12**, 555 (2020).
150. Van Lint, C., Emiliani, S., Ott, M. & Verdin, E. Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *EMBO J* **15**, 1112-1120 (1996).
151. du Chene, I., *et al.* Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J* **26**, 424-435 (2007).
152. Imai, K., Togami, H. & Okamoto, T. Involvement of histone H3 lysine 9 (H3K9) methyltransferase G9a in the maintenance of HIV-1 latency and its reactivation by BIX01294. *J Biol Chem* **285**, 16538-16545 (2010).
153. Kauder, S.E., Bosque, A., Lindqvist, A., Planelles, V. & Verdin, E. Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog* **5**, e1000495 (2009).
154. Broder, S. The development of antiretroviral therapy and its impact on the HIV-1/AIDS pandemic. *Antiviral Res* **85**, 1-18 (2010).
155. De Clercq, E. The nucleoside reverse transcriptase inhibitors, nonnucleoside reverse transcriptase inhibitors, and protease inhibitors in the treatment of HIV infections (AIDS). *Adv Pharmacol* **67**, 317-358 (2013).
156. Laskey, S.B. & Siliciano, R.F. A mechanistic theory to explain the efficacy of antiretroviral therapy. *Nat Rev Microbiol* **12**, 772-780 (2014).
157. Hammer, S.M., *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* **337**, 725-733 (1997).
158. Atta, M.G., De Seigneux, S. & Lucas, G.M. Clinical Pharmacology in HIV Therapy. *Clin J Am Soc Nephrol* **14**, 435-444 (2019).
159. Chandiwana, N.C., *et al.* Impact of long-acting therapies on the global HIV epidemic. *AIDS* **35**, S137-S143 (2021).
160. Phanuphak, N. & Gulick, R.M. HIV treatment and prevention 2019: current standards of care. *Curr Opin HIV AIDS* **15**, 4-12 (2020).
161. Pennings, P.S. HIV Drug Resistance: Problems and Perspectives. *Infect Dis Rep* **5**, e5 (2013).
162. Cohen, M.S., *et al.* Antiretroviral Therapy for the Prevention of HIV-1 Transmission. *N Engl J Med* **375**, 830-839 (2016).

163. Saez-Cirion, A., *et al.* Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS Pathog* **9**, e1003211 (2013).
164. Martin, G.E. & Frater, J. Post-treatment and spontaneous HIV control. *Curr Opin HIV AIDS* **13**, 402-407 (2018).
165. Maggiolo, F., Di Filippo, E., Comi, L. & Callegaro, A. Post-treatment controllers after treatment interruption in chronically HIV-infected patients. *AIDS* **32**, 623-628 (2018).
166. Choopanya, K., *et al.* Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet* **381**, 2083-2090 (2013).
167. Thomas, J., Ruggiero, A., Paxton, W.A. & Pollakis, G. Measuring the Success of HIV-1 Cure Strategies. *Front Cell Infect Microbiol* **10**, 134 (2020).
168. Frange, P., *et al.* HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. *Lancet HIV* **3**, e49-54 (2016).
169. Persaud, D., *et al.* Absence of detectable HIV-1 viremia after treatment cessation in an infant. *N Engl J Med* **369**, 1828-1835 (2013).
170. Luzuriaga, K., *et al.* Viremic relapse after HIV-1 remission in a perinatally infected child. *N Engl J Med* **372**, 786-788 (2015).
171. Scarborough, R.J., Goguen, R.P. & Gatignol, A. A second patient cured of HIV infection: hopes and limitations. *Virologie (Montrouge)* **23**, 1-4 (2019).
172. Liu, R., *et al.* Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367-377 (1996).
173. Samson, M., *et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**, 722-725 (1996).
174. Kalidasan, V. & Theva Das, K. Lessons Learned From Failures and Success Stories of HIV Breakthroughs: Are We Getting Closer to an HIV Cure? *Front Microbiol* **11**, 46 (2020).
175. Jensen, B.O., *et al.* In-depth virological and immunological characterization of HIV-1 cure after CCR5Delta32/Delta32 allogeneic hematopoietic stem cell transplantation. *Nature medicine* **29**, 583-587 (2023).
176. Hutter, G., *et al.* Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* **360**, 692-698 (2009).
177. Gupta, R.K., *et al.* HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* **568**, 244-248 (2019).
178. Dickter, J., *et al.* The 'City of Hope' Patient: prolonged HIV-1 remission without antiretrovirals (ART) after allogeneic hematopoietic stem cell transplantation (aHCT) of CCR5Δ32/Δ32 donor cells for acute myelogenous leukemia (AML). AIDS 2022 Abstract. https://programme.aids2022.org/Abstract/Abstract/?abstractid=12508. (2022).
179. Hsu, J.M., *et al.* HIV-1 remission with CCR5Δ32Δ32 haplo-cord transplant in a U.S. woman: IMPAACT P1107. [Abstract 65]. In Special Issue: Abstracts from the 2022 Conference on Retroviruses and Opportunistic Infections. *Top Antiv Med* **30**, 23 (2022).
180. Novembre, J., Galvani, A.P. & Slatkin, M. The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol* **3**, e339 (2005).
181. Hutter, G. More on shift of HIV tropism in stem-cell transplantation with CCR5 delta32/delta32 mutation. *N Engl J Med* **371**, 2437-2438 (2014).

182. Rodari, A., Darcis, G. & Van Lint, C.M. The Current Status of Latency Reversing Agents for HIV-1 Remission. *Annu Rev Virol* **8**, 491-514 (2021).
183. Bouchat, S., *et al.* Sequential treatment with 5-aza-2'-deoxycytidine and deacetylase inhibitors reactivates HIV-1. *EMBO Mol Med* **8**, 117-138 (2016).
184. Williams, S.A., *et al.* Prostratin antagonizes HIV latency by activating NF-kappaB. *J Biol Chem* **279**, 42008-42017 (2004).
185. Novis, C.L., *et al.* Reactivation of latent HIV-1 in central memory CD4(+) T cells through TLR-1/2 stimulation. *Retrovirology* **10**, 119 (2013).
186. Doyon, G., Zerbato, J., Mellors, J.W. & Sluis-Cremer, N. Disulfiram reactivates latent HIV-1 expression through depletion of the phosphatase and tensin homolog. *AIDS* **27**, F7-F11 (2013).
187. Zhu, J., *et al.* Reactivation of latent HIV-1 by inhibition of BRD4. *Cell Rep* **2**, 807-816 (2012).
188. Bartholomeeusen, K., Xiang, Y., Fujinaga, K. & Peterlin, B.M. Bromodomain and extra-terminal (BET) bromodomain inhibition activate transcription via transient release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein. *J Biol Chem* **287**, 36609-36616 (2012).
189. Abner, E. & Jordan, A. HIV "shock and kill" therapy: In need of revision. *Antiviral Res* **166**, 19-34 (2019).
190. Gama, L., *et al.* Reactivation of simian immunodeficiency virus reservoirs in the brain of virally suppressed macaques. *AIDS* **31**, 5-14 (2017).
191. Wu, G., *et al.* HDAC inhibition induces HIV-1 protein and enables immune-based clearance following latency reversal. *JCI Insight* **2**(2017).
192. Sung, J.A., *et al.* Vorinostat Renders the Replication-Competent Latent Reservoir of Human Immunodeficiency Virus (HIV) Vulnerable to Clearance by CD8 T Cells. *EBioMedicine* **23**, 52-58 (2017).
193. Jamaluddin, M.S., Hu, P.W., Jan, Y., Siwak, E.B. & Rice, A.P. Short Communication: The Broad-Spectrum Histone Deacetylase Inhibitors Vorinostat and Panobinostat Activate Latent HIV in CD4(+) T Cells In Part Through Phosphorylation of the T-Loop of the CDK9 Subunit of P-TEFb. *AIDS Res Hum Retroviruses* **32**, 169-173 (2016).
194. Rasmussen, T.A., *et al.* Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *Lancet HIV* **1**, e13-21 (2014).
195. Marban, C., *et al.* Targeting the Brain Reservoirs: Toward an HIV Cure. *Front Immunol* **7**, 397 (2016).
196. Chun, T.W., *et al.* Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy. *J Infect Dis* **197**, 714-720 (2008).
197. Damouche, A., *et al.* Adipose Tissue Is a Neglected Viral Reservoir and an Inflammatory Site during Chronic HIV and SIV Infection. *PLoS Pathog* **11**, e1005153 (2015).
198. Sacha, J.B. & Ndhlovu, L.C. Strategies to target non-T-cell HIV reservoirs. *Curr Opin HIV AIDS* **11**, 376-382 (2016).
199. Rocco, J., Mellors, J.W. & Macatangay, B.J. Regulatory T cells: the ultimate HIV reservoir? *J Virus Erad* **4**, 209-214 (2018).
200. Brockman, M.A., Jones, R.B. & Brumme, Z.L. Challenges and Opportunities for T-Cell-Mediated Strategies to Eliminate HIV Reservoirs. *Front Immunol* **6**, 506 (2015).

201. Halper-Stromberg, A., *et al.* Broadly neutralizing antibodies and viral inducers decrease rebound from HIV-1 latent reservoirs in humanized mice. *Cell* **158**, 989-999 (2014).
202. Limsirichai, P., Gaj, T. & Schaffer, D.V. CRISPR-mediated Activation of Latent HIV-1 Expression. *Mol Ther* **24**, 499-507 (2016).
203. Vansant, G., Bruggemans, A., Janssens, J. & Debyser, Z. Block-And-Lock Strategies to Cure HIV Infection. *Viruses* **12**, 84 (2020).
204. Da Rocha, M., Poirier, M., Bonhamn, B. & Bell, B. Bloquer la transcription du VIH pour verrouiller le virus. *Virologie (Montrouge)* **26**, 23-40 (2022).
205. Maina, E.K., Adan, A.A., Mureithi, H., Muriuki, J. & Lwembe, R.M. A Review of Current Strategies Towards the Elimination of Latent HIV-1 and Subsequent HIV-1 Cure. *Curr HIV Res* **19**, 14-26 (2021).
206. Ahlenstiel, C., *et al.* Novel RNA Duplex Locks HIV-1 in a Latent State via Chromatin-mediated Transcriptional Silencing. *Mol Ther Nucleic Acids* **4**, e261 (2015).
207. Janowski, B.A., *et al.* Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat Struct Mol Biol* **13**, 787-792 (2006).
208. Kim, D.H., Villeneuve, L.M., Morris, K.V. & Rossi, J.J. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* **13**, 793-797 (2006).
209. Hawkins, P.G., Santoso, S., Adams, C., Anest, V. & Morris, K.V. Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells. *Nucleic Acids Res* **37**, 2984-2995 (2009).
210. Suzuki, K., *et al.* Closed chromatin architecture is induced by an RNA duplex targeting the HIV-1 promoter region. *J Biol Chem* **283**, 23353-23363 (2008).
211. Castanotto, D., *et al.* Short hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells. *Mol Ther* **12**, 179-183 (2005).
212. Mendez, C., Ledger, S., Petoumenos, K., Ahlenstiel, C. & Kelleher, A.D. RNA-induced epigenetic silencing inhibits HIV-1 reactivation from latency. *Retrovirology* **15**, 67 (2018).
213. Higaki, K., *et al.* Generation of HIV-Resistant Macrophages from IPSCs by Using Transcriptional Gene Silencing and Promoter-Targeted RNA. *Mol Ther Nucleic Acids* **12**, 793-804 (2018).
214. Mousseau, G., *et al.* An analog of the natural steroidal alkaloid cortistatin A potently suppresses Tat-dependent HIV transcription. *Cell Host Microbe* **12**, 97-108 (2012).
215. Mousseau, G., *et al.* The Tat Inhibitor Didehydro-Cortistatin A Prevents HIV-1 Reactivation from Latency. *mBio* **6**, e00465 (2015).
216. Kessing, C.F., *et al.* In Vivo Suppression of HIV Rebound by Didehydro-Cortistatin A, a "Block-and-Lock" Strategy for HIV-1 Treatment. *Cell Rep* **21**, 600-611 (2017).
217. Kyei, G.B., *et al.* Splicing Factor 3B Subunit 1 Interacts with HIV Tat and Plays a Role in Viral Transcription and Reactivation from Latency. *mBio* **9**(2018).
218. Leoz, M., *et al.* HEXIM1-Tat chimera inhibits HIV-1 replication. *PLoS Pathog* **14**, e1007402 (2018).
219. Jadlowsky, J.K., Nojima, M., Okamoto, T. & Fujinaga, K. Dominant negative mutant cyclin T1 proteins that inhibit HIV transcription by forming a kinase inactive complex with Tat. *J Gen Virol* **89**, 2783-2787 (2008).
220. Meredith, L.W., Sivakumaran, H., Major, L., Suhrbier, A. & Harrich, D. Potent inhibition of HIV-1 replication by a Tat mutant. *PLoS One* **4**, e7769 (2009).

221. Wan, Z. & Chen, X. Triptolide inhibits human immunodeficiency virus type 1 replication by promoting proteasomal degradation of Tat protein. *Retrovirology* **11**, 88 (2014).
222. Paul, R., Dutta, D., Paul, R. & Dash, J. Target-Directed Azide-Alkyne Cycloaddition for Assembling HIV-1 TAR RNA Binding Ligands. *Angew Chem Int Ed Engl* **59**, 12407-12411 (2020).
223. Lacombe, B., Morel, M., Margottin-Goguet, F. & Ramirez, B.C. Specific Inhibition of HIV Infection by the Action of Spironolactone in T Cells. *J Virol* **90**, 10972-10980 (2016).
224. Schonhofer, C., *et al.* Flavonoid-based inhibition of cyclin-dependent kinase 9 without concomitant inhibition of histone deacetylases durably reinforces HIV latency. *Biochem Pharmacol* **186**, 114462 (2021).
225. Jean, M.J., *et al.* Curaxin CBL0100 Blocks HIV-1 Replication and Reactivation through Inhibition of Viral Transcriptional Elongation. *Front Microbiol* **8**, 2007 (2017).
226. Zhou, C.L., *et al.* A New Small-Molecule Compound, Q308, Silences Latent HIV-1 Provirus by Suppressing Tat- and FACT-Mediated Transcription. *Antimicrob Agents Chemother* **65**, e0047021 (2021).
227. Kim, H., Choi, M.S., Inn, K.S. & Kim, B.J. Inhibition of HIV-1 reactivation by a telomerase-derived peptide in a HSP90-dependent manner. *Sci Rep* **6**, 28896 (2016).
228. Besnard, E., *et al.* The mTOR Complex Controls HIV Latency. *Cell Host Microbe* **20**, 785-797 (2016).
229. Margolis, D.M., *et al.* Curing HIV: Seeking to Target and Clear Persistent Infection. *Cell* **181**, 189-206 (2020).
230. Kim, J., Vasan, S., Kim, J.H. & Ake, J.A. Current approaches to HIV vaccine development: a narrative review. *J Int AIDS Soc* **24 Suppl 7**, e25793 (2021).
231. Quinones-Kochs, M.I., Buonocore, L. & Rose, J.K. Role of N-linked glycans in a human immunodeficiency virus envelope glycoprotein: effects on protein function and the neutralizing antibody response. *J Virol* **76**, 4199-4211 (2002).
232. Pitisuttithum, P., *et al.* Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. *J Infect Dis* **194**, 1661-1671 (2006).
233. Flynn, N.M., *et al.* Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. *J Infect Dis* **191**, 654-665 (2005).
234. Buchbinder, S.P., *et al.* Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet* **372**, 1881-1893 (2008).
235. Gray, G.E., *et al.* Safety and efficacy of the HVTN 503/Phambili study of a clade-B-based HIV-1 vaccine in South Africa: a double-blind, randomised, placebo-controlled test-of-concept phase 2b study. *Lancet Infect Dis* **11**, 507-515 (2011).
236. Hammer, S.M., *et al.* Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. *N Engl J Med* **369**, 2083-2092 (2013).
237. Rerks-Ngarm, S., *et al.* Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. *N Engl J Med* **361**, 2209-2220 (2009).
238. Rerks-Ngarm, S., *et al.* Randomized, Double-Blind Evaluation of Late Boost Strategies for HIV-Uninfected Vaccine Recipients in the RV144 HIV Vaccine Efficacy Trial. *J Infect Dis* **215**, 1255-1263 (2017).

239. Easterhoff, D., *et al.* Boosting of HIV envelope CD4 binding site antibodies with long variable heavy third complementarity determining region in the randomized double blind RV305 HIV-1 vaccine trial. *PLoS Pathog* **13**, e1006182 (2017).
240. Bekker, L.G., *et al.* Subtype C ALVAC-HIV and bivalent subtype C gp120/MF59 HIV-1 vaccine in low-risk, HIV-uninfected, South African adults: a phase 1/2 trial. *Lancet HIV* **5**, e366-e378 (2018).
241. Fischer, W., *et al.* Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nature medicine* **13**, 100-106 (2007).
242. Moodie, Z., *et al.* Antibody and cellular responses to HIV vaccine regimens with DNA plasmid as compared with ALVAC priming: An analysis of two randomized controlled trials. *PLoS Med* **17**, e1003117 (2020).
243. Zhang, P., *et al.* A multiclade env-gag VLP mRNA vaccine elicits tier-2 HIV-1-neutralizing antibodies and reduces the risk of heterologous SHIV infection in macaques. *Nature medicine* **27**, 2234-2245 (2021).
244. Mehandru, S., *et al.* Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection. *J Virol* **81**, 11016-11031 (2007).
245. Trkola, A., *et al.* Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nature medicine* **11**, 615-622 (2005).
246. Caskey, M., *et al.* Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522**, 487-491 (2015).
247. Lynch, R.M., *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci Transl Med* **7**, 319ra206 (2015).
248. Ferrari, G., *et al.* Envelope-specific antibodies and antibody-derived molecules for treating and curing HIV infection. *Nat Rev Drug Discov* **15**, 823-834 (2016).
249. Leth, S., *et al.* Combined effect of Vacc-4x, recombinant human granulocyte macrophage colony-stimulating factor vaccination, and romidepsin on the HIV-1 reservoir (REDUC): a single-arm, phase 1B/2A trial. *Lancet HIV* **3**, e463-472 (2016).
250. Borducchi, E.N., *et al.* Antibody and TLR7 agonist delay viral rebound in SHIV-infected monkeys. *Nature* **563**, 360-364 (2018).
251. Rossi, J.J., June, C.H. & Kohn, D.B. Genetic therapies against HIV. *Nat Biotechnol* **25**, 1444-1454 (2007).
252. Baltimore, D. Gene therapy. Intracellular immunization. *Nature* **335**, 395-396 (1988).
253. Tebas, P., *et al.* Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* **370**, 901-910 (2014).
254. Xu, L., *et al.* CRISPR/Cas9-Mediated CCR5 Ablation in Human Hematopoietic Stem/Progenitor Cells Confers HIV-1 Resistance In Vivo. *Mol Ther* **25**, 1782-1789 (2017).
255. Kang, H., *et al.* CCR5 Disruption in Induced Pluripotent Stem Cells Using CRISPR/Cas9 Provides Selective Resistance of Immune Cells to CCR5-tropic HIV-1 Virus. *Mol Ther Nucleic Acids* **4**, e268 (2015).
256. Qi, C., *et al.* Inducing CCR5Delta32/Delta32 Homozygotes in the Human Jurkat CD4+ Cell Line and Primary CD4+ Cells by CRISPR-Cas9 Genome-Editing Technology. *Mol Ther Nucleic Acids* **12**, 267-274 (2018).
257. Kordelas, L., *et al.* Shift of HIV tropism in stem-cell transplantation with CCR5 Delta32 mutation. *N Engl J Med* **371**, 880-882 (2014).

258. Verheyen, J., *et al.* Rapid Rebound of a Preexisting CXCR4-tropic Human Immunodeficiency Virus Variant After Allogeneic Transplantation With CCR5 Delta32 Homozygous Stem Cells. *Clin Infect Dis* **68**, 684-687 (2019).
259. Chang, H.H.Y., Pannunzio, N.R., Adachi, N. & Lieber, M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* **18**, 495-506 (2017).
260. Azhagiri, M.K.K., Babu, P., Venkatesan, V. & Thangavel, S. Homology-directed gene-editing approaches for hematopoietic stem and progenitor cell gene therapy. *Stem Cell Res Ther* **12**, 500 (2021).
261. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646 (2010).
262. Wolfe, S.A., Nekludova, L. & Pabo, C.O. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* **29**, 183-212 (2000).
263. Urnov, F.D., *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646-651 (2005).
264. Perez, E.E., *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**, 808-816 (2008).
265. Holt, N., *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol* **28**, 839-847 (2010).
266. Didigu, C.A., *et al.* Simultaneous zinc-finger nuclease editing of the HIV coreceptors ccr5 and cxcr4 protects CD4+ T cells from HIV-1 infection. *Blood* **123**, 61-69 (2014).
267. Cathomen, T. & Keith Joung, J. Zinc-finger Nucleases: The Next Generation Emerges. *Mol Ther* **16**, 1200-1207 (2008).
268. Bhardwaj, A. & Nain, V. TALENs-an indispensable tool in the era of CRISPR: a mini review. *J Genet Eng Biotechnol* **19**, 125 (2021).
269. Mussolino, C., *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res* **39**, 9283-9293 (2011).
270. Mock, U., *et al.* mRNA transfection of a novel TAL effector nuclease (TALEN) facilitates efficient knockout of HIV co-receptor CCR5. *Nucleic Acids Res* **43**, 5560-5571 (2015).
271. Shi, B., *et al.* TALEN-Mediated Knockout of CCR5 Confers Protection Against Infection of Human Immunodeficiency Virus. *J Acquir Immune Defic Syndr* **74**, 229-241 (2017).
272. Barrangou, R., *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712 (2007).
273. Doudna, J.A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
274. Jinek, M., *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
275. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278 (2014).
276. Li, C., *et al.* Inhibition of HIV-1 infection of primary CD4+ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J Gen Virol* **96**, 2381-2393 (2015).
277. Ebina, H., Misawa, N., Kanemura, Y. & Koyanagi, Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep* **3**, 2510 (2013).
278. Hu, W., *et al.* RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc Natl Acad Sci U S A* **111**, 11461-11466 (2014).

279. Zhu, W., *et al.* The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology* **12**, 22 (2015).
280. Kaminski, R., *et al.* Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. *Sci Rep* **6**, 22555 (2016).
281. Trono, D., Feinberg, M.B. & Baltimore, D. HIV-1 Gag mutants can dominantly interfere with the replication of the wild-type virus. *Cell* **59**, 113-120 (1989).
282. Furuta, R.A., *et al.* HIV-1 capsid mutants inhibit the replication of wild-type virus at both early and late infection phases. *FEBS Lett* **415**, 231-234 (1997).
283. Lee, S.K., Harris, J. & Swanstrom, R. A strongly transdominant mutation in the human immunodeficiency virus type 1 gag gene defines an Achilles heel in the virus life cycle. *J Virol* **83**, 8536-8543 (2009).
284. Checkley, M.A., Luttge, B.G., Soheilian, F., Nagashima, K. & Freed, E.O. The capsid-spacer peptide 1 Gag processing intermediate is a dominant-negative inhibitor of HIV-1 maturation. *Virology* **400**, 137-144 (2010).
285. Malim, M.H., Bohnlein, S., Hauber, J. & Cullen, B.R. Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. *Cell* **58**, 205-214 (1989).
286. Plavec, I., *et al.* High transdominant RevM10 protein levels are required to inhibit HIV-1 replication in cell lines and primary T cells: implication for gene therapy of AIDS. *Gene Ther* **4**, 128-139 (1997).
287. Woffendin, C., Ranga, U., Yang, Z., Xu, L. & Nabel, G.J. Expression of a protective gene-prolongs survival of T cells in human immunodeficiency virus-infected patients. *Proc Natl Acad Sci U S A* **93**, 2889-2894 (1996).
288. Jin, H., *et al.* Strong In Vivo Inhibition of HIV-1 Replication by Nullbasic, a Tat Mutant. *mBio* **10**(2019).
289. Porcellini, S., *et al.* The F12-Vif derivative Chim3 inhibits HIV-1 replication in CD4+ T lymphocytes and CD34+-derived macrophages by blocking HIV-1 DNA integration. *Blood* **113**, 3443-3452 (2009).
290. Walker, R.C., Jr., *et al.* Identification of dominant negative human immunodeficiency virus type 1 Vif mutants that interfere with the functional inactivation of APOBEC3G by virus-encoded Vif. *J Virol* **84**, 5201-5211 (2010).
291. Fackler, O.T., d'Aloja, P., Baur, A.S., Federico, M. & Peterlin, B.M. Nef from human immunodeficiency virus type 1(F12) inhibits viral production and infectivity. *J Virol* **75**, 6601-6608 (2001).
292. Stremlau, M., *et al.* The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature* **427**, 848-853 (2004).
293. Stremlau, M., *et al.* Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A* **103**, 5514-5519 (2006).
294. Sawyer, S.L., Wu, L.I., Emerman, M. & Malik, H.S. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci U S A* **102**, 2832-2837 (2005).
295. Anderson, J. & Akkina, R. Human immunodeficiency virus type 1 restriction by human-rhesus chimeric tripartite motif 5alpha (TRIM 5alpha) in CD34(+) cell-derived macrophages in vitro and in T cells in vivo in severe combined immunodeficient (SCID-hu) mice transplanted with human fetal tissue. *Hum Gene Ther* **19**, 217-228 (2008).

296. Yang, A.G., Bai, X., Huang, X.F., Yao, C. & Chen, S. Phenotypic knockout of HIV type 1 chemokine coreceptor CCR-5 by intrakines as potential therapeutic approach for HIV-1 infection. *Proc Natl Acad Sci U S A* **94**, 11567-11572 (1997).
297. Schroers, R., Davis, C.M., Wagner, H.J. & Chen, S.Y. Lentiviral transduction of human T-lymphocytes with a RANTES intrakine inhibits human immunodeficiency virus type 1 infection. *Gene Ther* **9**, 889-897 (2002).
298. Melikyan, G.B., Egelhofer, M. & von Laer, D. Membrane-anchored inhibitory peptides capture human immunodeficiency virus type 1 gp41 conformations that engage the target membrane prior to fusion. *J Virol* **80**, 3249-3258 (2006).
299. van Lunzen, J., *et al.* Transfer of autologous gene-modified T cells in HIV-infected patients with advanced immunodeficiency and drug-resistant virus. *Mol Ther* **15**, 1024-1033 (2007).
300. DiGiusto, D.L., *et al.* Preclinical development and qualification of ZFN-mediated CCR5 disruption in human hematopoietic stem/progenitor cells. *Mol Ther Methods Clin Dev* **3**, 16067 (2016).
301. Ranga, U., *et al.* Enhanced T cell engraftment after retroviral delivery of an antiviral gene in HIV-infected individuals. *Proc Natl Acad Sci U S A* **95**, 1201-1206 (1998).
302. Podsakoff, G.M., *et al.* Selective survival of peripheral blood lymphocytes in children with HIV-1 following delivery of an anti-HIV gene to bone marrow CD34(+) cells. *Mol Ther* **12**, 77-86 (2005).
303. Morgan, R.A., *et al.* Preferential survival of CD4+ T lymphocytes engineered with anti-human immunodeficiency virus (HIV) genes in HIV-infected individuals. *Hum Gene Ther* **16**, 1065-1074 (2005).
304. Kang, E.M., *et al.* Nonmyeloablative conditioning followed by transplantation of genetically modified HLA-matched peripheral blood progenitor cells for hematologic malignancies in patients with acquired immunodeficiency syndrome. *Blood* **99**, 698-701 (2002).
305. Hayakawa, J., *et al.* Long-term vector integration site analysis following retroviral mediated gene transfer to hematopoietic stem cells for the treatment of HIV infection. *PLoS One* **4**, e4211 (2009).
306. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).
307. Wong-Staal, F., Poeschla, E.M. & Looney, D.J. A controlled, Phase 1 clinical trial to evaluate the safety and effects in HIV-1 infected humans of autologous lymphocytes transduced with a ribozyme that cleaves HIV-1 RNA. *Hum Gene Ther* **9**, 2407-2425 (1998).
308. Macpherson, J.L., *et al.* Long-term survival and concomitant gene expression of ribozyme-transduced CD4+ T-lymphocytes in HIV-infected patients. *J Gene Med* **7**, 552-564 (2005).
309. Amado, R.G., *et al.* Anti-human immunodeficiency virus hematopoietic progenitor cell-delivered ribozyme in a phase I study: myeloid and lymphoid reconstitution in human immunodeficiency virus type-1-infected patients. *Hum Gene Ther* **15**, 251-262 (2004).
310. Amado, R.G., *et al.* A phase I trial of autologous CD34+ hematopoietic progenitor cells transduced with an anti-HIV ribozyme. *Hum Gene Ther* **10**, 2255-2270 (1999).
311. Mitsuyasu, R.T., *et al.* Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nature medicine* **15**, 285-292 (2009).

312. Kohn, D.B., *et al.* A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children. *Blood* **94**, 368-371 (1999).
313. Fire, A., *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811 (1998).
314. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).
315. Hu, B., *et al.* Therapeutic siRNA: state of the art. *Signal Transduct Target Ther* **5**, 101 (2020).
316. Setten, R.L., Rossi, J.J. & Han, S.P. The current state and future directions of RNAi-based therapeutics. *Nat Rev Drug Discov* **18**, 421-446 (2019).
317. Alles, J., *et al.* An estimate of the total number of true human miRNAs. *Nucleic Acids Res* **47**, 3353-3364 (2019).
318. Brennecke, J., Stark, A., Russell, R.B. & Cohen, S.M. Principles of microRNA-target recognition. *PLoS Biol* **3**, e85 (2005).
319. Lim, L.P., *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-773 (2005).
320. Lee, Y., *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060 (2004).
321. Lin, S.L., Miller, J.D. & Ying, S.Y. Intronic microRNA (miRNA). *J Biomed Biotechnol* **2006**, 26818 (2006).
322. Saliminejad, K., Khorram Khorshid, H.R., Soleymani Fard, S. & Ghaffari, S.H. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *J Cell Physiol* **234**, 5451-5465 (2019).
323. Garcia-Lopez, J., Brieno-Enriquez, M.A. & Del Mazo, J. MicroRNA biogenesis and variability. *Biomol Concepts* **4**, 367-380 (2013).
324. Yi, R., Qin, Y., Macara, I.G. & Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-3016 (2003).
325. Daniels, S.M. & Gatignol, A. The multiple functions of TRBP, at the hub of cell responses to viruses, stress, and cancer. *Microbiol Mol Biol Rev* **76**, 652-666 (2012).
326. Eichhorn, S.W., *et al.* mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Mol Cell* **56**, 104-115 (2014).
327. Moore, C.B., Guthrie, E.H., Huang, M.T. & Taxman, D.J. Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *Methods Mol Biol* **629**, 141-158 (2010).
328. Goguen, R.P., *et al.* Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK, and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021).
329. Lambeth, L.S. & Smith, C.A. Short hairpin RNA-mediated gene silencing. *Methods Mol Biol* **942**, 205-232 (2013).
330. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).

331. Singh, S., Narang, A.S. & Mahato, R.I. Subcellular fate and off-target effects of siRNA, shRNA, and miRNA. *Pharm Res* **28**, 2996-3015 (2011).
332. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366 (2001).
333. Tian, Y., *et al.* A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Mol Cell* **53**, 606-616 (2014).
334. Chendrimada, T.P., *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740-744 (2005).
335. Lau, P.W., Potter, C.S., Carragher, B. & MacRae, I.J. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* **17**, 1326-1332 (2009).
336. Macrae, I.J., *et al.* Structural basis for double-stranded RNA processing by Dicer. *Science* **311**, 195-198 (2006).
337. Lau, P.W., *et al.* The molecular architecture of human Dicer. *Nat Struct Mol Biol* **19**, 436-440 (2012).
338. Paturi, S. & Deshmukh, M.V. A Glimpse of "Dicer Biology" Through the Structural and Functional Perspective. *Front Mol Biosci* **8**, 643657 (2021).
339. Tomari, Y., Matranga, C., Haley, B., Martinez, N. & Zamore, P.D. A protein sensor for siRNA asymmetry. *Science* **306**, 1377-1380 (2004).
340. Gatignol, A., Buckler-White, A., Berkhout, B. & Jeang, K.T. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* **251**, 1597-1600 (1991).
341. Gatignol, A., Duarte, M., Daviet, L., Chang, Y.N. & Jeang, K.T. Sequential steps in Tat trans-activation of HIV-1 mediated through cellular DNA, RNA, and protein binding factors. *Gene Expr* **5**, 217-228 (1996).
342. Daviet, L., *et al.* Analysis of a binding difference between the two dsRNA-binding domains in TRBP reveals the modular function of a KR-helix motif. *Eur J Biochem* **267**, 2419-2431 (2000).
343. Laraki, G., *et al.* Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions. *RNA Biol* **5**, 92-103 (2008).
344. Gatignol, A., Buckler, C. & Jeang, K.T. Relatedness of an RNA-binding motif in human immunodeficiency virus type 1 TAR RNA-binding protein TRBP to human P1/dsI kinase and Drosophila staufen. *Mol Cell Biol* **13**, 2193-2202 (1993).
345. Daniels, S.M., *et al.* Characterization of the TRBP domain required for dicer interaction and function in RNA interference. *BMC Mol Biol* **10**, 38 (2009).
346. Chakravarthy, S., Sternberg, S.H., Kellenberger, C.A. & Doudna, J.A. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *J Mol Biol* **404**, 392-402 (2010).
347. Masliah, G., *et al.* Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains. *EMBO J* **37**(2018).
348. Benoit, M.P., *et al.* The RNA-binding region of human TRBP interacts with microRNA precursors through two independent domains. *Nucleic Acids Res* **41**, 4241-4252 (2013).
349. Wilson, R.C., *et al.* Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis. *Mol Cell* **57**, 397-407 (2015).
350. Sudhakar, A., *et al.* Phosphorylation of serine 51 in initiation factor 2 alpha (eIF2 alpha) promotes complex formation between eIF2 alpha(P) and eIF2B and causes inhibition in the guanine nucleotide exchange activity of eIF2B. *Biochemistry* **39**, 12929-12938 (2000).

351. McCormick, C. & Khaperskyy, D.A. Translation inhibition and stress granules in the antiviral immune response. *Nat Rev Immunol* **17**, 647-660 (2017).
352. Park, H., *et al.* TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proc Natl Acad Sci U S A* **91**, 4713-4717 (1994).
353. Benkirane, M., *et al.* Oncogenic potential of TAR RNA binding protein TRBP and its regulatory interaction with RNA-dependent protein kinase PKR. *EMBO J* **16**, 611-624 (1997).
354. Ong, C.L., *et al.* Low TRBP levels support an innate human immunodeficiency virus type 1 resistance in astrocytes by enhancing the PKR antiviral response. *J Virol* **79**, 12763-12772 (2005).
355. Li, S., *et al.* Molecular basis for PKR activation by PACT or dsRNA. *Proc Natl Acad Sci U S A* **103**, 10005-10010 (2006).
356. Daher, A., *et al.* TRBP control of PACT-induced phosphorylation of protein kinase R is reversed by stress. *Mol Cell Biol* **29**, 254-265 (2009).
357. Bohmert, K., *et al.* AGO1 defines a novel locus of Arabidopsis controlling leaf development. *EMBO J* **17**, 170-180 (1998).
358. Hutvagner, G. & Simard, M.J. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* **9**, 22-32 (2008).
359. Olina, A.V., Kulbachinskiy, A.V., Aravin, A.A. & Esyunina, D.M. Argonaute Proteins and Mechanisms of RNA Interference in Eukaryotes and Prokaryotes. *Biochemistry (Mosc)* **83**, 483-497 (2018).
360. Vaucheret, H., Vazquez, F., Crete, P. & Bartel, D.P. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18**, 1187-1197 (2004).
361. Peters, L. & Meister, G. Argonaute proteins: mediators of RNA silencing. *Mol Cell* **26**, 611-623 (2007).
362. Kalmykova, A.I., Klenov, M.S. & Gvozdev, V.A. Argonaute protein PIWI controls mobilization of retrotransposons in the Drosophila male germline. *Nucleic Acids Res* **33**, 2052-2059 (2005).
363. Jin, S., Zhan, J. & Zhou, Y. Argonaute proteins: structures and their endonuclease activity. *Mol Biol Rep* **48**, 4837-4849 (2021).
364. Sheu-Gruttadauria, J. & MacRae, I.J. Structural Foundations of RNA Silencing by Argonaute. *J Mol Biol* **429**, 2619-2639 (2017).
365. Frank, F., Sonenberg, N. & Nagar, B. Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818-822 (2010).
366. Yan, K.S., *et al.* Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468-474 (2003).
367. Song, J.J., Smith, S.K., Hannon, G.J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-1437 (2004).
368. Sasaki, T., Shiohama, A., Minoshima, S. & Shimizu, N. Identification of eight members of the Argonaute family in the human genome. *Genomics* **82**, 323-330 (2003).
369. Meister, G., *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15**, 185-197 (2004).
370. Liu, J., *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-1441 (2004).

371. Schirle, N.T. & MacRae, I.J. The crystal structure of human Argonaute2. *Science* **336**, 1037-1040 (2012).
372. Das, A.T., *et al.* Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J Virol* **78**, 2601-2605 (2004).
373. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol Ther* **14**, 883-892 (2006).
374. Westerhout, E.M., Ooms, M., Vink, M., Das, A.T. & Berkhout, B. HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res* **33**, 796-804 (2005).
375. Centlivre, M., *et al.* Preclinical in vivo evaluation of the safety of a multi-shRNA-based gene therapy against HIV-1. *Mol Ther Nucleic Acids* **2**, e120 (2013).
376. McIntyre, G.J., *et al.* 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* **6**, 55 (2009).
377. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
378. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. & Gatignol, A. Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp* **115**, e54486 (2016).
379. Lee, N.S., *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20**, 500-505 (2002).
380. Shimizu, S., *et al.* Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**, 8 (2009).
381. Anderson, J. & Akkina, R. CXCR4 and CCR5 shRNA transgenic CD34+ cell derived macrophages are functionally normal and resist HIV-1 infection. *Retrovirology* **2**, 53 (2005).
382. Herrera-Carrillo, E. & Berkhout, B. Novel AgoshRNA molecules for silencing of the CCR5 co-receptor for HIV-1 infection. *PLoS One* **12**, e0177935 (2017).
383. Zhou, N., Fang, J., Mukhtar, M., Acheampong, E. & Pomerantz, R.J. Inhibition of HIV-1 fusion with small interfering RNAs targeting the chemokine coreceptor CXCR4. *Gene Ther* **11**, 1703-1712 (2004).
384. De Clercq, E. AMD3100/CXCR4 Inhibitor. *Front Immunol* **6**, 276 (2015).
385. Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci U S A* **112**, E4007-4016 (2015).
386. Yi, R., Doehle, B.P., Qin, Y., Macara, I.G. & Cullen, B.R. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA* **11**, 220-226 (2005).
387. Grimm, D., *et al.* Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *J Clin Invest* **120**, 3106-3119 (2010).
388. Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H. & Williams, B.R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**, 834-839 (2003).
389. Singhania, R., Khairuddin, N., Clarke, D. & McMillan, N.A. RNA interference for the treatment of papillomavirus disease. *Open Virol J* **6**, 204-215 (2012).
390. Liu, T., Zhang, L., Joo, D. & Sun, S.C. NF-kappaB signaling in inflammation. *Signal Transduct Target Ther* **2**(2017).

391. Gopalan, V., Jarrous, N. & Krasilnikov, A.S. Chance and necessity in the evolution of RNase P. *RNA* **24**, 1-5 (2018).
392. Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz, T.A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* **289**, 905-920 (2000).
393. Riccitelli, N. & Luptak, A. HDV family of self-cleaving ribozymes. *Prog Mol Biol Transl Sci* **120**, 123-171 (2013).
394. LaRoche-Johnston, F., Monat, C., Coulombe, S. & Cousineau, B. Bacterial group II introns generate genetic diversity by circularization and trans-splicing from a population of intron-invaded mRNAs. *PLoS Genet* **14**, e1007792 (2018).
395. Kruger, K., *et al.* Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**, 147-157 (1982).
396. Esakova, O. & Krasilnikov, A.S. Of proteins and RNA: the RNase P/MRP family. *RNA* **16**, 1725-1747 (2010).
397. Cech, T.R. Structural biology. The ribosome is a ribozyme. *Science* **289**, 878-879 (2000).
398. Edgell, D.R., Chalamcharla, V.R. & Belfort, M. Learning to live together: mutualism between self-splicing introns and their hosts. *BMC Biol* **9**, 22 (2011).
399. de la Pena, M. & Garcia-Robles, I. Ubiquitous presence of the hammerhead ribozyme motif along the tree of life. *RNA* **16**, 1943-1950 (2010).
400. Webb, C.H. & Luptak, A. HDV-like self-cleaving ribozymes. *RNA Biol* **8**, 719-727 (2011).
401. Fedor, M.J. Structure and function of the hairpin ribozyme. *J Mol Biol* **297**, 269-291 (2000).
402. Ferre-D'Amare, A.R. The hairpin ribozyme. *Biopolymers* **73**, 71-78 (2004).
403. Lilley, D.M. The Varkud satellite ribozyme. *RNA* **10**, 151-158 (2004).
404. Savinov, A. & Block, S.M. Self-cleavage of the glmS ribozyme core is controlled by a fragile folding element. *Proc Natl Acad Sci U S A* **115**, 11976-11981 (2018).
405. Roth, A., *et al.* A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nat Chem Biol* **10**, 56-60 (2014).
406. Teixeira, A., *et al.* Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. *Nature* **432**, 526-530 (2004).
407. Citti, L. & Rainaldi, G. Synthetic hammerhead ribozymes as therapeutic tools to control disease genes. *Curr Gene Ther* **5**, 11-24 (2005).
408. Park, S.V., *et al.* Catalytic RNA, ribozyme, and its applications in synthetic biology. *Biotechnol Adv* **37**, 107452 (2019).
409. Bergeron, L.J. & Perreault, J.P. Target-dependent on/off switch increases ribozyme fidelity. *Nucleic Acids Res* **33**, 1240-1248 (2005).
410. Laine, S., *et al.* In vitro and in vivo cleavage of HIV-1 RNA by new SOFA-HDV ribozymes and their potential to inhibit viral replication. *RNA Biol* **8**, 343-353 (2011).
411. Rossi, J.J. Therapeutic ribozymes: principles and applications. *BioDrugs* **9**, 1-10 (1998).
412. Sun, L.Q., Wang, L., Gerlach, W.L. & Symonds, G. Target sequence-specific inhibition of HIV-1 replication by ribozymes directed to tat RNA. *Nucleic Acids Res* **23**, 2909-2913 (1995).
413. Scarborough, R.J., Levesque, M.V., Perreault, J.P. & Gatignol, A. Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* **1103**, 31-43 (2014).

414. Muller-Kuller, T., Capalbo, G., Klebba, C., Engels, J.W. & Klein, S.A. Identification and characterization of a highly efficient anti-HIV pol hammerhead ribozyme. *Oligonucleotides* **19**, 265-272 (2009).
415. Ojwang, J.O., Hampel, A., Looney, D.J., Wong-Staal, F. & Rappaport, J. Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme. *Proc Natl Acad Sci U S A* **89**, 10802-10806 (1992).
416. Cagnon, L. & Rossi, J.J. Downregulation of the CCR5 beta-chemokine receptor and inhibition of HIV-1 infection by stable VA1-ribozyme chimeric transcripts. *Antisense Nucleic Acid Drug Dev* **10**, 251-261 (2000).
417. Goguen, R.P., Gatignol, A. & Scarborough, R.J. Cloning and Detection of Aptamer-Ribozyme Conjugations. *Methods Mol Biol* **2167**, 253-267 (2021).
418. Berkhout, B., Gao, Z. & Herrera-Carrillo, E. Design and Evaluation of Guide RNA Transcripts with a 3'-Terminal HDV Ribozyme to Enhance CRISPR-Based Gene Inactivation. *Methods Mol Biol* **2167**, 205-224 (2021).
419. Berkhout, B. & Herrera-Carrillo, E. Design and Evaluation of AgoshRNAs with 3'-Terminal HDV Ribozymes to Enhance the Silencing Activity. *Methods Mol Biol* **2167**, 225-252 (2021).
420. Goguen, R.P., Chen, M.J., Dunkley, O.R.S., Gatignol, A. & Scarborough, R.J. Gene therapy to cure HIV infection. *Virologie*, In press (2023).
421. Ellington, A.D. & Szostak, J.W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822 (1990).
422. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510 (1990).
423. Nimjee, S.M., White, R.R., Becker, R.C. & Sullenger, B.A. Aptamers as Therapeutics. *Annu Rev Pharmacol Toxicol* **57**, 61-79 (2017).
424. Huang, J., *et al.* Competition-mediated pyrene-switching aptasensor: probing lysozyme in human serum with a monomer-excimer fluorescence switch. *Anal Chem* **82**, 10158-10163 (2010).
425. Shi, H., *et al.* Activatable aptamer probe for contrast-enhanced in vivo cancer imaging based on cell membrane protein-triggered conformation alteration. *Proc Natl Acad Sci U S A* **108**, 3900-3905 (2011).
426. Mehlhorn, A., Rahimi, P. & Joseph, Y. Aptamer-Based Biosensors for Antibiotic Detection: A Review. *Biosensors (Basel)* **8**(2018).
427. Zhu, G. & Chen, X. Aptamer-based targeted therapy. *Adv Drug Deliv Rev* **134**, 65-78 (2018).
428. Duclair, S., Gautam, A., Ellington, A. & Prasad, V.R. High-affinity RNA Aptamers Against the HIV-1 Protease Inhibit Both In Vitro Protease Activity and Late Events of Viral Replication. *Mol Ther Nucleic Acids* **4**, e228 (2015).
429. Lange, M.J., Nguyen, P.D.M., Callaway, M.K., Johnson, M.C. & Burke, D.H. RNA-protein interactions govern antiviral specificity and encapsidation of broad spectrum anti-HIV reverse transcriptase aptamers. *Nucleic Acids Res* **45**, 6087-6097 (2017).
430. Ramalingam, D., *et al.* RNA aptamers directed to human immunodeficiency virus type 1 Gag polyprotein bind to the matrix and nucleocapsid domains and inhibit virus production. *J Virol* **85**, 305-314 (2011).

431. Pang, K.M., Castanotto, D., Li, H., Scherer, L. & Rossi, J.J. Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. *Nucleic Acids Res* **46**, e6 (2018).
432. Sanchez-Luque, F.J., Stich, M., Manrubia, S., Briones, C. & Berzal-Herranz, A. Efficient HIV-1 inhibition by a 16 nt-long RNA aptamer designed by combining in vitro selection and in silico optimisation strategies. *Sci Rep* **4**, 6242 (2014).
433. Sullenger, B.A., Gallardo, H.F., Ungers, G.E. & Gilboa, E. Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication. *Cell* **63**, 601-608 (1990).
434. Michienzi, A., Li, S., Zaia, J.A. & Rossi, J.J. A nucleolar TAR decoy inhibitor of HIV-1 replication. *Proc Natl Acad Sci U S A* **99**, 14047-14052 (2002).
435. Bahner, I., Kearns, K., Hao, Q.L., Smogorzewska, E.M. & Kohn, D.B. Transduction of human CD34+ hematopoietic progenitor cells by a retroviral vector expressing an RRE decoy inhibits human immunodeficiency virus type 1 replication in myelomonocytic cells produced in long-term culture. *J Virol* **70**, 4352-4360 (1996).
436. Siomi, H., Shida, H., Maki, M. & Hatanaka, M. Effects of a highly basic region of human immunodeficiency virus Tat protein on nucleolar localization. *J Virol* **64**, 1803-1807 (1990).
437. Li, Y.P. Protein B23 is an important human factor for the nucleolar localization of the human immunodeficiency virus protein Tat. *J Virol* **71**, 4098-4102 (1997).
438. Michienzi, A., De Angelis, F.G., Bozzoni, I. & Rossi, J.J. A nucleolar localizing Rev binding element inhibits HIV replication. *AIDS Res Ther* **3**, 13 (2006).
439. Will, C.L. & Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**(2011).
440. Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. *Mol Ther Nucleic Acids* **18**, 815-830 (2019).
441. Wilkinson, M.E., Charenton, C. & Nagai, K. RNA Splicing by the Spliceosome. *Annu Rev Biochem* **89**, 359-388 (2020).
442. Vagner, S., Ruegsegger, U., Gunderson, S.I., Keller, W. & Mattaj, I.W. Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA* **6**, 178-188 (2000).
443. Gunderson, S.I., Polycarpou-Schwarz, M. & Mattaj, I.W. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell* **1**, 255-264 (1998).
444. Donadon, I., *et al.* Exon-specific U1 snRNAs improve ELP1 exon 20 definition and rescue ELP1 protein expression in a familial dysautonomia mouse model. *Hum Mol Genet* **27**, 2466-2476 (2018).
445. Yanaizu, M., Sakai, K., Tosaki, Y., Kino, Y. & Satoh, J.I. Small nuclear RNA-mediated modulation of splicing reveals a therapeutic strategy for a TREM2 mutation and its post-transcriptional regulation. *Sci Rep* **8**, 6937 (2018).
446. Tanner, G., *et al.* Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. *Hum Mutat* **30**, 255-263 (2009).
447. Blazquez, L., *et al.* In vitro correction of a pseudoexon-generating deep intronic mutation in LGMD2A by antisense oligonucleotides and modified small nuclear RNAs. *Hum Mutat* **34**, 1387-1395 (2013).

448. Blazquez, L., *et al.* Increased in vivo inhibition of gene expression by combining RNA interference and U1 inhibition. *Nucleic Acids Res* **40**, e8 (2012).
449. Knoepfel, S.A., Abad, A., Abad, X., Fortes, P. & Berkhout, B. Design of modified U1i molecules against HIV-1 RNA. *Antiviral Res* **94**, 208-216 (2012).
450. Sajic, R., *et al.* Use of modified U1 snRNAs to inhibit HIV-1 replication. *Nucleic Acids Res* **35**, 247-255 (2007).
451. Mandal, D., Feng, Z. & Stoltzfus, C.M. Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J Virol* **84**, 12790-12800 (2010).
452. Abad, X., Razquin, N., Abad, A. & Fortes, P. Combination of RNA interference and U1 inhibition leads to increased inhibition of gene expression. *Nucleic Acids Res* **38**, e136 (2010).
453. Chen, M.J., Gatignol, A. & Scarborough, R. The discovery and development of RNA-based therapies for treatment of HIV-1 infection. *Expert Opin Drug Discov* **18,** 163-179 (2023).
454. Blaese, R.M., *et al.* T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* **270**, 475-480 (1995).
455. Cavazzana-Calvo, M., *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-672 (2000).
456. Puumalainen, A.M., *et al.* Beta-galactosidase gene transfer to human malignant glioma in vivo using replication-deficient retroviruses and adenoviruses. *Hum Gene Ther* **9**, 1769-1774 (1998).
457. Sibbald, B. Death but one unintended consequence of gene-therapy trial. *CMAJ* **164**, 1612 (2001).
458. Hacein-Bey-Abina, S., *et al.* Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med* **346**, 1185-1193 (2002).
459. Hacein-Bey-Abina, S., *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415-419 (2003).
460. Fischer, A., Hacein-Bey-Abina, S. & Cavazzana-Calvo, M. 20 years of gene therapy for SCID. *Nat Immunol* **11**, 457-460 (2010).
461. Yin, H., *et al.* Non-viral vectors for gene-based therapy. *Nat Rev Genet* **15**, 541-555 (2014).
462. Al-Dosari, M.S. & Gao, X. Nonviral gene delivery: principle, limitations, and recent progress. *AAPS J* **11**, 671-681 (2009).
463. Santana-Armas, M.L. & Tros de Ilarduya, C. Strategies for cancer gene-delivery improvement by non-viral vectors. *Int J Pharm* **596**, 120291 (2021).
464. Wu, G.Y. & Wu, C.H. Receptor-mediated gene delivery and expression in vivo. *J Biol Chem* **263**, 14621-14624 (1988).
465. Urbiola, K., Sanmartin, C., Blanco-Fernandez, L. & Tros de Ilarduya, C. Efficient targeted gene delivery by a novel PAMAM/DNA dendriplex coated with hyaluronic acid. *Nanomedicine (Lond)* **9**, 2787-2801 (2014).
466. Baum, C., Kustikova, O., Modlich, U., Li, Z. & Fehse, B. Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors. *Hum Gene Ther* **17**, 253-263 (2006).
467. Bessis, N., GarciaCozar, F.J. & Boissier, M.C. Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. *Gene Ther* **11 Suppl 1**, S10-17 (2004).
468. Thomas, C.E., Ehrhardt, A. & Kay, M.A. Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet* **4**, 346-358 (2003).

469. Kamimura, K., Suda, T., Zhang, G. & Liu, D. Advances in Gene Delivery Systems. *Pharmaceut Med* **25**, 293-306 (2011).
470. Vetrini, F. & Ng, P. Gene therapy with helper-dependent adenoviral vectors: current advances and future perspectives. *Viruses* **2**, 1886-1917 (2010).
471. Shirley, J.L., de Jong, Y.P., Terhorst, C. & Herzog, R.W. Immune Responses to Viral Gene Therapy Vectors. *Mol Ther* **28**, 709-722 (2020).
472. Atchison, R.W., Casto, B.C. & Hammon, W.M. Adenovirus-Associated Defective Virus Particles. *Science* **149**, 754-756 (1965).
473. Chirmule, N., *et al.* Immune responses to adenovirus and adeno-associated virus in humans. *Gene Ther* **6**, 1574-1583 (1999).
474. Smith, R.H. Adeno-associated virus integration: virus versus vector. *Gene Ther* **15**, 817-822 (2008).
475. Daya, S. & Berns, K.I. Gene therapy using adeno-associated virus vectors. *Clin Microbiol Rev* **21**, 583-593 (2008).
476. Burton, E.A., Fink, D.J. & Glorioso, J.C. Gene delivery using herpes simplex virus vectors. *DNA Cell Biol* **21**, 915-936 (2002).
477. Friedman, G.K., Pressey, J.G., Reddy, A.T., Markert, J.M. & Gillespie, G.Y. Herpes simplex virus oncolytic therapy for pediatric malignancies. *Mol Ther* **17**, 1125-1135 (2009).
478. Shah, A.C., Benos, D., Gillespie, G.Y. & Markert, J.M. Oncolytic viruses: clinical applications as vectors for the treatment of malignant gliomas. *J Neurooncol* **65**, 203-226 (2003).
479. Kay, M.A., Glorioso, J.C. & Naldini, L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nature medicine* **7**, 33-40 (2001).
480. Yee, J.K., Friedmann, T. & Burns, J.C. Generation of high-titer pseudotyped retroviral vectors with very broad host range. *Methods Cell Biol* **43 Pt A**, 99-112 (1994).
481. Finkelshtein, D., Werman, A., Novick, D., Barak, S. & Rubinstein, M. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. *Proc Natl Acad Sci U S A* **110**, 7306-7311 (2013).
482. Barquinero, J., Eixarch, H. & Perez-Melgosa, M. Retroviral vectors: new applications for an old tool. *Gene Ther* **11 Suppl 1**, S3-9 (2004).
483. Cone, R.D. & Mulligan, R.C. High-efficiency gene transfer into mammalian cells: generation of helper-free recombinant retrovirus with broad mammalian host range. *Proc Natl Acad Sci U S A* **81**, 6349-6353 (1984).
484. Matrai, J., Chuah, M.K. & VandenDriessche, T. Recent advances in lentiviral vector development and applications. *Mol Ther* **18**, 477-490 (2010).
485. Levine, B.L., *et al.* Gene transfer in humans using a conditionally replicating lentiviral vector. *Proc Natl Acad Sci U S A* **103**, 17372-17377 (2006).
486. Tebas, P., *et al.* Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV. *Blood* **121**, 1524-1533 (2013).
487. Yu, S.F., *et al.* Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc Natl Acad Sci U S A* **83**, 3194-3198 (1986).
488. Cramer, P., *et al.* Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**, 337-352 (2008).

489. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**, 2593-2620 (2002).
490. Good, P.D., *et al.* Expression of small, therapeutic RNAs in human cell nuclei. *Gene Ther* **4**, 45-54 (1997).
491. Egloff, S. & Murphy, S. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**, 280-288 (2008).
492. Lewis, J.D. & Izaurralde, E. The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* **247**, 461-469 (1997).
493. Banerjee, A.K. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev* **44**, 175-205 (1980).
494. Coller, J.M., Gray, N.K. & Wickens, M.P. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev* **12**, 3226-3235 (1998).
495. Paul, C.P., *et al.* Localized expression of small RNA inhibitors in human cells. *Mol Ther* **7**, 237-247 (2003).
496. Lee, N.S., *et al.* Functional and intracellular localization properties of U6 promoter-expressed siRNAs, shRNAs, and chimeric VA1 shRNAs in mammalian cells. *RNA* **14**, 1823-1833 (2008).
497. Bertrand, E., *et al.* The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *RNA* **3**, 75-88 (1997).
498. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. A Single H1 Promoter Can Drive Both Guide RNA and Endonuclease Expression in the CRISPR-Cas9 System. *Mol Ther Nucleic Acids* **14**, 32-40 (2019).
499. Ma, H., *et al.* Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids* **3**, e161 (2014).
500. An, D.S., *et al.* Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors. *Mol Ther* **14**, 494-504 (2006).
501. Makinen, P.I., *et al.* Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* **8**, 433-441 (2006).
502. Sun, C.P., *et al.* Studies of efficacy and liver toxicity related to adeno-associated virus-mediated RNA interference. *Hum Gene Ther* **24**, 739-750 (2013).
503. Spanevello, F., *et al.* Development of Lentiviral Vectors Simultaneously Expressing Multiple siRNAs Against CCR5, vif and tat/rev Genes for an HIV-1 Gene Therapy Approach. *Mol Ther Nucleic Acids* **5**, e312 (2016).
504. ter Brake, O., *et al.* Lentiviral Vector Design for Multiple shRNA Expression and Durable HIV-1 Inhibition. *Mol Ther* **16**, 557-564 (2008).

Rationale and hypothesis

Anti-HIV-1 gene therapy has been demonstrably safe in clinical trials but there is limited evidence that current gene therapies can effectively inhibit the virus in the absence of cART. It is therefore crucial to individually optimize anti-HIV-1 RNAs to allow for the establishment of a combination gene therapy with an effective therapeutic outcome. We hypothesize that a combination of optimized antiviral RNAs will result in long-term inhibition of HIV-1, no cellular toxicity and no emergence of resistant virus.

Objectives

The general objective of this thesis was to optimize several anti-HIV-1 RNAs to maximize inhibitory capabilities while avoiding cytotoxicity, allowing for the establishment of an effective combination anti-HIV-1 gene therapy. Initially, we investigated the use of various promoters to express antiviral RNAs. Although similar work has been done by other research groups, we felt it was important to conduct our own investigations due to inconsistencies in the literature as to which promoters express the most potent therapeutic molecules. To further optimize anti-HIV-1 RNAs for gene therapy, we also explored various molecular designs including the incorporation of aptamers within the terminal loop of shRNAs to generate aptamer-shRNA chimeras. Finally, we applied our optimization work to generate effective double molecule combinations. Compared to similar work reported in the literature, we elucidate that each anti-HIV-1 RNA molecule must be independently optimized to achieve maximal effectiveness rather than following a universal rule pertaining to promoter selection or molecular design.

Objective 1: Identify the best RNA Pol III promoter for use in gene therapy.

Our first objective was to express anti-HIV-1 RNAs from the RNA Pol III promoters U6, 7SK and H1 in various cell lines to compare antiviral potency, effects on cellular viability and the metabolic profile of each molecule when expressed from the different promoters. Both transient transfection and lentivirus mediated transduction were evaluated.

*Objective 2: Identify the best molecular design of anti-HIV-*1 *RNAs.*

Our next objective was to express anti-HIV-1 aptamers without or with flanking hairpins, flanking hammerhead ribozymes, flanking hepatitis delta virus ribozymes structures as well as conjugated to shRNAs to determine which constructs are the most efficacious.

*Objective 3: Identify optimal combinations of different classes of anti-HIV-*1 *molecules to prevent viral resistance.*

Using our findings from objectives 1 & 2, we assessed the occurrence of viral escape during long-term combination therapy to be able to identify an optimal combination of antiviral molecules.

Chapter II

Efficacy, accumulation and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK and H1 promoters

This chapter is adapted from the following manuscript: Goguen, R.P., Del Corpo, O., Malard, C.M.G., Daher, A., Alpuche-Lazcano, S.P., Chen, M.J., Scarborough, R.J., Gatignol, A. 2021. Efficacy, accumulation and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034.

2.1 Preface.

Expression of shRNAs is typically mediated by the RNA polymerase III promoters H1, 7SK and U6. Transcripts produced from these promoters do not include post-transcriptional modifications which could impede the functionality of the shRNA. Several studies have evaluated the efficacy of shRNAs expressed from these promoters. Reports of which promoter produces the most active shRNAs has been conflicting, with some research groups stating that shRNAs are more active when expressed from the U6 promoter while other research groups have found maximal shRNA activity during expression from the H1 promoter. In addition, there has been a study showing no difference in shRNA activity during expression from the U6, 7SK or H1 promoter.

As the literature is inconsistent regarding which promoter is best to drive shRNA expression, we hypothesized that the nucleotide sequence of the shRNAs may have an effect as to whether expression from the U6, 7SK or H1 promoter leads to maximal shRNA activity without cytotoxicity. Our laboratory has previously designed an anti-HIV shRNA targeting the *gag* region of HIV RNA which we originally expressed from the H1 promoter. In an attempt to increase its antiviral effects, we decided to also express it from the U6 and 7SK promoter. Therefore, the main purpose of this study was to explore various expression strategies of anti-HIV shRNAs. We investigated how the expression from different promoters affects the inhibitory capabilities of shRNAs as well as their cytotoxic potential. Additionally, we characterized the transcription of shRNAs from different promoters to show that transcriptional activity and the accuracy in using the +1 transcriptional start site varies between promoters.

The results from this study allowed us to conclude that expression of anti-HIV shRNAs from the U6 and 7SK promoters did indeed lead to more potent antiviral effects, but also caused cytotoxicity due to higher expression levels of the shRNAs. The collected information in this paper provides

critical information to select appropriate expression strategies of anti-HIV RNAs for use in gene therapy to achieve a functional HIV cure.

CONFLICT OF INTEREST:

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We would like to thank Dr. John Rossi (City of Hope, Duarte, CA, USA) for provision of lentiviral plasmids. We would also like to thank Christian Young (Lady Davis Institute flow cytometry manager) for assistance in cell sorting and François Lefebvre (Canadian Centre for Computational Genomics Bioinformatics manager) for assistance in RNA sequencing work-flow design and data analysis. This work was supported by grant PJT-148704 awarded to AG and RJS from the Canadian Institutes of Health Research (CIHR). RPG is supported by a Doctoral fellowship from the Fonds de Recherche du Québec-Santé (FRQS). ODC and CMM were supported by a Frederick Banting and Charles Best Canada Graduate Scholarship-Master's (CGS-M) from the CIHR and by a MSc scholarship from the FRQS, successively. SPAL was supported by a Doctoral fellowship from the Consejo Nacional de Ciencia y Tecnologia (CONACYT) (Mexico). RJS was a recipient of the Wares Family post-doctoral award and a post-doctoral fellowship from the Richard and Edith Strauss Canadian Foundation through the McGill University Department of Medicine.

2.2 Abstract.

The expression of short hairpin RNAs (shRNAs) in cells has many potential therapeutic applications including as a functional cure for HIV. The RNA polymerase III promoters H1, 7SK and U6 have all been used to express shRNAs. However, there have been no direct and simultaneous comparisons of shRNA potency, expression level and transcriptional profile between the promoters. We show that the 7SK and U6 promoters result in higher shRNA levels and potency compared to the H1 promoter, but that in transduced T lymphocytes, higher expression levels can also lead to growth defects. We present evidence that Dicer cleavage of shRNAs is measured from the first base pair in the shRNA stem, rather than from the 5' end as previously shown for structurally related microRNAs. As a result, guide strand identity was unaffected by variations in 5' transcription start sites among the different promoters making expression levels the main determinant of shRNA potency. While all promoters generated shRNAs with variable start sites, the U6 promoter was the most accurate in using its intended +1 position. Our results have implications for the development of therapeutic small RNAs for gene therapy and for our understanding of how shRNAs are processed in cells.

2.3 Introduction.

Combination antiretroviral therapy (cART) used to treat Human Immunodeficiency Virus type 1 (HIV-1) infection involves a cocktail of several drugs targeting HIV-1 reverse transcriptase, protease, integrase and cellular entry receptors[1]. However, cART cannot clear an infection and is associated with multiple short-term and long-term side effects[2]. Alternative treatments for HIV-1 infection that do not require chronic and life-long drug administration are extremely desirable. Anti-HIV-1 RNA molecules, expressed from a gene introduced into patient cells, provide an alternative approach to treat the infection, with the potential to provide long-term control of HIV-1 replication[3-7]. The goal is to confer cellular resistance to the virus, resembling the cases of the Berlin and London patients where a hematopoietic stem cell (HSC) transplant from a donor harboring a homozygous 32 base pair deletion in the CCR5 gene, conferred resistance to HIV and cured the infection[8-10]. HSCs are the ultimate target of gene therapy as they would provide long-term inhibition of viral replication by allowing all future differentiated HIV-1 target cells to carry the viral resistance phenotype. This could be accomplished by using lentiviral vectors to transduce HSCs *ex vivo* with antiviral genes and then transplanting the cells back into the patient[4,11-14]. Antiviral small RNAs are among the top candidates for gene therapy. They include short hairpin RNAs (shRNAs), ribozymes, RNA decoys, RNA aptamers and U1 interference RNAs[15-20]. To ensure that these RNAs are expressed in an efficient and safe manner, it is important that the promoters used to express them in cells be optimized to maximize antiviral effects and minimize toxic effects.

In eukaryotic cells, three different types of RNA polymerase (Pol) promoters exist and are recognized by an equal number of RNA Pol enzymes. The interaction between enzyme and promoter mediates transcription of protein-coding mRNAs as well as a diverse array of non-coding

RNAs such as transfer RNAs (tRNAs), ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs). Specifically, the Pol I promoters drive the transcription of most rRNAs, while the Pol II promoters generate mRNAs from protein-coding genes and some small RNAs such as snRNAs and microRNAs. Finally, the Pol III promoters, which are separated into three types, transcribe exclusively non-coding RNAs[21]. The type 1 Pol III promoter transcribes the 5S rRNA gene while type 2 transcribes tRNA genes[22]. Unlike type 1 and 2 Pol III promoters that contain part of their transcripts within the promoter sequence, the type 3 Pol III promoters utilize a generally specific +1 transcriptional start site just after the end of the promoter sequence[23]. The termination sequence for RNA Pol III is a stretch of several thymidines (Ts) that results in a variable tail of one to six uridines (Us) on the 3' end of the transcript[24]. The type 3 Pol III promoters are most suited for expressing artificially designed small RNAs that require a defined 5' and 3' end, because they are the only human promoters that do not result in the addition of longer extraneous RNA sequences from the transcription start and/or termination signals. Commonly used type 3 Pol III promoters include the U6, H1 and 7SK promoters, which express the U6 snRNA, the RNase P RNA and the 7SK RNA, respectively.

Several studies have investigated the activity of shRNAs when expressed from the different Pol III promoters. However, the results have not been consistent, with three studies showing the U6 promoter produces more active shRNAs compared to the H1 promoter[25-27], another study showing the H1 promoter produces more active shRNAs compared to the U6 and 7SK promoters[28] and another study showing no difference in the activity of four different shRNAs expressed from the H1, U6 or 7SK promoters[29]. Furthermore, although it is expected that the type 3 RNA Pol III promoters would drive the expression of shRNAs from their intended +1 start sites, recent evidence demonstrates that this is not always the case. For example, one study showed that the initiation site

of transcription from human and mouse U6 promoters is affected by the surrounding sequence[30]. Further investigations have elaborated on this to identify how the precise nucleotide sequence around the +1 position affects the transcriptional efficiency of the Pol III promoters as well as which exact nucleotide position is used as the transcriptional start site[31]. Accurate expression of designed small RNAs is critical for them to perform their function and to limit potential off target effects from unintended transcripts. In the case of shRNAs, their processing by the RNA interference (RNAi) machinery takes place at specific nucleotide (nt) positions. Since it has been shown that the human Dicer endonuclease processes double stranded RNAs by measuring ~22 nt downstream from the 5' end, vastly different RNA molecules could arise when transcription is altered by even a single nt position[32,33]. Having defined start sites for the transcription of ribozymes, decoys, aptamers and clustered regularly interspaced short palindromic repeats (CRISPR) nuclease guide RNAs is also important for these molecules to function properly[34,35].

We have previously shown that an shRNA targeting a conserved sequence in HIV-1 RNA coding for the Gag polyprotein can inhibit viral replication from diverse HIV-1 strains[17]. This shRNA targets a site that begins at nucleotide position 1498 of HIV-1 NL4-3 DNA and was called sh1498. So far the only shRNA targeting HIV-1 RNA that has entered clinical trials targets the tat/rev coding region[4,36], and its target site begins at nucleotide position 5983 of HIV-1 NL4-3 DNA. In our previous study, we denoted this tat/rev shRNA as sh5983 and identified a similar anti-HIV-1 potency when compared to sh1498 using the H1 promoter to drive the expression of both shRNAs[17]. In this study, we have used the same plasmid backbone (psiRNA GFP::Zeo, InvivoGen) to directly compare the antiviral potency of sh1498 and sh5983 when expressed from the promoters H1, U6 and 7SK. As the design of vectors with the H1 and U6 promoters included eight nucleotides from the 7SK sequence directly upstream of the intended transcription start site,

we also generated plasmids with the natural H1 and U6 sequences at these positions. Our results show that the U6 and 7SK shRNA cassettes are more potent compared to the H1 shRNA cassettes and that this correlates with the expression level of the shRNAs. We also show that changing the eight nucleotides upstream of the transcription start site from the 7SK sequence to the natural H1 and U6 sequences does not enhance the cassette's potency and for the H1 promoter leads to a slight decrease in potency. Finally, by RNA sequencing (RNA Seq) analysis, we identify differences in transcriptional profiles of shRNAs expressed from the different promoters, particularly in the transcriptional start site. Usage of the expected +1 transcriptional start site was most accurate from the U6 promoter, while the accuracy in using this site from the 7SK and H1 promoters was variable based on the specific molecule being transcribed.

2.4 Results.

2.4.1 Anti-HIV-1 shRNAs are more potent when expressed from the 7SK and U6 promoters compared to the H1 promoter

To evaluate if the promoter used to express anti-HIV-1 RNAs affects the antiviral activity of these molecules, we expressed the sh1498 and sh5983 molecules from the H1, U6 and 7SK promoters. To rule-out nonspecific activity of the shRNAs contributing to the inhibition of HIV-1 production, a non-sense shRNA (shNS) that does not target HIV-1 RNA was also evaluated. The ability of the shRNAs to suppress HIV-1 production was measured after co-transfections with HIV-1 molecular clone pNL4-3 and each plasmid construct in parallel with the corresponding empty plasmid, psiRNA-H1GFP::Zeo, psiRNA-7SKGFP::Zeo or psiRNA-U6GFP::Zeo. Relative HIV-1 production was estimated by measuring the activity of HIV-1 reverse transcriptase (RT) in the cell supernatants normalized to RT activity in the supernatants of cells co-transfected with the empty

psiRNA vectors (Figure 2.1). The shRNAs expressed from the H1 promoter had an IC_{50} between 0.5-5ng (Figure 2.1A) while the IC_{50} of the shRNAs expressed from both the 7SK and U6 promoter were about 10-fold lower, between 0.05-0.5ng (Figure 2.1B, C). These results indicate that the 7SK and U6 shRNA cassettes are more potent compared to the H1 shRNA cassettes. HIV-1 production was not significantly affected by the presence of shNS when expressed from any promoter (Figure 2.1A-C) and there were no major differences in HIV-1 production in cells co-transfected with the different empty plasmids (Figure 2.2).

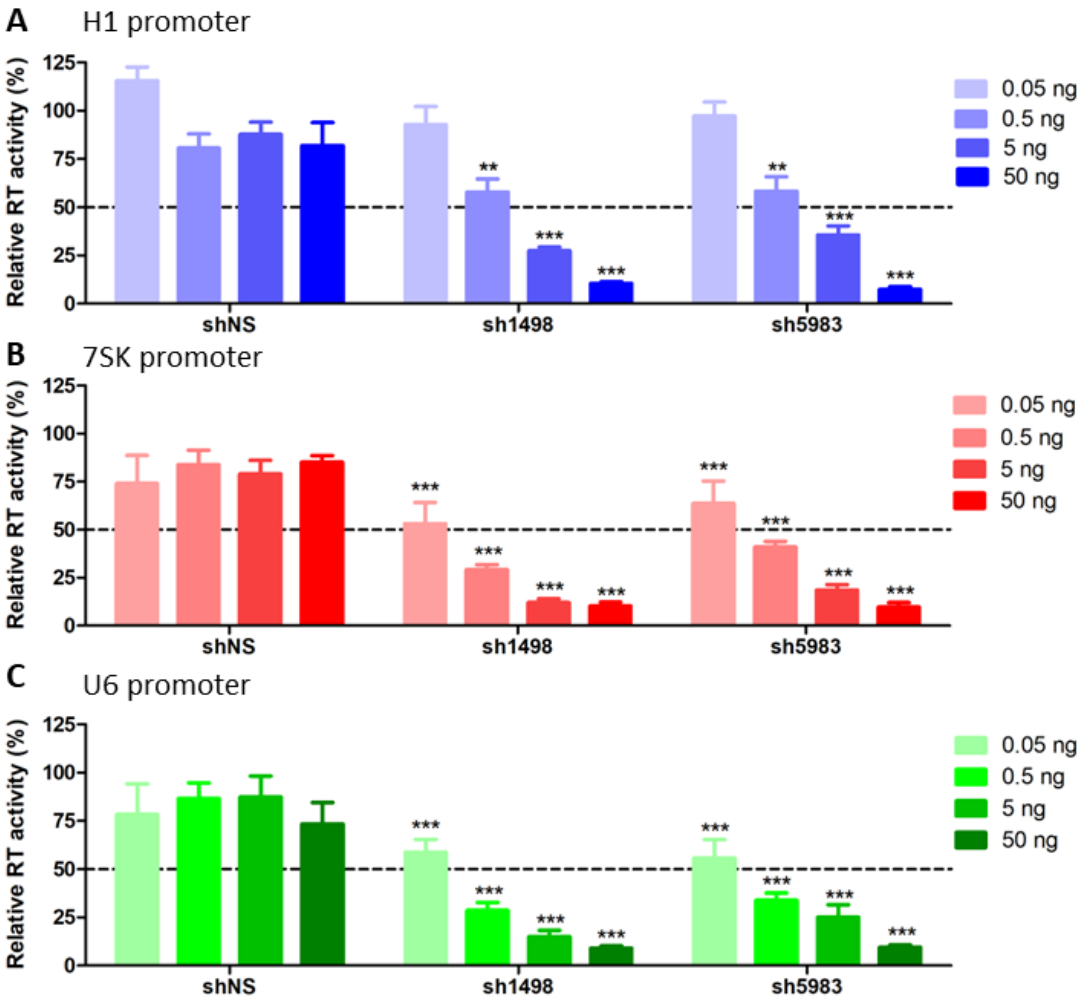

Figure 2.1 Inhibition of HIV-1 production by anti-HIV shRNAs expressed from different promoters. HEK 293T cells were cotransfected with the HIV-1 molecular clone pNL4-3 along with one of the plasmids containing the RNA pol III H1 promoter (A), 7SK promoter (B) or U6 promoter (C) and expressing a shRNA. Supernatants were collected 48 h post-transfection and virus production was estimated by measuring HIV-1 RT activity. Data are expressed as a percentage of RT activity in cells cotransfected with the respective empty shRNA expression plasmid. Each data point represents the mean +/- standard error mean (SEM) from at least two independent experiments with 2 replicates (n = 4-16). Effect of shRNAs on virus production are shown when expressed from each RNA pol III promoter. A two-way ANOVA with Bonferroni post test was used to compare means to means of empty vector transfected cells. Significance ($P<0.05$ *, $P<0.01$ **, $P<0.001$ ***) is shown for those data points that were significantly different.

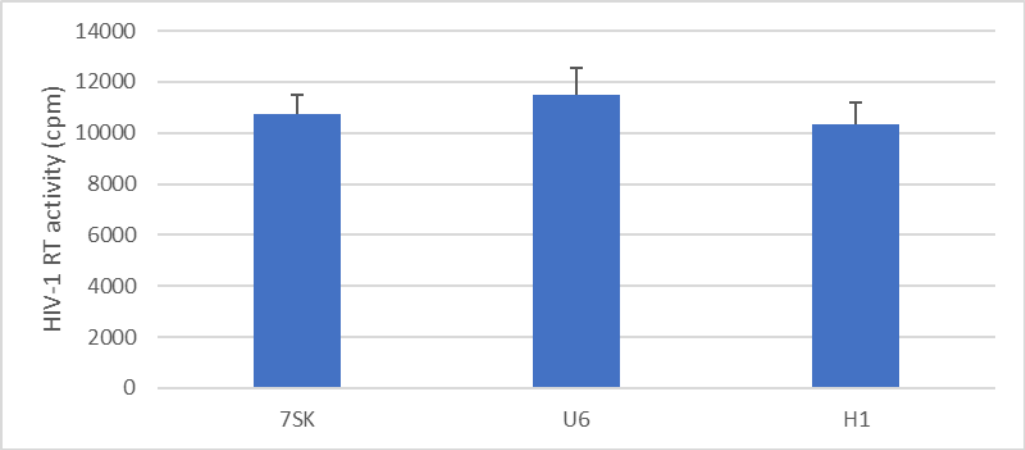

Figure 2.2 RT activity following co-transfection of HIV-1 pNL4-3 and different empty psiRNA vectors. HEK293T cells were cotransfected with the HIV-1 molecular clone pNL4-3 (100 ng) along with 100 ng of one of the empty psiRNA expression plasmids containing the RNA pol III 7SK, U6 or H1 promoter. Supernatants were collected 48 h post-transfection and virus production was estimated by measuring HIV-1 RT activity. Data are expressed as counts per minute. Each data point represents the mean +/- standard error mean (SEM) from two independent experiments with 3 replicates (n = 6). An unpaired two-tailed t-test was used to compare the mean of 7SK empty vector transfected cell RT activity to the means of the other promoters. No comparisons were significantly different ($P<0.05$).

2.4.2 Anti-HIV-1 shRNAs expressed from the different Pol III promoters do not affect cell viability in HEK 293T cells

To determine if the differences in effects on HIV-1 production were related to cytotoxicity, we measured cell viability in HEK 293T cells using a WST-1 metabolism assay two days after transfection of the different shRNA expression vectors. shRNA cassettes were transfected into HEK 293T cells at 300 to 40,000 times their IC_{50} levels (0.05 to 5 ng, Figure 2.1). WST-1 metabolism was normalized to the metabolism in cells transfected with the empty psiRNA vectors. Cell viability was not impaired at doses of up to 2 µg when the shRNAs were expressed from any of the promoters (Figure 2.3A-C), demonstrating that the shRNAs are not cytotoxic when transfected at levels well above their effective amounts in HEK 293T cells. These results confirm

that differences in the inhibition of HIV-1 production by the shRNAs expressed from the different promoters is not a consequence of differences in cytotoxicity.

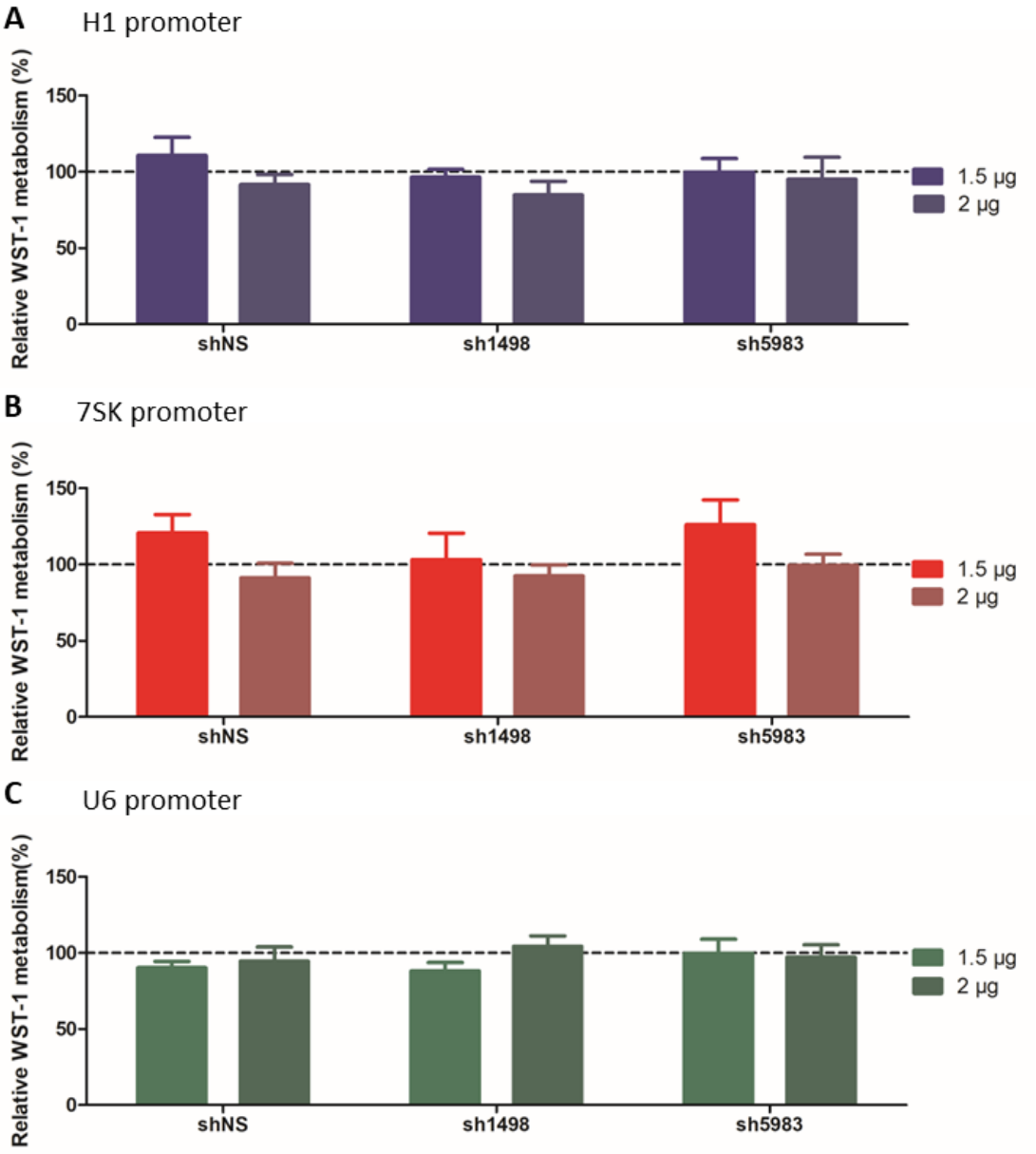

Figure 2.3 shRNAs do not affect cell viability when expressed from RNA pol III promoters. HEK 293T cells were transfected with 0.05 µg – 2 µg of plasmid containing the RNA pol III H1 promoter (A), 7SK promoter (B) and U6 promoter (C) and expressing a shRNA. Cell viability was

quantified by measuring the metabolism of WST-1 48 h post-transfection. Effect of shRNAs on cell viability is expressed as a percentage of WST-1 metabolism. Each data point represents the mean +/- SEM from at least three independent experiments with 1-2 replicates (n = 5-8). A two-way ANOVA with Bonferroni post-test was used to compare means to means of empty vector transfected cells. No data points were significantly different ($P<0.05$).

2.4.3 The potency of anti-HIV-1 shRNAs is not improved when expressed from Pol III promoters containing the complete natural human sequences

shRNA expression plasmids are typically designed to include a restriction site immediately before the intended +1 transcription start site to facilitate cloning of different shRNA sequences downstream of the promoter (see Table 2.1 for the upstream sequences in some commonly used vectors). However, the precise start site for transcription can be affected by the surrounding nucleotide sequence[30,31] and these sequences may also affect the transcription efficiency. The commercially available psiRNA plasmid (InvivoGen) used in this study contains the 7SK promoter, which happens to include a KpnI restriction site two nucleotides upstream from the +1 transcription start site. This KpnI site was used by InvivoGen to replace the 7SK promoter with the H1 promoter and by us to replace the 7SK promoter with the U6 promoter for this study. As a result, both the U6 and H1 promoters contain eight nucleotides of the 7SK promoter sequence directly upstream of their intended +1 transcription start sites (Figure 2.4A). To evaluate if the antiviral potency of the shRNAs is affected by these eight nucleotides, site-directed mutagenesis was used to change them to the natural H1 and U6 promoter sequences, which we called the "humanized (h)" hH1 and hU6 promoters (Figure 2.4A). Inhibition of HIV-1 production was measured using HIV-1 RT activity following co-transfections of plasmid constructs expressing anti-HIV-1 shRNAs and HIV-1 pNL4-3. All data were normalized to the empty psiRNA vectors at each dose. Interestingly, the H1 shRNA cassettes with the 7SK sequence upstream of the transcription start site were more potent compared to the hH1 shRNA cassettes with the natural H1 sequence (Figure 2.4B). The IC_{50}s of the H1 shRNA cassettes were 0.24 ng for sh5983 and

0.59 ng for sh1498 compared to the IC_{50}s of the hH1 shRNA cassettes, which were 4.75 ng for sh5983 and 2.49 ng for sh1498. In contrast, the IC_{50}s of the U6 shRNA cassettes were similar between the hU6 and U6 promoters (Figure 2.4C).

Table 2.1 Promoter sequences upstream of the +1 start site in commercially available vectors.

Natural H1	**ACTCTTTCCC+1**	**Human**
pSUPER	ACAGATCTAA+1	Oligoengine
psiRNA	ACGGTACCTC+1	InVivogen
pSilencer	ACTCGGATCC+1	Life Tech.
Natural U6	**ACGAAACACC+1**	**Human**
pSIREN-DNR	ACGAGGATCC+1	Clontech
pSilencer	ACGCGGGATC+1	Life Tech.

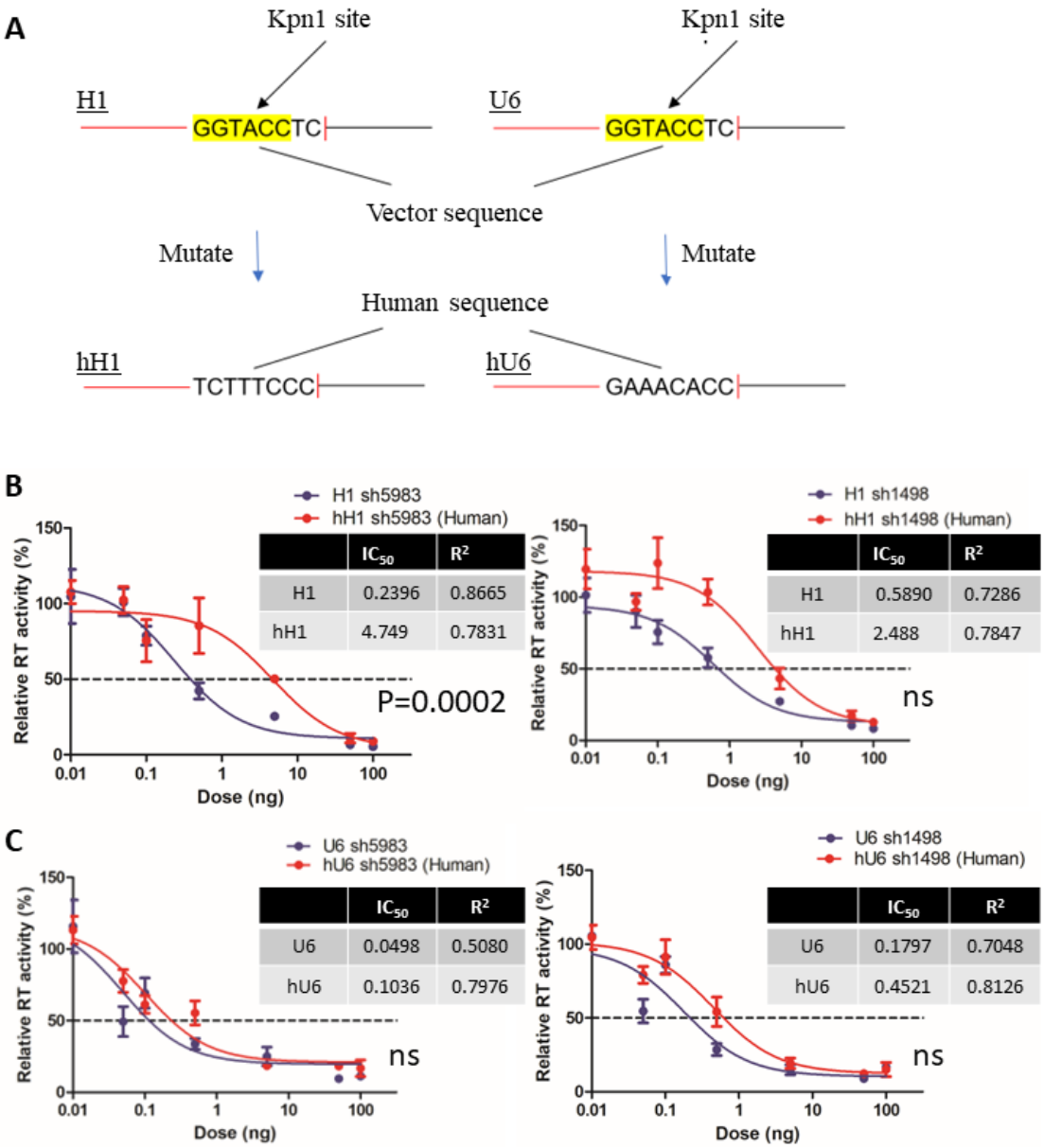

Figure 2.4 Antiviral effect of shRNAs when expressed from "humanized" promoters. The KpnI sites upstream of the transcriptional start site were replaced by the natural promoter sequence (A). The effect of the molecules on HIV-1 production was measured exactly as in Figure 2.1. Antiviral potency of shRNAs was compared when expressed from promoters containing the KpnI site and "humanized" promoters for H1 (B) and U6 (C). Each data point represents the mean +/- SEM from at least two independent experiments with 1-2 replicates (n = 3-15). A nonlinear regression log(inhibitor) versus response equation with least-square (ordinary) fit was applied to the log transformed data and statistical significance between LogIC50s was determined using extra sum-of-squares F test where $P<0.05$ is considered not significant (ns).

2.4.4 Expression levels of shRNAs from each Pol III promoter correlates with their antiviral activity

We then hypothesized that the differences in antiviral potency of the shRNA cassettes with the different Pol III promoters may be due to differing efficiencies of transcription. To determine whether this is the case, RNA was recovered from the transfected cells and the expected shRNA guide strands were measured by Northern blot. The relative expression level of shNS, sh5983 and sh1498 when expressed by the different Pol III promoters was then determined by quantifying the different band intensities normalized to the 5S rRNA loading control using Fiji software (Figure 2.5). We observed that gene expression levels were similar between the 7SK, U6 and hU6 promoters when they express the three different shRNAs. The expression levels between the H1 and the hH1 promoters expressing these same shRNAs were similar, but much lower compared to 7SK, U6 and hU6 promoters (Figure 2.5A-C). This expression pattern closely mirrors the trend of antiviral potency observed when the shRNAs were expressed from the different Pol III promoters (Figures 2.1 and 2.4), which suggests that the differences in antiviral effects observed is most likely due to the varying transcription efficiencies of the promoters.

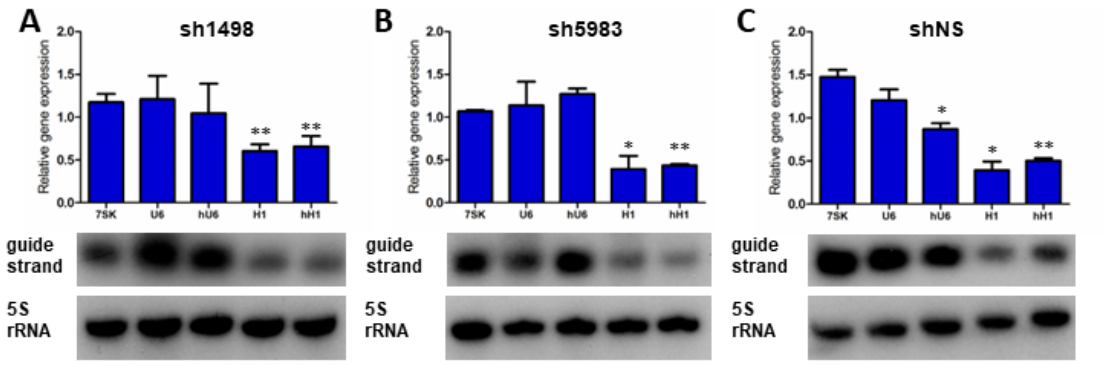

Figure 2.5 The 7SK, U6 and hU6 promoters express higher levels of shRNA guide strands than the H1 and hH1 promoters. RNA harvested from shRNA transfected HEK 293T cells was migrated in a 15% polyacrylamide-urea gel. shRNA as well as 5S RNA were detected with ^{32}P-labeled RNA probes and band intensities were analyzed with Fiji software to generate numerical values from two independent experiments. Intended guide strand RNA expression levels from the Pol III promoters is shown for sh1498 (A), sh5983 (B) and shNS (C). Each data point represents the mean +/- SEM from two independent experiments (n=2). An unpaired two-tailed t-test was used to compare the mean of 7SK gene expression values to the mean of the other promoters. Significance ($P<0.05$ *, $P<0.01$ **, $P<0.001$ ***) is shown for gene expression means that were significantly different compared to the 7SK gene expression means.

2.4.5 Cells transduced with a U6 and 7SK driven anti-HIV-1 shRNA restrict viral replication but have a severe growth disadvantage compared to untransduced cells

To compare the efficacy of an anti-HIV-1 shRNA driven by the different Pol III promoters against HIV-1 replication in a T cell line, SupT1 cells were transduced with lentiviral vectors (HIV7-GFP[37]) carrying U6, 7SK and H1 driven sh1498 and shNS. Following transduction, cells were sorted for GFP expression (Gating shown in Figure 2.8) and infected with HIV-1 NL4-3. Viral replication was measured using HIV-1 RT activity in the culture supernatants over time. Regardless of the promoter, shNS transduced cells had similar replication kinetics compared to the empty vector transduced cells (Figure 2.6A). Cells transduced with H1 driven sh1498 also had similar HIV-1 replication kinetics compared to the empty vector (Figure 2.6B). In contrast, no viral replication was detected in cells transduced with U6 or 7SK driven sh1498 up to 60 days post infection. Similar results were obtained from an independent transduction (Gating shown in Figure

2.9) except that one replicate infection for U6 sh1498 transduced cells started to produce detectable RT activity at 26 days post infection (Figure 2.6C).

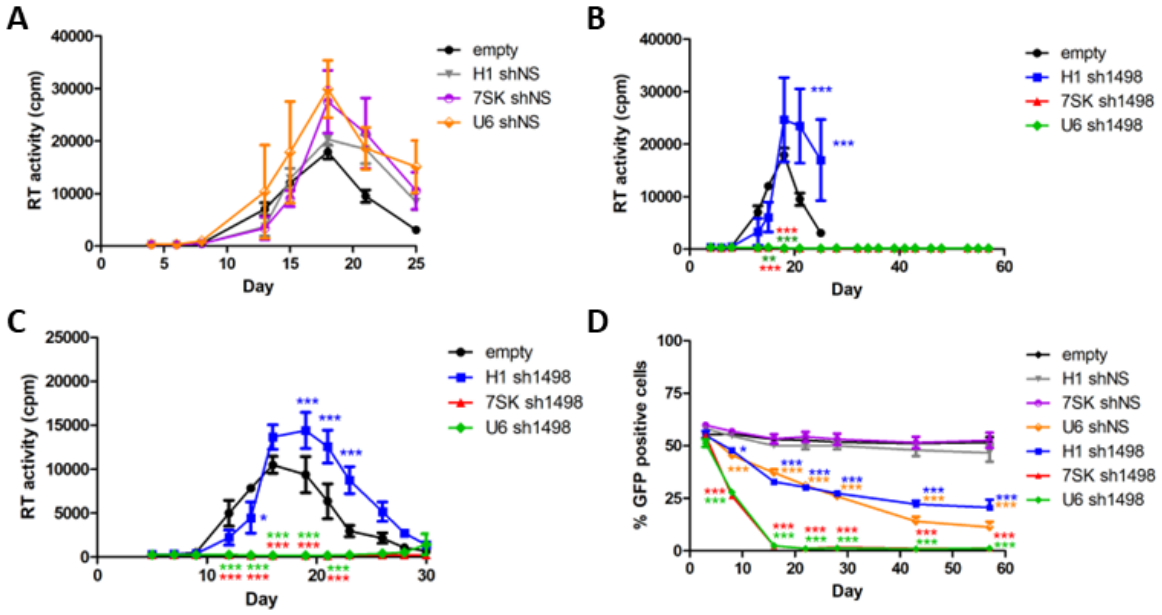

Figure 2.6 HIV-1 replication is restricted in SupT1 cells expressing sh1498 from the 7SK and U6 promoters, but the cells have a severe growth disadvantage compared to untransduced SupT1 cells. SupT1 cells were transduced with HIV-7-EGFP lentiviral vectors expressing shNS (A) or sh1498 (B and C) from the H1, U6 and 7SK promoters and infected with HIV-1 NL4-3 at 1750 cpm/mL. The mean RT activity (cpm) was measured in culture supernatants at various days post infection. Each data point represents the mean +/- SEM from three infections (n=3). (D) Transduced SupT1 cells were mixed with untransduced SupT1 cells and the percentage of GFP positive cells was measured at various days post-mixing. Each data point represents the mean +/- SEM from three experiments (n=3). A two-way ANOVA with Bonferroni post-test was used to compare means to means of empty vector transduced cells. Significance ($P<0.05$ *, $P<0.01$ **, $P<0.001$ ***) is shown for those data points that were significantly different from the empty vector transduced controls.

To evaluate the potential for shRNA expression to affect cell growth we also mixed GFP sorted transduced SupT1 cells with unsorted SupT1 cells that had also been passed through the flow cytometer (same cells as for Figure 2.6C but not infected, gating shown in Figure 2.9). We then followed the % GFP positive cells over time using a competitive cell growth assay similar to a

previously described method[38] (Figure 2.6D). Mixed cultures of empty vector, H1 shNS and 7SK shNS remained at around 50% out to 57 days post sorting. In contrast, the % of GFP positive cells decreased over time for U6 shNS and H1 sh1498 cultures down to 11 and 21 %, respectively. They rapidly decreased to nearly 0 % in U6 sh1498 and 7SK sh1498 cultures by day 16. Because of the severe impact on cell growth observed for U6 and 7SK sh1498 mixed cultures (Figure 2.6D), it is difficult to conclude whether the failure of HIV-1 to replicate in these cells (Figures 2.6B and C) was a result of cell toxicity or the effect of the shRNAs. Likely, a combination of shRNAs targeting HIV-1 and loss of cells available to be infected because of toxicity stopped HIV-1 from establishing an infection at the beginning of the experiment.

2.4.6 Expression of shRNAs leads to the accumulation of multiple RNA sequences

shRNA expression cassettes are typically designed with the intended passenger strand at the 5' end separated by a loop from the intended guide strand at the 3' end. Following transcription, the complementary passenger and guide strands hybridize to form a hairpin structure, which is exported to the cytoplasm by Exportin 5 where the molecule will be processed by RNAi enzymes. This processing begins by Dicer removing the loop and then by Ago2 cleaving the passenger strand[16]. The guide strand can then direct Ago2 to cleave its target RNA by complementary base pairing. The accumulation of different RNA species, in addition to the intended active guide strand, results from the differential processing by the RNAi enzymes as well as from the potential for alternative start and end sites of transcription. To identify potential differences in the expression of these RNA sequences between the different Pol III promoters we sequenced the small RNAs of shRNA transfected HEK 293T cells.

RNA Seq data were obtained from cDNA libraries generated with a small RNA library preparation kit that uses adaptors that specifically ligate to 3' hydroxyl groups that are a hallmark of Dicer

cleavage. All sequences that aligned with each shRNA cassette were identified and organized by read number. The data were then expressed as a percentage of reads over total number of reads for each sequence identified. Using a cut-off of 0.1% of total reads, the sequences were organized into tables with their corresponding abundance tabulated for each promoter (Tables 2.2-2.4). The most common sequences (>0.5% of total reads) are illustrated for sh1498 and sh5983 in Tables 2.5 and 2.6, respectively. As expected, several sequences were identified for shRNAs expressed from all promoters corresponding predominantly to guide and passenger strands of different lengths. Notably, a more diverse set of sequences were observed for sh5983 compared to sh1498, due predominantly to more diverse Dicer cleavage sites seen at the 3' ends of passenger strands and the 5' ends of guide strands. Overall, the results demonstrate that a variety of sequences accumulate following shRNA transfection and that Dicer cleavage is more uniform for sh1498 compared to sh5983.

Table 2.2 Transcriptional profile of sh1498 expressed from different RNA Pol III promoters for reads occurring at > 0.1% of total reads.

sh1498						
	NNNGCAGGAACTACTAGTACCCT**ACTCGAGAA**GGGTACTAGTAGTTCCTGC**TTTTT**	7SK	U6	hU6	H1	hH1
	CTCGCAGGAACTACTAGTACCCT	1.3				
	CTCGCAGGAACTACTAGTACCC	0.9				
	CTCGCAGGAACTACTAGTACC	0.3				
	CTCGCAGGAACTACTAGTAC	0.1				
	TCGCAGGAACTACTAGTACCCTA	0.3				
	TCGCAGGAACTACTAGTACCCT	11.6			6.7	
	TCGCAGGAACTACTAGTACCC	1.7			1.2	
	TCGCAGGAACTACTAGTACC	0.4			0.4	
	CGCAGGAACTACTAGTACCCT	1.8	0.8	2.8	4.6	2.1
	CGCAGGAACTACTAGTACCC	0.2	0.1	0.4	0.8	0.3
	CGCAGGAACTACTAGTACC			0.1	0.2	
	GCAGGAACTACTAGTACCCTACTCG		0.2		0.2	
	GCAGGAACTACTAGTACCCTACT		0.3	0.1	0.2	
	GCAGGAACTACTAGTACCCTAC		0.2			
Passenger	GCAGGAACTACTAGTACCCTA	0.3	0.5	0.3	0.2	0.3
	GCAGGAACTACTAGTACCCT	7.3	11.2	11.1	7.2	7.9
	GCAGGAACTACTAGTACCC	1.9	3.6	3.7	2.8	2.2
	GCAGGAACTACTAGTACC	0.2	0.5	0.4	0.4	0.2
	GCAGGAACTACTAGTA	0.2	0.7	0.4	0.7	0.3
	GCAGGAACTACTAGT	0.5	1.6	0.9	1.3	0.6
	GCAGGAACTACTAG	0.4	1.4	0.8	1.5	0.3
	GCAGGAACTACTA		0.4	0.2	0.4	
	GCAGGAACTACT	0.3	2.4	1.3	2.6	0.5
	GCAGGAACT		0.2			
	AGGAACTACTAGTACCCTACTCGA		0.1			
	AGGAACTACTAGTACCCTACTCG		0.1			
	AGGAACTACTAGTACCCTAC		0.3	0.1	0.1	
	AGGAACTACTAGTACCC		0.4	0.1	0.2	
	TACTCGAGAAGG		0.2		0.1	
	GAGAAGGGTACTAGTAG		0.2		0.1	
	AGAAGGGTACTAGT		0.3		0.2	
	AGAAGGGTACTAGTA		0.2			
Other	AGAAGGGTACTAGTAG		1.4	0.5	0.8	
	AGAAGGGTACTAGTAGT		1.2	0.4	0.7	
	AGAAGGGTACTAGTAGTT		0.1			
	AGAAGGGTACTAGTAGTTCC		0.1			
	GGTACTAGTAGTTCCTGC	0.3	0.4	0.3	0.2	0.3
	AGAAGGGTACTAGTAGTTCCTGCT		0.2			
	GAAGGGTACTAGTAGTTCCTGCT		0.1			
	GGTACTAGTAGTTCCTGCT	3.2	8.5	8.2	9.1	4.8
	GTACTAGTAGTTCCTGCT	0.4	2.0	1.6	2.0	0.6
	GAAGGGTACTAGTAGTTCCTGCTT		0.1			
	GGTACTAGTAGTTCCTGCTT	7.9	11.7	11.7	7.7	10.3
	GTACTAGTAGTTCCTGCTT	0.5	1.5	1.3	1.2	0.7
	GGTACTAGTAGTTCCTGCTTT	33.8	22.6	26.8	18.8	38.6
Guide	GTACTAGTAGTTCCTGCTTT	4.1	6.3	6.3	5.9	5.4
	GGTACTAGTAGTTCCTGCTTTT	7.0	4.2	5.8	5.3	8.4
	GTACTAGTAGTTCCTGCTTTT	8.8	7.4	9.1	11.4	11.4
	GGTACTAGTAGTTCCTGCTTTTT	0.9	0.5	0.7	0.5	1.0
	GTACTAGTAGTTCCTGCTTTTT	1.4	0.9	1.1	0.9	1.7

Table 2.3 Transcriptional profile of sh5983 expressed from different RNA Pol III promoters for reads occurring at > 0.1% of total reads.

sh5983						
	NNNGCGGAGACAGCGACGAAGAGGCTCGAGGCTCTTCGTCGCTGTCTCCGC**TTTTT**	7SK	U6	hU6	H1	hH1
Passenger	CCCGCGGAGACAGCGACGA					0.2
	CCCGCGGAGACAGCGACGAA					0.1
	CCCGCGGAGACAGCGACGAAG					1.1
	CCCGCGGAGACAGCGACGAAGA					4.6
	CTCGCGGAGACAGCGACGAAGA	1.0				
	CCCGCGGAGACAGCGACGAAGAG					0.2
	CCGCGGAGACAGCGACGA					0.1
	CCGCGGAGACAGCGACGAAG					0.7
	TCGCGGAGACAGCGACGAAG	0.2			0.1	
	CCGCGGAGACAGCGACGAAGA					1.8
	TCGCGGAGACAGCGACGAAGA	0.8			0.2	
	CCGCGGAGACAGCGACGAAGAG					1.6
	TCGCGGAGACAGCGACGAAGAG	0.6			0.2	
	CCGCGGAGACAGCGACGAAGAGG					0.2
	TCGCGGAGACAGCGACGAAGAGG	0.3				
	CGCGGAGACAGCGACGAAG	0.1		0.3	0.4	0.4
	CGCGGAGACAGCGACGAAGA	0.3	0.2	0.6	0.9	1.0
	CGCGGAGACAGCGACGAAGAG	0.3	0.2	0.5	0.9	0.7
	CGCGGAGACAGCGACGAAGAGG	0.2	0.1	0.3	0.9	0.5
	GCGGAGAC				0.1	
	GCGGAGACA		0.1		0.1	
	GCGGAGACAG		0.2		0.1	
	GCGGAGACAGCG	0.3	2.3	1.1	1.0	0.3
	GCGGAGACAGCGA		0.2	0.1		
	GCGGAGACAGCGAC		0.3	0.1		
	GCGGAGACAGCGACG		0.1		0.1	
	GCGGAGACAGCGACGA	0.3	0.6	0.8	0.8	0.2
	GCGGAGACAGCGACGAA	0.4	0.9	0.9	0.8	0.3
	GCGGAGACAGCGACGAAG	2.6	3.9	6.1	6.7	1.9
	GCGGAGACAGCGACGAAGA	5.5	6.5	9.7	9.9	3.6
	GCGGAGACAGCGACGAAGAG	0.7	0.7	0.9	1.2	0.4
	GCGGAGACAGCGACGAAGAGG	0.5	0.6	0.7	1.3	0.3
	GCGGAGACAGCGACGAAGAGGC	1.5	2.7	2.5	5.4	1.1
	GCGGAGACAGCGACGAAGAGGCT		0.2	0.3	0.5	
	GCGGAGACAGCGACGAAGAGGCTC		0.2	0.2	0.4	
	GCGGAGACAGCGACGAAGAGGCTCG	0.2	0.8	0.5	1.4	0.2
	CGGAGACAGCGACGAAG		0.2	0.1	0.2	
	CGGAGACAGCGACGAAGA		0.1			
	CGGAGACAGCGACGAAGAGGC		0.2	0.1	0.3	
	CGGAGACAGCGACGAAGAGGCT		0.2		0.2	
	GGAGACAGCGACGAAGAGGC		0.2		0.2	
Other	GAGACAGCGACGAAGAGGC		0.4	0.3	0.3	
	GAGACAGCGACGAAGAGGCT		0.1			
	AGACAGCGACGAAG		0.1		0.2	
	AGACAGCGACGAAGA		0.1			
	AGACAGCGACGAAGAGGC	0.1	2.8	1.2	2.8	0.2
	AGACAGCGACGAAGAGGCT		0.3	0.1	0.2	
	AGCGACGAAGAGGC		0.1			
	ACGAAGAGGC		0.1			
	ACGAAGAGGCTCG		0.5	0.3	0.2	0.1
	TCGAGGCTCTTCGT		0.2		0.2	
	TCGAGGCTCTTCGTC		0.1			
	TCGAGGCTCTTCGTCG		0.3	0.2	0.3	
	TCGAGGCTCTTCGTCGC		1.3	0.7	1.2	0.1
	TCGAGGCTCTTCGTCGCT		0.1	0.1	0.1	
	CGAGGCTCTTCGT		0.2		0.3	
	CGAGGCTCTTCGTCG		0.1			
	CGAGGCTCTTCGTCGC		0.8	0.3	0.7	
	CGAGGCTCTTCGTCGCT		0.6	0.3	0.3	
	GAGGCTCTTCGTCGCT		0.1			
	AGGCTCTTCGT		0.1		0.1	
	AGGCTCTTCGTCGCT		0.1			
	AGGCTCTTCGTCGCTG		0.2			
	AGGCTCTTCGTCGCTGT		0.4	0.2	0.1	
	AGGCTCTTCGTCGCTGTCT		0.2	0.1		
	AGGCTCTTCGTCGCTGTCTC		0.2			
	AGGCTCTTCGTCGCTGTCTCC		0.7	0.3	0.4	
	AGGCTCTTCGTCGCTGTCTCCG		0.6	0.3	0.5	
	GGCTCTTCGTCGCTGT		0.1			
	GGCTCTTCGTCGCTGTCT		0.1			
	TCTTCGTCGCTG		0.1			
	TCTTCGTCGCTGTCTCC		0.2	0.1	0.1	

Table 2.3 (continued) Transcriptional profile of sh5983 expressed from different RNA Pol III promoters for reads occurring at > 0.1% of total reads.

Guide						
	AGGCTCTTCGTCGCTGTCTCCGC		0.4	0.1	0.3	
	CTCTTCGTCGCTGTCTCCGC		0.1			
	TCTTCGTCGCTGTCTCCGC	0.1	0.3	0.2	0.2	
	AGGCTCTTCGTCGCTGTCTCCGCT		0.9	0.4	0.7	
	GGCTCTTCGTCGCTGTCTCCGCT		0.2	0.1	0.2	
	GCTCTTCGTCGCTGTCTCCGCT		0.2			
	CTCTTCGTCGCTGTCTCCGCT	0.8	1.4	1.1	1.1	0.5
	TCTTCGTCGCTGTCTCCGCT	1.8	4.8	3.3	2.7	1.4
	CTTCGTCGCTGTCTCCGCT	0.8	1.1	0.9	0.8	0.7
	TTCGTCGCTGTCTCCGCT		0.1	0.1		
	AGGCTCTTCGTCGCTGTCTCCGCTT	0.3	1.6	0.9	2.5	0.3
	GGCTCTTCGTCGCTGTCTCCGCTT		0.4	0.2	0.4	0.1
	GCTCTTCGTCGCTGTCTCCGCTT		0.3	0.2	0.2	
	CTCTTCGTCGCTGTCTCCGCTT	1.8	1.6	1.3	1.3	1.0
	TCTTCGTCGCTGTCTCCGCTT	6.9	9.5	8.3	7.3	6.6
	CTTCGTCGCTGTCTCCGCTT	5.4	4.6	4.7	3.7	5.2
	TTCGTCGCTGTCTCCGCTT	0.2	0.3	0.3	0.2	0.2
	GGCTCTTCGTCGCTGTCTCCGCTTT	0.1	0.3	0.2	0.3	0.2
	GCTCTTCGTCGCTGTCTCCGCTTT		0.2			
	CTCTTCGTCGCTGTCTCCGCTTT	1.2	0.7	0.7	0.6	0.7
	TCTTCGTCGCTGTCTCCGCTTT	16.6	9.5	10.8	8.5	15.0
	CTTCGTCGCTGTCTCCGCTTT	26.7	13.7	16.9	12.2	26.3
	TTCGTCGCTGTCTCCGCTTT	1.5	1.7	1.9	1.4	1.7
	CTCTTCGTCGCTGTCTCCGCTTTT	0.2	0.1	0.2	0.1	0.1
	TCTTCGTCGCTGTCTCCGCTTTT	1.5	1.1	1.5	1.0	1.4
	CTTCGTCGCTGTCTCCGCTTTT	11.6	4.2	7.3	4.8	10.7
	TTCGTCGCTGTCTCCGCTTTT	1.5	0.9	1.3	0.9	1.5
	TCTTCGTCGCTGTCTCCGCTTTTT		0.1	0.1		
	CTTCGTCGCTGTCTCCGCTTTTT	0.8	0.3	0.5	0.4	0.8
	TTCGTCGCTGTCTCCGCTTTTT	0.3	0.1	0.2	0.2	0.3

Table 2.4 Transcriptional profile of shNS expressed from different RNA Pol III promoters for reads occurring at > 0.1% of total reads.

shNS						
	NNNGTACCGCACGTCATTCGTATCCTCGAGCATACGAATGACGTGCGGTAC**TTTTT**	7SK	U6	hU6	H1	hH1
	CTCGTACCGCACGTCATTCGTA	0.2				0.2
	CTCGTACCGCACGTCATTCGT	0.1				0.1
	GTACCGCACGTCATTCGTATCCT		0.2		0.2	
	GTACCGCACGTCATTCGTATCCTCG	0.1	1.8	0.7	2.1	0.1
	GTACCGCACGTCATTCGTATCCTC		0.5	0.3	0.7	
	TACCGCACGTCATTCGTATCCTCGA		0.1			
	TACCGCACGTCATTCGTATCCTCG		0.4	0.2	0.4	
	TATCCTCGAGCATACGAATGACGTG				0.1	
	TCGAGCATACGAATGACG		0.1			
	TCGAGCATACGAATGACGT		0.4	0.3	0.2	
	CGAGCATACGAATGACGT		0.6	0.2	0.2	
	GAGCATACGAATGACGT		0.1			
	TACGAATGACGTGCGGT			0.1		
	ACGAATGACGTGCGGT					
	TACGAATGACGTGCGGTA			0.1		
	ACGAATGACGTGCGGTA					0.1
	TACGAATGACGTGCGGTAC		0.2	0.1	0.1	
	AGCATACGAATGACGTGCGGTACT		0.7	0.4	1.1	
	CATACGAATGACGTGCGGTACT		0.3	0.1	0.3	
	ATACGAATGACGTGCGGTACT		0.1		0.1	
	TACGAATGACGTGCGGTACT	1.5	4.5	3.7	3.3	1.5
	ACGAATGACGTGCGGTACT	0.7	1.7	1.7	1.2	0.8
	CGAATGACGTGCGGTACT	0.2	0.4	0.4	0.3	0.2
	AATGACGTGCGGTACT		0.1		0.1	
	AGCATACGAATGACGTGCGGTACTT		0.5	0.3	0.5	
	CATACGAATGACGTGCGGTACTT		0.4	0.2	0.4	
	ATACGAATGACGTGCGGTACTT		0.2	0.1	0.2	
	TACGAATGACGTGCGGTACTT	10.2	19.5	17.0	16.1	10.2
	ACGAATGACGTGCGGTACTT	7.6	11.8	11.8	10.2	8.3
	CGAATGACGTGCGGTACTT	0.6	0.5	0.6	0.5	1.8
	AATGACGTGCGGTACTT		0.1		0.1	
	CATACGAATGACGTGCGGTACTTT		0.1		0.1	
	TACGAATGACGTGCGGTACTTT	27.9	19.8	21.1	21.7	24.9
	ACGAATGACGTGCGGTACTTT	30.2	22.0	25.5	23.5	30.9
	CGAATGACGTGCGGTACTTT	1.7	1.2	1.4	1.2	2.1
	TACGAATGACGTGCGGTACTTTT	0.4	0.4	0.4	0.5	0.4
	ACGAATGACGTGCGGTACTTTT	13.9	6.4	8.4	9.4	13.2
	CGAATGACGTGCGGTACTTTT	1.7	0.9	1.1	1.2	2.0
	ACGAATGACGTGCGGTACTTTTT	1.2	0.6	0.8	0.8	1.3
	CGAATGACGTGCGGTACTTTTT	0.5	0.2	0.3	0.3	0.5

Table 2.5 Transcriptional profile of sh1498 expressed from different RNA Pol III promoters.

sh1498	Passenger / Loop / Guide	Promoter (% of total reads)				
	NNNGCAGGAACTACTAGTACCCT**ACTCGAGA**AGGGTACTAGTAGTTCCTGC**TTTTT**	7SK	U6	hU6	H1	hH1
Passenger	CTCGCAGGAACTACTAGTACCCT	1.3				
	CTCGCAGGAACTACTAGTACCC	0.9				
	TCGCAGGAACTACTAGTACCCT	11.6			6.7	
	TCGCAGGAACTACTAGTACCC	1.7			1.2	
	CGCAGGAACTACTAGTACCCT	1.8	0.8	2.8	4.6	2.1
	CGCAGGAACTACTAGTACCC				0.8	
	GCAGGAACTACTAGTACCCTA		0.5			
	GCAGGAACTACTAGTACCCT	7.3	11.2	11.1	7.2	7.9
	GCAGGAACTACTAGTACCC	1.9	3.6	3.7	2.8	2.2
	GCAGGAACTACTAGTACC		0.5			
	GCAGGAACTACTAGTA		0.7		0.7	
	GCAGGAACTACTAGT	0.5	1.6	0.9	1.3	0.6
	GCAGGAACTACTAG		1.4	0.8	1.5	
	GCAGGAACTACT		2.4	1.3	2.6	0.5
Other	AGAAGGGTACTAGTAG		1.4	0.5	0.8	
	AGAAGGGTACTAGTAGT		1.2		0.7	
Guide	GGTACTAGTAGTTCCTGCT	3.2	8.5	8.2	9.1	4.8
	GTACTAGTAGTTCCTGCT		2.0	1.6	2.0	0.6
	GGTACTAGTAGTTCCTGCTT	7.9	11.7	11.7	7.7	10.3
	GTACTAGTAGTTCCTGCTT	0.5	1.5	1.3	1.2	0.7
	GGTACTAGTAGTTCCTGCTTT	33.8	22.6	26.8	18.8	38.6
	GTACTAGTAGTTCCTGCTTT	4.1	6.3	6.3	5.9	5.4
	GGTACTAGTAGTTCCTGCTTTT	7.0	4.2	5.8	5.3	8.4
	GTACTAGTAGTTCCTGCTTTT	8.8	7.4	9.1	11.4	11.4
	GGTACTAGTAGTTCCTGCTTTTT	0.9	0.5	0.7	0.5	1.0
	GTACTAGTAGTTCCTGCTTTTT	1.4	0.9	1.1	0.9	1.7

Table 2.6 Transcriptional profile of sh5983 expressed from different RNA Pol III promoters.

sh5983	Passenger / Loop / Guide	Promoter (% of total reads)				
	NNNGCGGAGACAGCGACGAAGAG**GCTCGAGG**CTCTTCGTCGCTGTCTCCGC**TTTTT**	7SK	U6	hU6	H1	hH1
Passenger	CCCGCGGAGACAGCGACGAAGA					4.6
	CTCGCGGAGACAGCGACGAAGA	1.0				
	CCCGCGGAGACAGCGACGAAG					1.1
	CCGCGGAGACAGCGACGAAGAG					1.6
	TCGCGGAGACAGCGACGAAGAG	0.6				
	CCGCGGAGACAGCGACGAAGA					1.8
	TCGCGGAGACAGCGACGAAGA	0.8				
	CCGCGGAGACAGCGACGAAG					0.7
	CGCGGAGACAGCGACGAAGAGG				0.9	
	CGCGGAGACAGCGACGAAGAG				0.9	0.7
	CGCGGAGACAGCGACGAAGA			0.6	0.9	1.0
	GCGGAGACAGCGACGAAGAGGCTCG		0.8	0.5	1.4	
	GCGGAGACAGCGACGAAGAGGCT				0.5	
	GCGGAGACAGCGACGAAGAGGC	1.5	2.7	2.5	5.4	1.1
	GCGGAGACAGCGACGAAGAGG		0.6	0.7	1.3	
	GCGGAGACAGCGACGAAGAG	0.7	0.7	0.9	1.2	
	GCGGAGACAGCGACGAAGA	5.5	6.5	9.7	9.9	3.6
	GCGGAGACAGCGACGAAG	2.6	3.9	6.1	6.7	1.9

	GCGGAGACAGCGACGAA		0.9	0.9	0.8	
	GCGGAGACAGCGACGA		0.6	0.8	0.8	
	GCGGAGACAGCG		2.3	1.1	1.0	
Other	AGACAGCGACGAAGAGGC	0.1	2.8	1.2	2.8	
	ACGAAGAGGCTCG		0.5			
	TCGAGGCTCTTCGTCGC		1.3	0.7	1.2	
	CGAGGCTCTTCGTCGC		0.8		0.7	
	CGAGGCTCTTCGTCGCT		0.6			
	AGGCTCTTCGTCGCTGTCTCC		0.7			
	AGGCTCTTCGTCGCTGTCTCCG		0.6		0.5	
Guide	AGGCTCTTCGTCGCTGTCTCCGCT		0.9		0.7	
	CTCTTCGTCGCTGTCTCCGCT	0.8	1.4	1.1	1.1	
	TCTTCGTCGCTGTCTCCGCT	1.8	4.8	3.3	2.7	1.4
	CTTCGTCGCTGTCTCCGCT	0.8	1.1	0.9	0.8	0.7
	AGGCTCTTCGTCGCTGTCTCCGCTT		1.6	0.9	2.5	
	CTCTTCGTCGCTGTCTCCGCTT	1.8	1.6	1.3	1.3	1.0
	TCTTCGTCGCTGTCTCCGCTT	6.9	9.5	8.3	7.3	6.6
	CTTCGTCGCTGTCTCCGCTT	5.4	4.6	4.7	3.7	5.2
	CTCTTCGTCGCTGTCTCCGCTTT	1.2	0.7	0.7	0.6	0.7
	TCTTCGTCGCTGTCTCCGCTTT	16.6	9.5	10.8	8.5	15.0
	CTTCGTCGCTGTCTCCGCTTT	26.7	13.7	16.9	12.2	26.3
	TTCGTCGCTGTCTCCGCTTT	1.5	1.7	1.9	1.4	1.7
	TCTTCGTCGCTGTCTCCGCTTTT	1.5	1.1	1.5	1.0	1.4
	CTTCGTCGCTGTCTCCGCTTTT	11.6	4.2	7.3	4.8	10.7
	TTCGTCGCTGTCTCCGCTTTT	1.5	0.9	1.3	0.9	1.5
	CTTCGTCGCTGTCTCCGCTTTTT	0.8		0.5		0.8

2.4.7 The U6 promoter is the most accurate in using the +1 transcriptional start site but guide strand identity is not affected by differences in the +1 site usage

Using the abundance values in Tables 2.2-2.4 we calculated the percentage of reads corresponding to passenger, guide and other strands, where other strands were defined as those strands that start after the first 3 nucleotides following the expected +1 transcription start site (+4 and over) and end before the first nucleotide preceding the transcription termination signal of five Ts (Figure 2.7A). For all shRNAs expressed from all promoters, guide strands were the most abundant sequences. For both sh1498 and sh5983 the proportion of guide strand reads ranged from about 60 to 80 percent, whereas for shNS almost all the reads corresponded to guide strands and this was consistent for all of the promoters.

The percentage of reads with different end sites and start sites were calculated as a fraction of the total number of the guide strands and passenger strands, respectively (Figure 2.7B and C). For all

shRNAs expressed from all promoters the most common end sequence was UUU, suggesting that regardless of the promoter or shRNA sequences, UUU is the most common shRNA tail. No major differences in transcription termination were evident between the promoters nor the different shRNA sequences. In contrast, the different promoters resulted in major differences in transcription start sites (Figure 2.7C). Both versions of the U6 promoter were extremely precise at generating transcripts of all shRNAs from the intended +1 transcriptional start site. Meanwhile, the 7SK and H1 promoters varied in their use of this site depending on which shRNA was transcribed. Specifically, the 7SK promoter was reliable in using the +1 site solely when expressing sh5983 while the H1 promoter was reliable in using this site when expressing sh5983 and shNS but not sh1498. The hH1 promoter was reliable in using the +1 site only when transcribing sh1498. Overall, the U6 and hU6 promoters gave rise to more accurate transcriptional start sites when expressing the different shRNAs.

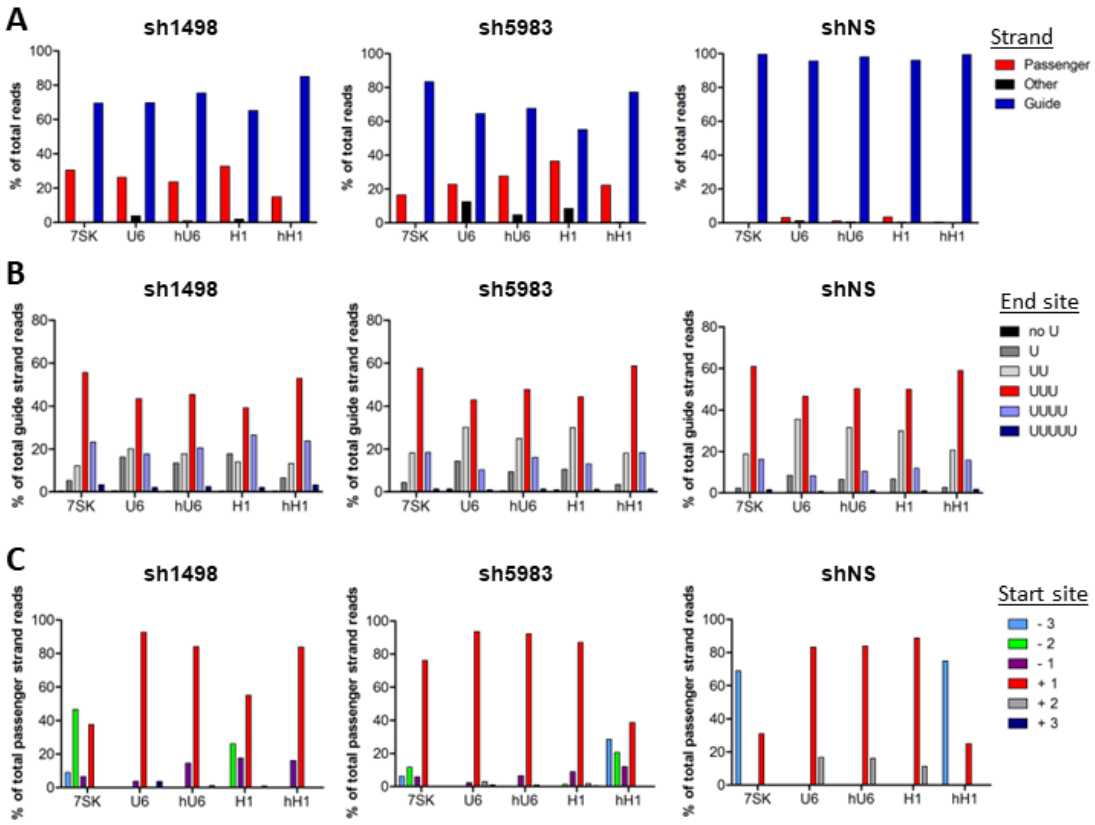

Figure 2.7 Proportion of reads corresponding to different RNA variants of sh1498, sh5983 and shNS. (A) The proportion of variants corresponding to expected passenger, guide and other strands are shown. (B) The proportion of guide strands ending in different numbers of Uridines is shown. (C) The proportion of passenger strands starting at different positions relative to the expected +1 transcriptional start site are shown.

An unexpected observation was that despite differences in transcriptional start sites between the promoters (Figure 2.7C) there were not major differences in the identity of the guides strands (Tables 2.5 and 2.6), suggesting that Dicer cleavage is unaffected by the length of the shRNA 5' end. This observation can most easily be appreciated when looking at sh1498 sequences (Table 2.5) for which, despite major differences in the start site between the different promoters, there are only two main Dicer cleavage sites revealed when looking at the 5' end of the guide strands. Similarly, when looking at the 3' end of the passenger strands, two major Dicer cleavage sites are

apparent for passenger strands that start at the -3 to +1 positions. These results suggest that, for shRNAs, Dicer cleavage sites are measured from the first base pair of the duplex rather than from the 5' end as has previously been demonstrated for structurally related microRNAs[32].

2.5 Discussion.

The RNA Pol III type 3 promoters U6, H1 and 7SK are typically used to express small RNAs such as shRNAs and CRISPR guide RNAs as transcription from these promoters has defined start and end sites and results in the addition of only a few uridines to the 3' end of transcripts. However, there has been limited work comparing these promoters for the expression of therapeutic RNAs. Additionally, of the studies that have utilized Pol III promoters for gene therapy, most have only focused on one or two promoters at a time[25,39,40]. For gene therapy to treat HIV-1 infection, multiple antiviral RNAs will be needed to avoid the development of viral resistance[15,18,28,29,41,42]. If a single promoter is used to express a combination of antiviral RNAs, deletions of the therapeutic genes could occur as a consequence of recombination between the different transcriptional units using the same Pol III promoter[28,29]. Therefore, a different Pol III promoter expressing each therapeutic RNA may be required to avoid recombination. With this in mind, it is important to properly evaluate which promoter expresses the most active antiviral RNAs as well as characterizing the transcriptional profile of the promoters U6, 7SK and H1.

Reports on which Pol III type 3 promoter express the most active shRNAs have not been consistent. Specifically, an shRNA targeting the HIV-1 vif coding sequence and a long hairpin RNA targeting the tat/rev viral sequences were shown to be more active when expressed from the H1 promoter compared to both the U6 and 7SK promoters[28]. In another study, two CCR5-specific shRNAs were shown to be more active when expressed from the U6 promoter compared to the H1

promoter, but the U6 driven shRNAs were cytotoxic in primary human blood cells[25]. Similar results were obtained in two other studies where both the expression level and activity of shRNAs were higher for the U6 promoter compared to the H1 promoter[26,27]. In contrast, in a study that compared four different anti-HIV-1 shRNAs expressed from the H1, U6 or 7SK promoters, similar activities were observed regardless of promoter choice[29]. To further complicate comparisons of the different Pol III promoters, it has been found that transcription start sites vary depending on both the promoter choice as well as the nucleotides upstream of the intended +1 transcription start sites[30], which can be problematic as these upstream nucleotides are often altered to accommodate a restriction enzyme recognition site (Table 2.1). In another study comparing several 7SK, H1 and U6 promoters, the nucleotide identity of the +1 transcriptional start site of a small unstructured RNA was found to affect both transcriptional activity and start site identity with the H1 promoters resulting in the most variable +1 start site.[31] While important observations have been made, there has not been any direct simultaneous comparison of shRNA activity, expression levels and transcriptional profile between the U6, 7SK and H1 promoters.

In this study, the antiviral potencies of an anti-HIV-1 shRNA targeting Tat/Rev (sh5983) and another targeting Gag (sh1498) were compared when expressed from the 7SK promoter as well as the U6 and H1 promoters, which included eight nucleotides of the 7SK sequence upstream of the +1 start site (Figure 2.1) and when expressed from the humanized hU6 and hH1 promoters with the complete natural promoter sequences (Figure 2.4). Consistent with studies that simultaneously compared both the expression level and activity of H1 and U6 driven shRNAs[25-27], our results suggest that shRNAs expressed from the U6 promoter are more potent and are expressed at a higher level when compared to shRNAs expressed from the H1 promoter (Figures 2.1 and 2.5), regardless of the nucleotide identity upstream of the transcription start site (Figure 2.4). While two

other studies reported conflicting observations for shRNA activity[28,29], neither study compared expression levels and the activity level of the shRNAs were compared at only one dose. It is possible that the doses selected in those studies were already above the level required for maximum target suppression and so the differences in activity related to differences in shRNA expression were not apparent. Within our inhibitory assays (Figures 2.1 and 2.4) we tested several doses and although the two anti-HIV shRNAs expressed from all promoters were able to almost completely suppress HIV-1 production at the higher doses in these assays, there was a clear and consistent difference in potency, with the H1 promoter shRNA constructs being less potent compared to the U6 and 7SK shRNA constructs. Our data therefore strongly suggest that the U6 promoter is indeed more transcriptionally active compared to the H1 promoter, at least in HEK 293T cells. We also show that the 7SK promoter is more transcriptionally active compared to the H1 promoter, providing similar expression levels and activities of shRNAs as the U6 promoter.

Although the transcriptional profiling revealed major differences in the localization of the transcription start sites for the different promoter-shRNA constructs (Figure 2.7C), our data suggest that the variations in RNA expression levels between the promoters (Figure 2.5) is the primary contributing factor towards the differences in antiviral potency observed (Figures 2.1 and 2.4). Differences in the transcriptional start sites were observed, for example the 7SK promoter was only accurate in its usage of the +1 site when expressing sh5983 and was much less accurate when expressing sh1498, while the U6 promoter was overall the most accurate in using the +1 transcriptional start site (Figure 2.7C). Because sh1498 expressed from the 7SK promoter had antiviral capabilities similar to those expressed from the U6 promoter (Figure 2.1B and C), despite its transcription from the 7SK promoter not beginning reliably at the +1 site, we conclude that expression from the +1 transcription start site is not a primary contributing factor to the antiviral

potency. Nonetheless, the identification of U6 as the most accurate promoter will have strong implications for molecules where an accurate +1 transcription start site is required for function, such as for CRISPR guide RNAs. As we compared only three shRNAs in one cell type, additional studies are needed to confirm if this observation can be widely applied to different molecules and in different cells.

It has been shown that human Dicer measures approximately 22 nucleotides from the 5' end of microRNAs (5' counting rule) to locate its cleavage site[32] and would therefore be expected to yield different siRNA duplexes when the +1 site at the 5' end is changed. The different duplexes should then give rise to different guide strands, which, in contrast with our conclusion, would be expected to have different RNAi activities. However, our sequencing results suggest that Dicer cleavage sites are not altered when the +1 site is changed and that regardless of the +1 start site the identity of the guide strands remains the same as can be seen most clearly for sh1498 in Table 2.5. Since the 5' counting rule for human Dicer is dependent on a 5' terminal phosphate, it may be that shRNAs transcribed from RNA Pol III promoters do not follow this rule because they have a 5' terminal triphosphate instead of the monophosphate typical of primary microRNAs[32]. Alternatively, since the 5' counting rule was established using *in vitro* cleavage experiments, it may not apply to Dicer cleavage in live cells and further studies are needed to better characterize Dicer cleavage of different hairpin RNAs in different environments. Interestingly, our results also demonstrate that independently from the promoter, Dicer cleavage was more uniform for sh1498 compared to sh5983, with only two major Dicer cleavage sites apparent for sh1498 compared to four for sh5983 (Table 2.5 and 2.6). Thus, the identity of the shRNA sequence can affect the number of Dicer products and, consequently, the diversity of guide and passenger strands. Importantly, our results show that certain promoter-shRNA cassettes have different start sites but

that this does not affect the identity and corresponding RNAi activity of the guide strand. Rather, RNAi activity is dependant only on the expression level, with both the U6 and 7SK promoters producing more active shRNAs compared to the H1 promoters.

Although anti-HIV shRNAs are potent inhibitors of viral replication and are among the top candidates for combination anti-HIV gene therapy, there is a need to ensure that they do not negatively impact the cells they are expressed in (HSCs and their progeny cells). A potential negative attribute of using a promoter with a more promiscuous start site is that potential off-targeting could be increased due to the increased diversity of the transcripts. While our results suggest the guide strand identities are not affected by the more promiscuous transcriptional start sites of some promoter-shRNA combinations, there is certainly an increase in the diversity of the passenger strands, which could be a source of off targeting. Several potential mechanisms of shRNA toxicity have been described, including activation of innate immune responses[43-45], off-target effects on human RNAs[46,47] and saturation of components of the RNAi pathway[48]. Two studies showed that adverse effects of shRNAs expressed from the U6 promoter could be prevented by using the H1 promoter instead[25,26]. For this reason an shRNA targeting CCR5 expressed from the H1 promoter[40,42] was chosen for a combination anti-HIV gene therapy clinical trial[49]. While toxicity of shRNAs expressed from the U6 promoter has been observed, an shRNA targeting tat/rev expressed from the U6 promoter was shown to be safe in preclinical and clinical studies[4,36]. This suggests that toxicity of U6 promoted shRNAs may be related only to particular shRNA sequences.

We evaluated if there was cellular toxicity in response to the shRNAs expressed from the different promoters to determine whether the inhibition of HIV-1 production seen in response to co-transfection was a consequence of cell death. Our experiments confirmed that the expression of

the different shRNAs does not cause cytotoxicity in HEK293T cells (Figure 2.3). In contrast, when we transduced different shNS and sh1498 promoted constructs into SupT1 cells, cytotoxicity was evident for some constructs by a marked decrease in transduced cells (GFP positive) versus untransduced cells (GFP negative) over time (Figure 2.6D). Regardless of the promoter used, sh1498 transduced cells had a greater growth disadvantage compared to shNS transduced cells, suggesting that the mechanism of the growth defect in sh1498 transduced cells was at least partially sequence dependent. Interestingly, while the decline in GFP positive cells was similar for U6 and 7SK sh1498 cultures, only U6 shNS cultures had a noticeable decline with both H1 and 7SK shNS cultures remaining stable at approximately 50% out to 57 days. Further studies will be needed to determine if the mechanism of the growth defect in GFP positive cells is a direct toxicity triggered by the different promoter-shRNA combinations and to evaluate whether different results can be obtained using different anti-HIV-1 shRNA sequences.

Although RNA Pol III promoters are typically used to express therapeutic small RNAs such as CRISPR guide RNAs and shRNAs, few studies have compared the three commonly used promoters for efficacy, RNA accumulation and transcriptional profile of their intended therapeutic products. Using three different shRNAs we show that both the U6 and 7SK promoters consistently express higher amounts of shRNAs compared to the H1 promoter and that this leads to an increase in potency in the case of the two anti-HIV shRNAs. Consistent with other studies our RNA seq data shows that all RNA Pol III promoters produce shRNA transcripts with different start[30] and end[24] sites. We also show that the U6 promoter consistently uses the +1 transcriptional start site most frequently but that in a T lympohcyte cell line both an anti-HIV-1 and a nonsense shRNA expressed from this promoter confers a negative impact on cell growth. To the best of our knowledge, this is the first study to use RNA seq data to examine the Dicer cleavage sites of

shRNAs. These results suggest that contrary to established rules for Dicer cleavage in mammalian cells[32], shRNAs are not cleaved at a set distance from their 5' ends, but rather at a set distance from the first base pair in their stem. Consequently, the guide strands produced from different promoters are largely the same, regardless of how accurate the promoter is at starting transcription at the intended +1 position. Furthermore, we show that the shRNA sequence affects the uniformity of Dicer cleavage.

Overall, our results highlight the unpredictability of shRNA transcription and processing in human cells as well as underscore the importance of evaluating different promoters for any particular shRNA gene therapy candidate. For an HIV-1 functional cure using shRNAs several parameters need to be considered to ensure the safe and effective creation of HIV-1 resistant cells. Our results highlight the importance of shRNA promoter choice and demonstrate that expression level is most important for shRNA activity and toxicity. Testing alternative anti-HIV-1 shRNAs using the different promoters described in this study will be required to identify a combination of shRNA and promoter that is effective in safely generating HIV-1 resistant cells in combination with other anti-HIV-1 RNAs.

2.6 Materials and methods.

2.6.1 Cell culture

Human embryonic kidney (HEK) 293T cells were grown in Dulbecco's modified Eagle's medium (HyClone, Logan, UT) supplemented with 10% fetal bovine serum (FBS, HyClone), 50 μg/ml streptomycin, and 50 U/ml penicillin (Life Technologies, Carlsbad, CA). SupT1 cells were grown in Roswell Park Memorial Institute Medium 1640 (HyClone), supplemented with 10% heat inactivated FBS (HyClone), 50 μg/ml streptomycin, and 50 U/ml penicillin (Life Technologies).

2.6.2 Vector construction

The U6 promoter was amplified by PCR using the pSIREN-shuttle vector (Clontech Laboratories, Mountain View, CA) as template and the following primers:

Forward: 5'-GCGCTATCGATGGAAGAGGCTATTTCCCA-3'

Reverse: 5'-GCGGAGGTACCGTCCTTTCCACAAGATAT-3'

These amplicons were digested with Acc651 (isoschisomer of KpnI) and ClaI then ligated into Acc651 and ClaI digested psiRNA-7SKGFP::Zeo (InvivoGen, San Diego, CA) plasmid to create psiRNA-U6GFP::Zeo. The DNA inserts coding for the shRNAs were generated by annealing complementary oligonucleotides as described previously[17,50]. These DNA inserts were then ligated into BbsI digested psiRNA-7SKGFP::Zeo, as well as into BbsI digested psiRNA-U6GFP::Zeo. Constructs containing the H1 promoter were previously described[17].

2.6.3 Mutagenesis

All constructed plasmids originating from the psiRNA-H1GFP::Zeo (InvivoGen) and the constructed psiRNA-U6GFP::Zeo plasmids were mutated using the QuikChange II site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. The primers used in the mutagenesis, with the nucleotides which conferred mutations within the promoter highlighted in bold are shown in Table 2.7.

Table 2.7 Primers used for site directed mutagenesis.

Primer name	Primer sequence 5'-3' (Mutagenesis nucleotides in bold)
H1sh1498F	TCTTATAAGTTCTGTATGAGACCAC**TCTTTCCC**GCAGGAACTACTAGTACCCTACTC
H1sh1498R	GAGTAGGGTACTAGTAGTTCCTGC**GGGAAAGA**GTGGTCTCATACAGAACTTATAAGA
H1sh5983F	TATAAGTTCTGTATGAGACCAC**TCTTTCCC**GCGGAGACAGCGACGAAGAGG
H1sh5983R	CCTCTTCGTCGCTGTCTCCGC**GGGAAAGA**GTGGTCTCATACAGAACTTATA
U6sh1498F	CTTTATATATCTTGTGGAAAGGAC**GAAACACC**GCAGGAACTACTAGTACCCTACTC
U6sh1498R	GAGTAGGGTACTAGTAGTTCCTGC**GGTGTTTC**GTCCTTTCCACAAGATATATAAAG
U6sh5983F	GGCTTTATATATCTTGTGGAAAGGAC**GAAACACC**GCGGAGACAGCGACG

U6sh5983R	CGTCGCTGTCTCCGC**GGTGTTTC**GTCCTTTCCACAAGATATATAAAGCC
H1shNSF	CTTATAAGTTCTGTATGAGACCAC**TCTTTCCC**GTACCGCACGTCATTCGTATCCT
H1shNSR	AGGATACGAATGACGTGCGGTAC**GGGAAAGA**GTGGTCTCATACAGAACTTATAAG
U6shNSF	ATATATCTTGTGGAAAGGAC**GAAACACC**GTACCGCACGTCATTCGTATCC
U6shNSR	GGATACGAATGACGTGCGGTAC**GGTGTTTC**GTCCTTTCCACAAGATATAT

2.6.4 Transfections and RT assay

HEK 293T cells were plated into a 96 well plate at 5×10^4 cells/well, 24 h prior to transfections. Co-transfections were performed with 50 ng HIV-1 molecular clone (pNL4-3) and each plasmid construct at 0.05 to 50 ng using TransIT-LT1 (Mirus Bio, Madison, WI) according to the manufacturer's protocol. Culture supernatants were collected 48 h after transfection and viral production was measured by RT assay as previously described[51,52]. Briefly, 5 µl of supernatant was incubated with 50 µl of RT cocktail containing a poly(A) template (Roche, Basel, Switzerland), an oligo(dT) primer (Life Technologies) and [^{32}P] dTTP (3,000 Ci/mmol; Perkin Elmer, Waltham, MA) for 2 h at 37°C. The poly dT RT product was then detected by spotting 5 µl of the reaction mixture onto DEAE filter mats (Perkin Elmer), washing away unincorporated [^{32}P] dTTP with 2× SSC and measuring counts per minute (cpm) on a microplate scintillation counter (MicroBeta TriLux; Perkin Elmer). The amount of HIV-1 RT enzyme in the supernatants is proportional to the cpm readout.

2.6.5 WST-1 assay

HEK 293T cells were plated into a 96 well plate at 5×10^4 cells/well, 24 h prior to transfections. Transfections were performed with 1.5 or 2 µg of each plasmid construct and TransIT-LT1 (Mirus) according to the manufacturer's instructions. Water soluble tetrazolium salt (WST-1) assay (Roche Applied Science, Penzberg, Germany) was performed 48 h after transfection according to the manufacturer's protocol. Cell viability was measured 2 h after the addition of the WST-1 reagent as previously described[53].

2.6.6 RNA extraction for Northern blot

HEK 293T cells were plated in a 10 mL tissue culture dish at 3.5×10^6 cells/dish, 24 h prior to the transfections with 1 μg of plasmid. Cell lysates were recovered 48 h after transfection using TRIzol reagent (Life Technologies). Total RNA from the recovered cell lysates was obtained by phenol chloroform extraction followed by clean up with a miRNeasy mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

2.6.7 Northern blot

Following the RNA extraction, a total of 15 μg of total RNA from each sample was mixed with equal volume of 2× gel loading buffer. RNA was resolved in a 15% polyacrylamide-urea gel as described previously[54]. Briefly, the RNA was transferred to a neutral charged nylon membrane (Hybond-N, Amersham Biosciences, Little Chalfont, UK) with a semi-dry electroblotter (40 min, 4°C, 20 V). Membranes were prehybridized in prehybridization buffer composed of 6× SSC, 2× Denhardt's solution and 0.1% SDS. Hybridization of the ^{32}P-labeled RNA probes to the membrane was done at 37°C overnight in hybridization buffer of identical composition as the prehybridization buffer. Once probe hybridization was complete, the membrane was washed for 15 min in wash buffer #1 (2× SSC), 15 minutes in wash buffer #2 (1× SSC), and 15 min in wash buffer #3 (0.1× SSC), all at 37°C. Northern bands were exposed on radiographic film and analyzed with the Fiji software[55] to generate raw numbers proportional to the exposure intensity of the bands. The band corresponding to each shRNA was standardized to the band corresponding to the 5S RNA loading control of the samples to generate percentage-based values of the shRNA band intensities compared to the 5S band intensities.

2.6.8 Probe labeling

To generate radioactively labeled RNA probes, *in vitro* transcription was performed using the HiScribe T7 kit (NEB) according to the manufacturer's instructions with annealed template DNA and [α-32P] CTP (800 Ci/mmol; Perkin Elmer). Following *in vitro* transcription, the reaction mixture was treated with DNase I (New England Biolabs, Ipswich, MA) to eliminate the template DNA. The labeled RNA probes were then purified with ProbeQuant G50 Micro Columns (GE Healthcare, Little Chalfont, UK) to remove unincorporated nucleotides. The probe sequences with the T7 promoter sequence in bold were:

sh1498-antisense:GGGTACTAGTAGTTCCTGCC**TATAGTGAGTCGTATTA**ATTTC,

sh5983-antisense:TCTTCGTCGCTGTCTCCGCC**TATAGTGAGTCGTATTA**ATTTC,

shNS-antisense:TACGAATGACGTGCGGTACC**TATAGTGAGTCGTATTA**ATTTC,

5S-rRNA-antisense:GGGAATACCGGGTGCTGTAGGCTTTCC**TATAGTGAGTCGTATTA**ATTTC

and T7-promoter-sense:GAAAT**TAATACGACTCACTATA**

2.6.9 SupT1 T cell transduction, infection and competitive growth

Promoter shRNA cassettes were subcloned from the psiRNA plasmids into the lentiviral transfer vector HIV-7-EGFP (donated by Dr. J. Rossi)[37] using forward primer 5'-TATGCGGCCGCAGGGATTTTGGTCATGTTCTTAATCGATACTA-3' and reverse primer 5'-GTAACGCCTGCAGGTTAATTAAGTCTAGAAGCTTTTCCAA-3' and restriction sites NotI and XbaI. Lentiviral transfer vectors were cotransfected into HEK293T cells with a plasmid expressing vesicular stomatitis virus G protein (from Dr. J. Rossi) and the packaging plasmid psPAX2 (Addgene, no. 12260). The supernatant was collected 48 h post-transfection and the

lentiviruses were concentrated using Lenti-X (Clontech laboratories) following the manufacturer's protocol. Lentivirus titres were determined using percent GFP positive SupT1 cells transduced with a range of dilutions (1 in 4 to 1 in 2048). In 5 mL cultures, 1.05×10^6 SupT1 cells were transduced with lentiviruses at a MOI of 1 with 8 µg/mL Polybrene (Sigma-Aldrich, St. Louis, MO). Cells were sorted 72-96 h after transduction for GFP expression using the gates shown in Figures 2.8 and 2.9 with a FACSAria Fusion cell sorter (BD biosciences Franklin Lakes, NJ). Cells were then plated in 96 well round bottom plates at 2×10^4 cells/well for HIV-1 infection and for competitive growth. Cells plated for HIV-1 infection were infected with HIV-1 NL4-3 (1750 cpm/mL, determined using the HIV-1 RT assay) 24 h later and HIV-1 RT activity was determined on various days post infection in the culture supernatant at which times 100 µL of supernatant was collected and replaced with 100 to 110 µL of fresh media. Cells plated for competitive growth were immediately mixed with 2×10^4 cells/well of untransduced SupT1 cells that had been passed through the flow cytometer without sorting. GFP positive cell percentage was determined at various days post mixing using a LSRFortessa flow cytometer (BD biosciences) at which times 100 µL of cells were collected and replaced with 100 to 110 µL of fresh media.

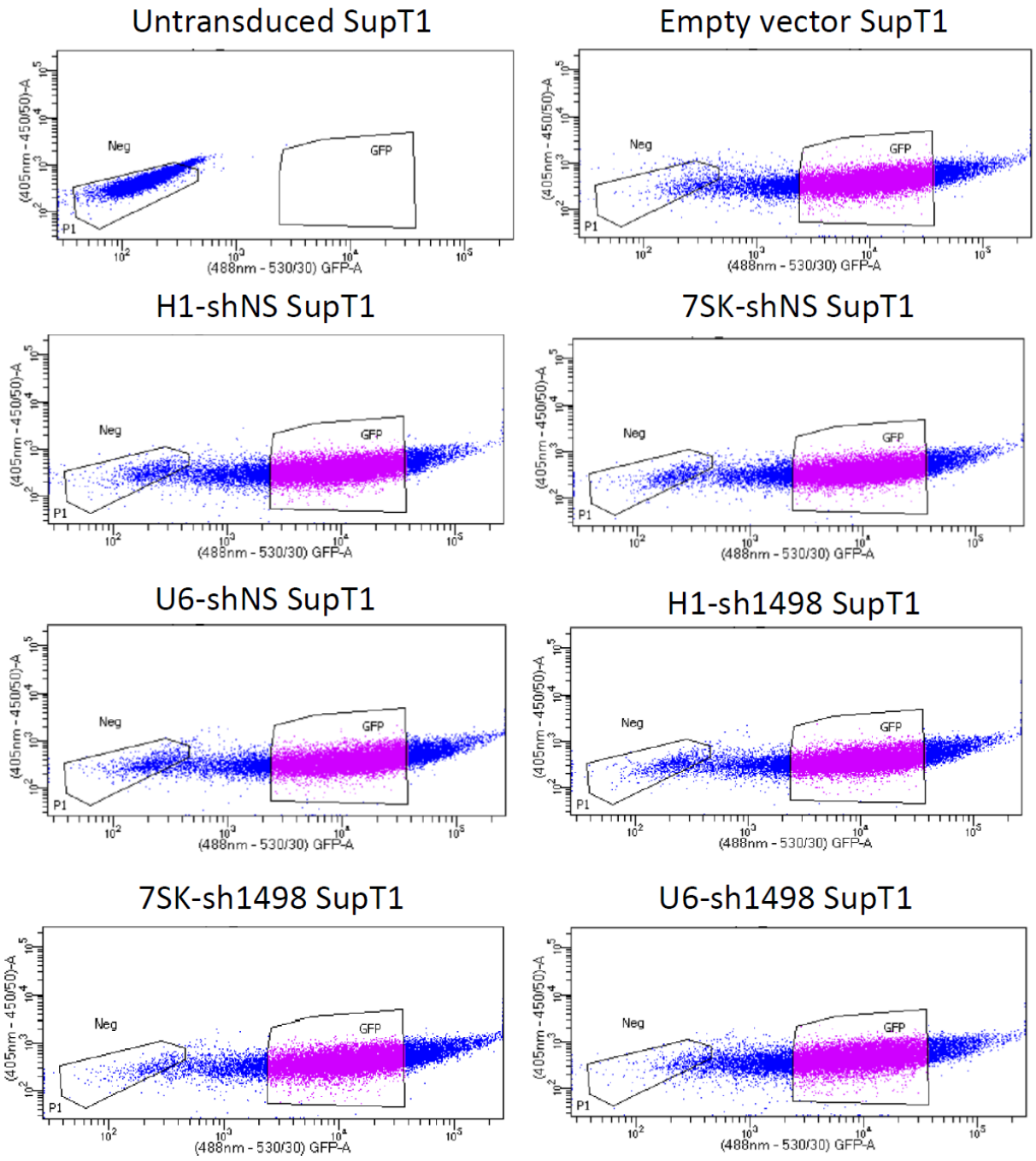

Figure 2.8 Gating parameters used to sort transduced cells for GFP expression. Parameters used to sort transduced cells prior to infection for results shorn in Figure 2.6A and B.

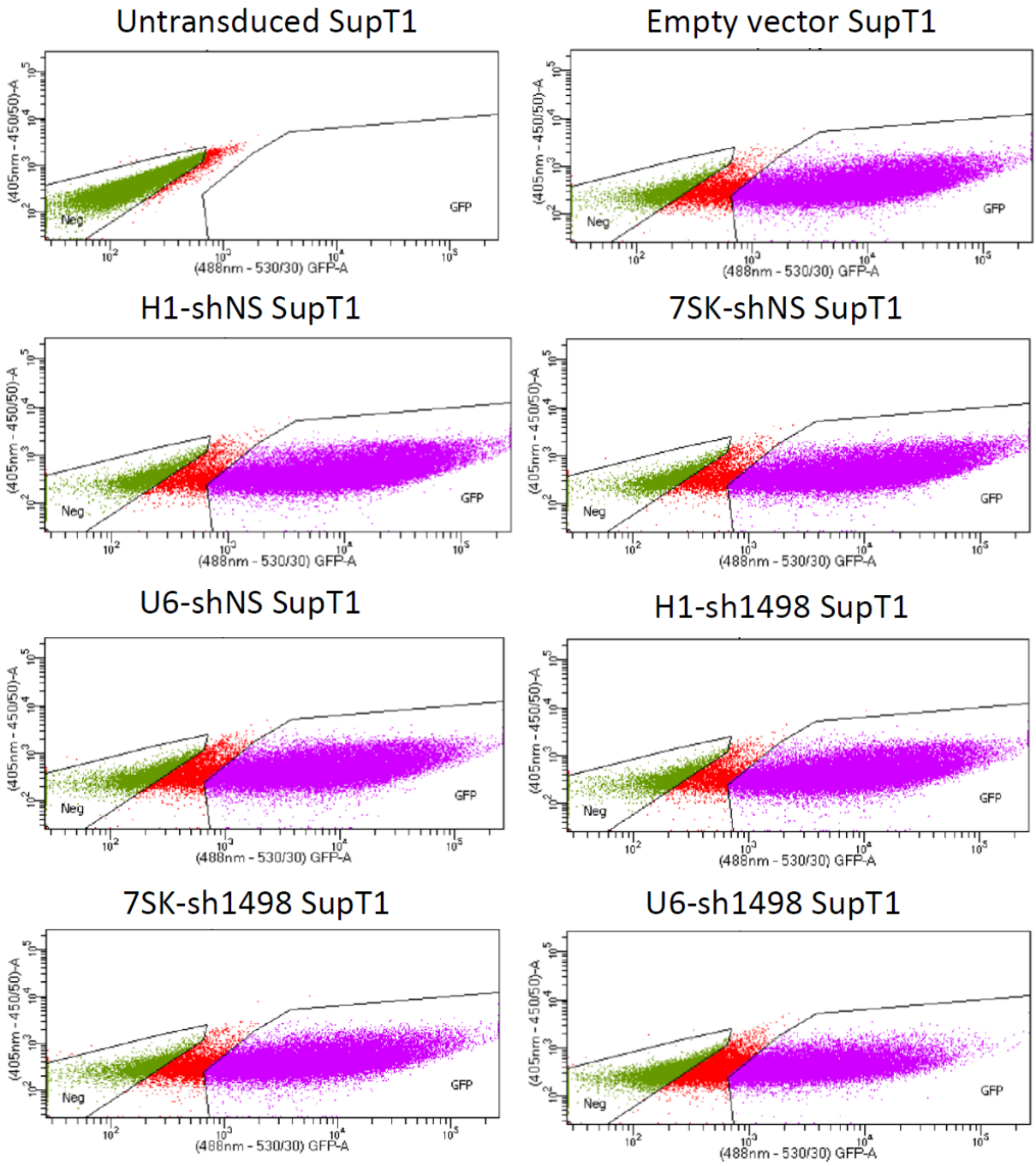

Figure 2.9 Gating parameters used to sort transduced cells for GFP expression. Parameters used to sort transduced cells prior to infection and competitive growth assays for results shown in Figure 2.6C and D.

2.6.10 RNA extraction for RNA seq

HEK 293T cells were plated in a flask at 5×10^6 cells/flask, 24 h prior to the transfection. The cells were then transfected with 5 µg of plasmids expressing sh1498, shNS and sh5983 expressed from the promoters U6, 7SK, H1, hU6 or hH1. An individual flask was used to transfect the three shRNAs, one flask for each promoter. Cell lysates were recovered after 48 h, using TRIzol reagent (Life Technologies). Total RNA was isolated using phenol chloroform extraction followed by clean up with a RNeasy mini kit (Qiagen) according to the manufacturer's instructions.

2.6.11 RNA sequencing

Libraries were prepared from total RNA by Genome Quebec (Montréal, Canada) using a NEB small RNA library kit with size selection. The libraries were run in an Illumina HiSeq4000 SR50 sequencing lane. Data analysis was carried out by the Canadian Center for Computational Genomics (Montréal, Canada). Briefly, adaptor sequences were clipped from the reads, but reads were not trimmed, to avoid introducing false variants. Reads were mapped to the different promoter shRNA sequences. Reads smaller than 6 base pairs and singletons were removed, and the remaining sequences were arranged by read number. Reads were then expressed as a percentage of total reads for each promoter shRNA sequence (Supplementary Files S1 to S5, found at the end of the following link: https://www.cell.com/molecular-therapy-family/nucleic-acids/fulltext/S2162-2531(20)30403-0). A single table was created for each shRNA, reporting the % of total reads for each variant that occurred at greater than 0.1% of total reads (Tables 2.2-2.4).

2.6.12 Statistical Analysis

A two-way ANOVA with a Bonferroni post-test was used to compare replicate means of each test shRNA to replicate means of empty psiRNA transfected cells in Figures 2.1 and 2.3. The same test was used to compare replicate means of infection and competitive growth time course data to

means of empty vector transduced cells (Figure 2.6). Non-linear regression extra sum-of-squares F test was used to compare LogIC50s in Figure 2.4. Unpaired student t-tests were used to compare mean RT activity and relative gene expression data in Figure 2.2 and 2.5, respectively. Graph Pad Prism Version 5.03 was used to perform all statistical analyses (GraphPad, San Diego, CA).

2.7 References.

1. De Clercq, E. Antiretroviral drugs. *Curr Opin Pharmacol* **10**, 507-515 (2010).
2. Le Douce, V., *et al.* Achieving a cure for HIV infection: do we have reasons to be optimistic? *J Antimicrob Chemother* **67**, 1063-1074 (2012).
3. An, D.S., *et al.* Stable reduction of CCR5 by RNAi through hematopoietic stem cell transplant in non-human primates. *Proc Natl Acad Sci U S A* **104**, 13110-13115 (2007).
4. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).
5. Li, M.J., *et al.* Inhibition of HIV-1 infection by lentiviral vectors expressing Pol III-promoted anti-HIV RNAs. *Mol Ther* **8**, 196-206 (2003).
6. Li, M.J., *et al.* 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**, 900-909 (2005).
7. Kohn, D.B., *et al.* A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children. *Blood* **94**, 368-371 (1999).
8. Hutter, G., *et al.* Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* **360**, 692-698 (2009).
9. Gupta, R.K., *et al.* HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* **568**, 244-248 (2019).
10. Scarborough, R.J., Goguen, R.P. & Gatignol, A. A second patient cured of HIV infection: hopes and limitations. *Virologie (Montrouge)* **23**, 1-4 (2019).
11. Rossi, J.J., June, C.H. & Kohn, D.B. Genetic therapies against HIV. *Nat Biotechnol* **25**, 1444-1454 (2007).
12. Burnett, J.C. & Rossi, J.J. RNA-based therapeutics: current progress and future prospects. *Chem Biol* **19**, 60-71 (2012).
13. Macpherson, J.L., *et al.* Long-term survival and concomitant gene expression of ribozyme-transduced CD4+ T-lymphocytes in HIV-infected patients. *J Gene Med* **7**, 552-564 (2005).
14. Mitsuyasu, R.T., *et al.* Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nature medicine* **15**, 285-292 (2009).
15. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).

16. Scarborough, R.J. & Gatignol, A. RNA Interference Therapies for an HIV-1 Functional Cure. *Viruses* **10**, 8 (2018).
17. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
18. Pang, K.M., Castanotto, D., Li, H., Scherer, L. & Rossi, J.J. Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. *Nucleic Acids Res* **46**, e6 (2018).
19. Muller-Kuller, T., Capalbo, G., Klebba, C., Engels, J.W. & Klein, S.A. Identification and characterization of a highly efficient anti-HIV pol hammerhead ribozyme. *Oligonucleotides* **19**, 265-272 (2009).
20. Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. *Mol Ther Nucleic Acids* **18**, 815-830 (2019).
21. Cramer, P., *et al.* Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**, 337-352 (2008).
22. Arimbasseri, A.G., Rijal, K. & Maraia, R.J. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* **5**, e27639 (2014).
23. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**, 2593-2620 (2002).
24. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. Delineation of the Exact Transcription Termination Signal for Type 3 Polymerase III. *Mol Ther Nucleic Acids* **10**, 36-44 (2018).
25. An, D.S., *et al.* Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors. *Mol Ther* **14**, 494-504 (2006).
26. Sun, C.P., *et al.* Studies of efficacy and liver toxicity related to adeno-associated virus-mediated RNA interference. *Hum Gene Ther* **24**, 739-750 (2013).
27. Makinen, P.I., *et al.* Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* **8**, 433-441 (2006).
28. Spanevello, F., *et al.* Development of Lentiviral Vectors Simultaneously Expressing Multiple siRNAs Against CCR5, vif and tat/rev Genes for an HIV-1 Gene Therapy Approach. *Mol Ther Nucleic Acids* **5**, e312 (2016).
29. Brake, O.T., *et al.* Lentiviral Vector Design for Multiple shRNA Expression and Durable HIV-1 Inhibition. *Mol Ther* **16**, 557-564 (2008).
30. Ma, H., *et al.* Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids* **3**, e161 (2014).
31. Gao, Z., Harwig, A., Berkhout, B. & Herrera-Carrillo, E. Mutation of nucleotides around the +1 position of type 3 polymerase III promoters: The effect on transcriptional activity and start site usage. *Transcription* **8**, 275-287 (2017).
32. Park, J.E., *et al.* Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**, 201-205 (2011).
33. Macrae, I.J., *et al.* Structural basis for double-stranded RNA processing by Dicer. *Science* **311**, 195-198 (2006).
34. Keefe, A.D., Pai, S. & Ellington, A. Aptamers as therapeutics. *Nature Reviews Drug Discovery* **9**, 537-550 (2010).

35. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. Improvement of the CRISPR-Cpf1 system with ribozyme-processed crRNA. *RNA Biol* **15**, 1458-1467 (2018).
36. Lee, N.S., *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20**, 500-505 (2002).
37. Yam, P.Y., *et al.* Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol Ther* **5**, 479-484 (2002).
38. Eekels, J.J., *et al.* A competitive cell growth assay for the detection of subtle effects of gene transduction on cell proliferation. *Gene Ther* **19**, 1058-1064 (2012).
39. Qin, X.F., An, D.S., Chen, I.S. & Baltimore, D. 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**, 183-188 (2003).
40. Shimizu, S., *et al.* Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**, 8 (2009).
41. Chung, J., *et al.* Endogenous MCM7 microRNA cluster as a novel platform to multiplex small interfering and nucleolar RNAs for combinational HIV-1 gene therapy. *Hum Gene Ther* **23**, 1200-1208 (2012).
42. Burke, B.P., *et al.* Engineering Cellular Resistance to HIV-1 Infection In Vivo Using a Dual Therapeutic Lentiviral Vector. *Mol Ther Nucleic Acids* **4**, e236 (2015).
43. Fish, R.J. & Kruithof, E.K. Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. *BMC Mol Biol* **5**, 9 (2004).
44. Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**, 263-264 (2003).
45. Olejniczak, M., *et al.* Sequence-non-specific effects generated by various types of RNA interference triggers. *Biochim Biophys Acta* **1859**, 306-314 (2016).
46. Jackson, A.L., *et al.* Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**, 635-637 (2003).
47. Scacheri, P.C., *et al.* Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc Natl Acad Sci U S A* **101**, 1892-1897 (2004).
48. Grimm, D., *et al.* Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**, 537-541 (2006).
49. Delville, M., *et al.* Safety of CD34(+) Hematopoietic Stem Cells and CD4(+) T Lymphocytes Transduced with LVsh5/C46 in HIV-1 Infected Patients with High-Risk Lymphoma. *Mol Ther Methods Clin Dev* **13**, 303-309 (2019).
50. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).
51. Scarborough, R.J., Levesque, M.V., Perreault, J.P. & Gatignol, A. Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* **1103**, 31-43 (2014).
52. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. & Gatignol, A. Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp* **115**, e54486 (2016).
53. Alpuche-Lazcano, S.P., *et al.* Higher Cytopathic Effects of a Zika Virus Brazilian Isolate from Bahia Compared to a Canadian-Imported Thai Strain. *Viruses* **10**, 53 (2018).

54. Goguen, R.P., Gatignol, A. & Scarborough, R.J. Cloning and Detection of Aptamer-Ribozyme Conjugations. *Methods Mol Biol* **2167**, 253-267 (2021).
55. Schindelin, J., *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).

Chapter III

Aptamer-shRNA chimeras as a molecular design to prevent shRNA-mediated cytotoxicity.

This chapter is adapted from the following manuscript: Goguen R.P., Chen, M.J., Daher, A., Gatignol, A., Scarborough, R.J. Aptamer-shRNA chimeras as a molecular design to prevent shRNA-mediated cytotoxicity. In preparation.

3.1 Preface.

The study presented in chapter II investigated various expression strategies of anti-HIV shRNAs. Although antiviral activity was maximized when shRNAs were expressed from the U6 and 7SK promoters, this came at a cost of severe cytotoxicity. Optimization of an antiviral therapy must not only focus on maximizing therapeutic effects, but also eliminating adverse effects. With this in mind, our anti-HIV gene therapy must be further optimized by reducing its associated cytotoxicity. Through our investigation of the transcriptional activity of the U6, 7SK and H1 promoter, we observed higher shRNA expression levels from the U6 and 7SK promoters compared to the H1 promoter and concluded that these higher expression levels can cause cytotoxicity. Based on this finding, reducing the levels of circulating shRNAs and its metabolic products in cells may serve as a means of preventing cytotoxicity.

This chapter evaluates the effects of various molecular designs of anti-HIV aptamers and shRNAs. Here, we show that the generation of aptamer-shRNA chimeras by incorporating aptamers within the terminal loop of shRNAs can eliminate the cytotoxicity observed from the expression of canonical shRNAs. Additionally, we show preliminary work to elucidate the mechanism of this effect where guide strands are present in greater quantities in cells expressing canonical shRNAs and thus mediate cytotoxicity.

The results from this study demonstrate that anti-HIV gene therapy can be optimized by exploring alternative molecular designs to avoid cytotoxicity. In addition, these results further support our chapter II hypothesis that shRNA-mediated cytotoxicity caused from the expression of the U6 and 7SK promoters are due to higher expression levels. In the case of either shRNA expression from the H1 promoter or expression of aptamer-shRNA chimeras, low amounts of guide strands were present in cells which corresponds to a low cytotoxic potential. These findings identify a unique

characteristic of aptamer-shRNA conjugation that decreases the cytotoxicity of the shRNA, which contributes to their safety in a gene therapy setting.

CONFLICT OF INTEREST:

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We would like to thank Dr. John Rossi (City of Hope, Duarte, CA, USA) for provision of lentiviral plasmids. We would also like to thank Christian Young (Lady Davis Institute flow cytometry manager) for assistance in cell sorting. This work was supported by grant PJT-148704 awarded to AG and RJS from the Canadian Institutes of Health Research (CIHR). RPG is supported by a Doctoral fellowship and MJC by a Masters fellowship from the Fonds de Recherche du Québec-Santé (FRQS).

3.2 Abstract.

Gene therapy to treat HIV infection has the potential to eliminate the need for cART, with clinical trials having been performed to assess the safety and efficacy of combination gene therapy using multiple anti-HIV RNAs. However, the limited effectiveness of these therapies thus far suggests that further optimization of individual antiviral RNAs is necessary. We explored numerous molecular designs of anti-HIV aptamers and shRNAs to maximize their inhibitory capabilities while avoiding cytotoxicity. Specifically, flanking hairpins and self-cleaving ribozymes were added at the 5' and 3' end of aptamers. However, regardless of the molecular design, expression of anti-HIV aptamers did not result in strong inhibition of HIV production. We then combined anti-HIV aptamers and shRNAs into a single transcript by incorporating the aptamers within the terminal loop of the shRNAs. These aptamer-shRNA chimeras could inhibit viral replication without inducing cytotoxicity, effectively eliminating the cytotoxicity that we previously reported from U6 and 7SK driven canonical shRNAs. We measured the amount of shRNA guide strands present in cells expressing aptamer-shRNA chimeras compared to cells expressing canonical shRNAs. We show that guide strands were more abundant in cells expressing canonical shRNAs, suggesting that the aptamer-shRNA chimeras reduce cytotoxicity, at least in part, through reducing the accumulation of guide strands in the cells.

3.3 Introduction.

Current therapies used to treat human immunodeficiency virus type 1 (HIV) infection can greatly prolong the lifespan of infected individuals and have revolutionized the prognosis of the infection[1]. However, combination antiretroviral therapy (cART) is not a curative strategy and lifelong chronic administration of therapies can accrue high financial costs along with the development of side effects. Engineering patient cells to express anti-HIV RNA molecules represents a promising treatment option to eliminate the need of cART and its complications[2,3]. Antiviral RNAs that have been used to inhibit HIV replication include short hairpin RNAs (shRNAs), aptamers, ribozymes and U1 interference RNA[4-8].

Inhibition of gene expression by shRNAs occurs by co-opting elements of the RNA interference (RNAi) pathway with a net result of mRNA target cleavage. The endogenous function of RNAi is to regulate gene expression through the use of microRNAs (miRNAs) and this has an important impact on cell growth, metabolism and other biological processes[9]. Similar to miRNAs, shRNAs form a stem-loop structure that is recognized and cleaved by the RNase III endonuclease Dicer. As shRNAs are designed to have perfect complementarity to their target, they recruit the RNA-induced silencing complex (RISC) to mediate target cleavage by the endonuclease Argonaute2 (Ago2)[10]. Various shRNAs have been designed to inhibit HIV replication by interfering with the expression of viral proteins, with target sites including the LTR, *env*, *tat*, *pol*, *rev* and *gag* regions of the viral genome[4,5,10-16].

Aptamers differ from shRNAs in that they do not bind to their target by nucleotide complementarity and do not utilize cellular proteins for their inhibitory effects[2]. These single stranded RNA molecules instead fold into specific three-dimensional structures which allow them to bind to a particular target. Binding of aptamers to HIV proteins or RNA motifs renders these

viral elements inert and unable to perform their intrinsic function. Aptamers have been designed to bind and inactivate the HIV protease[17], reverse transcriptase[6], gag polyprotein[18], integrase[7] and untranslated region (UTR)[19].

cART is required in the clinic to maintain low levels of viremia while avoiding the occurrence of drug resistant virus[20,21]. Similarly, it will be necessary to use multiple anti-HIV RNAs to prevent the emergence of viral resistance in the face of gene therapy[11,22]. The earliest attempt to clinically evaluate the efficacy of HIV combination gene therapy employed a triple molecule therapy consisting of an shRNA targeting the Tat/Rev region, a TAR decoy and a CCR5-specific ribozyme[16]. Ongoing HIV gene therapy clinical trials include the use of a CCR5-specific shRNA along with a viral fusion inhibitor called C46 [NCT01734850] as well as the use of a TAR RNA decoy, a chimeric TRIM5α along with a CCR5-specific shRNA [NCT02797470]. While several anti-HIV RNAs have been expressed in patient cells and evaluated in clinical trials, more potent combinations are needed to effectively inhibit HIV replication and effectively generate HIV resistant cells.

A novel molecular design for anti-HIV RNAs has been the inclusion of aptamers as the terminal loop of shRNAs[7]. These aptamer-shRNA chimeras allow for the expression of two anti-HIV RNAs from a single RNA transcript. Processing of the primary transcript by Dicer will result in cleavage of the stem-loop and separation of the aptamer from the shRNA. This design provides a potential additional advantage in localizing the aptamer to the cytoplasm as it will be exported from the nucleus along with the shRNA through Exportin-5[7,23].

Another format used for aptamers is the addition of intentional flanking structures such as hairpins and self-cleaving ribozymes. RNA hairpins have been added to the 5’ and 3’ ends of aptamers in an attempt to increase RNA stability[19,24]. Aptamers and other small RNAs such as shRNAs can be

expressed from RNA Pol III promoters that have defined transcriptional start and end sites[5,25,26]. However, a 3' U tail of variable length is included in transcripts derived from these promoters[24]. To remove these sequences and generate transcripts of uniform length, the hepatitis delta virus (HDV) ribozyme has been added to the 3' ends of functional RNAs. This expression strategy has been accomplished with Ago-shRNAs and clustered regularly interspaced short palindromic repeats (CRISPR) guide RNAs[27,28]. The removal of extraneous sequences from transcription is also possible through the inclusion of hammerhead ribozymes at both the 5' and 3' ends[29].

The type 3 RNA Pol III promoters U6, H1 and 7SK have been the most utilized promoters for expression of small RNAs. Selection of appropriate promoters is critical to ensure maximal antiviral efficacy in the absence of cytotoxic effects. With this in mind, studies have shown that the expression of shRNAs from certain promoters can induce cytotoxicity. Specifically, shRNAs used to downregulate expression of cellular CCR5, an HIV coreceptor, were found to be cytotoxic when expressed from the U6 promoter but not from the H1 promoter[30,31]. Similarly, we have previously developed an shRNA targeting a conserved sequence in the *gag* region of HIV RNA, sh1498, and shown that it is more potent when expressed from the 7SK and U6 promoters compared to the H1 promoter, but that the higher expression levels driving this difference in potency also results in growth defects in transduced T lymphocytes[4,5]. Nevertheless, this trend in promoter induced cytotoxicity is not absolutely consistent, with one study showing no associated toxicity to the U6 promoter[32]. From this conflicting information, it seems apparent that the therapeutic molecule of interest also heavily influences the potential for RNA induced cytotoxicity.

In this study, we show that various secondary structures at the 5' and 3' ends of anti-HIV aptamers have minimal inhibitory effects on HIV production. We also incorporated aptamers within the terminal loop of sh1498 and shNS to generate aptamer-shRNA chimeras. We observed that when

large aptamers are included within the loop, the previously observed growth defects in response to expression of sh1498 from the U6 and 7SK promoters is effectively eliminated while maintaining the inhibition of HIV replication. Additionally, upon comparison of gene expression levels of guide strands from the sh1498 and the aptamer-shRNA chimeras, we observed a lower abundance of fully processed sh1498 guide strands from the expression of the aptamer-sh1498 chimera. These results represent a novel strategy to eliminate shRNA induced cytotoxicity from the RNA Pol III promoters during gene therapy.

3.4 Results.

3.4.1 Evaluation of aptamers with different conjugations against HIV production

To compare the potency of previously described anti-HIV aptamers, we expressed top performing protease aptamers (PR10.1-8E, PR10.1-8A, PR10.9-8N) [17], a gag aptamer (DP6-12)[18] and a UTR aptamer (RNApt16)[19] from the 7SK promoter with either no flanking structure, flanking hairpins (HP), flanking hammerhead (HH) or flanking HDV ribozymes. All structures are shown in Figure 3.1. Co-transfections were performed in HEK293T cells with the HIV-1 molecular clone pNL4-3 and each plasmid expressing a particular aptamer with its flanking structures. The activity of HIV RT was then measured in cell supernatants of the co-transfections to estimate the relative HIV production. All data were normalized to the RT activity measured in co-transfections of pNL4-3 with the empty vector (psiRNA-7SKGFP::Zeo). Of all the aptamer constructs that were screened, only HP-PR10.1-8E, HDV-PR10.1-8A and HDV-RNApt16 moderately inhibited HIV production by approximately 40%, with a maximal inhibition of 45% for HDV-RNApt16 (Figure 3.2). The aptamers PR10.9-8N and DP6-12 either did not inhibit viral production or had only a minimal inhibitory effect (Figure 3.2).

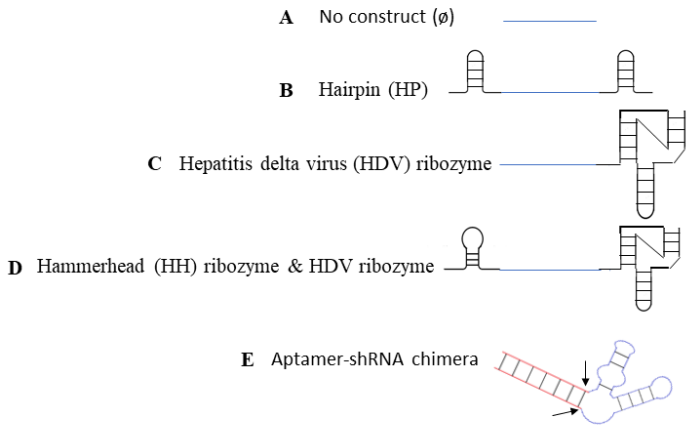

Figure 3.1 Schematic representation aptamers expressed with flanking secondary structure or incorporated within the terminal loop of shRNAs. Aptamer sequences depicted in blue are expressed either without secondary structures (A), with flanking hairpins (B), 3' HDV ribozyme (C) or 5' hammerhead ribozyme and 3' HDV ribozyme (D).

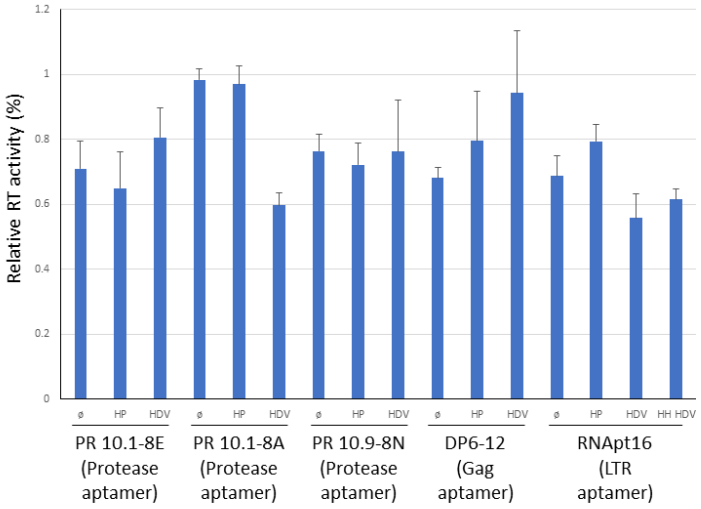

Figure 3.2 Aptamers expressed from 7SK promoter with various flanking secondary RNA structures are not potent inhibitors of HIV. HEK 293T cells were co-transfected with the HIV-1 molecular clone pNL4-3 along with 250ng of one of the plasmids expressing a particular aptamer from the 7SK promoter. Supernatants were collected 48 h posttransfection and virus production was estimated by measuring HIV RT activity. Data are expressed as a percentage of RT activity

in cells cotransfected with the empty expression plasmid. The antiviral potency of the aptamers was measured when these molecules were expressed with or without secondary structures. Each data point represents the mean +/- standard error mean (SEM) from at least one independent experiment with 2 replicates (n = 2-6).

3.4.2 Aptamer-shRNA chimeras act as inhibitors of viral replication

To evaluate the effects of shRNA-aptamer conjugation using our previously described sh1498, we incorporated the aptamers PR10.9-8N, RNApt16, Dp6.12 and S3R3 (protease, UTR, Gag and integrase[7] aptamers respectively) at the terminal loop of sh1498. These aptamers were also included in the loop of a non-sense shRNA (shNS), which does not target HIV RNA, to evaluate whether the aptamers can act as inhibitors of HIV replication without the presence of an antiviral shRNA. The structure of the various aptamer-shRNA chimeras is shown in Figure 3.3. Lentiviral vectors (HIV-7-GFP[33]) carrying each of these aptamer-shRNA chimeras driven from either the U6 or 7SK promoter were transduced into SupT1 lymphocytic cell line. To select for cells with properly integrated lentiviral vectors, cell sorting by GFP expression was performed and collected GFP positive cells were subsequently infected with HIV-1 NL4-3. HIV RT activity was measured in culture supernatants over time to evaluate the kinetics of viral replication in the transduced cells. All aptamer-shRNA chimeras except for 7SK driven S3R3-shNS had slight inhibitory effects on HIV replication compared to the empty vector transduced cells and no replication was detected in cells transduced by U6 driven RNApt16-sh1498 (Figure 3.4A-E).

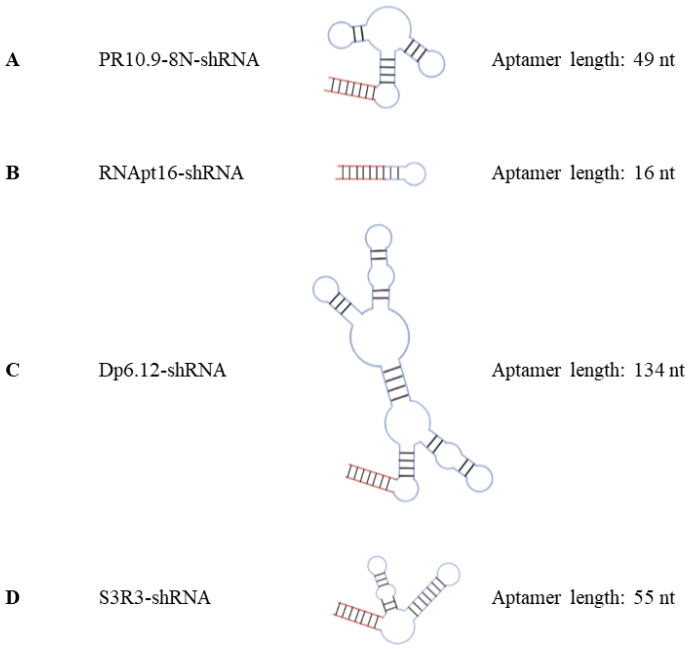

Figure 3.3 Schematic representation of aptamer-shRNA chimeras. Aptamers used to generate aptamer-shRNA chimeras in the study include PR10.9-8N (A), RNApt16 (B), Dp6.12 (C) and S3R3 (D). The aptamer sequence is represented in blue and the shRNA sequence is represented in red.

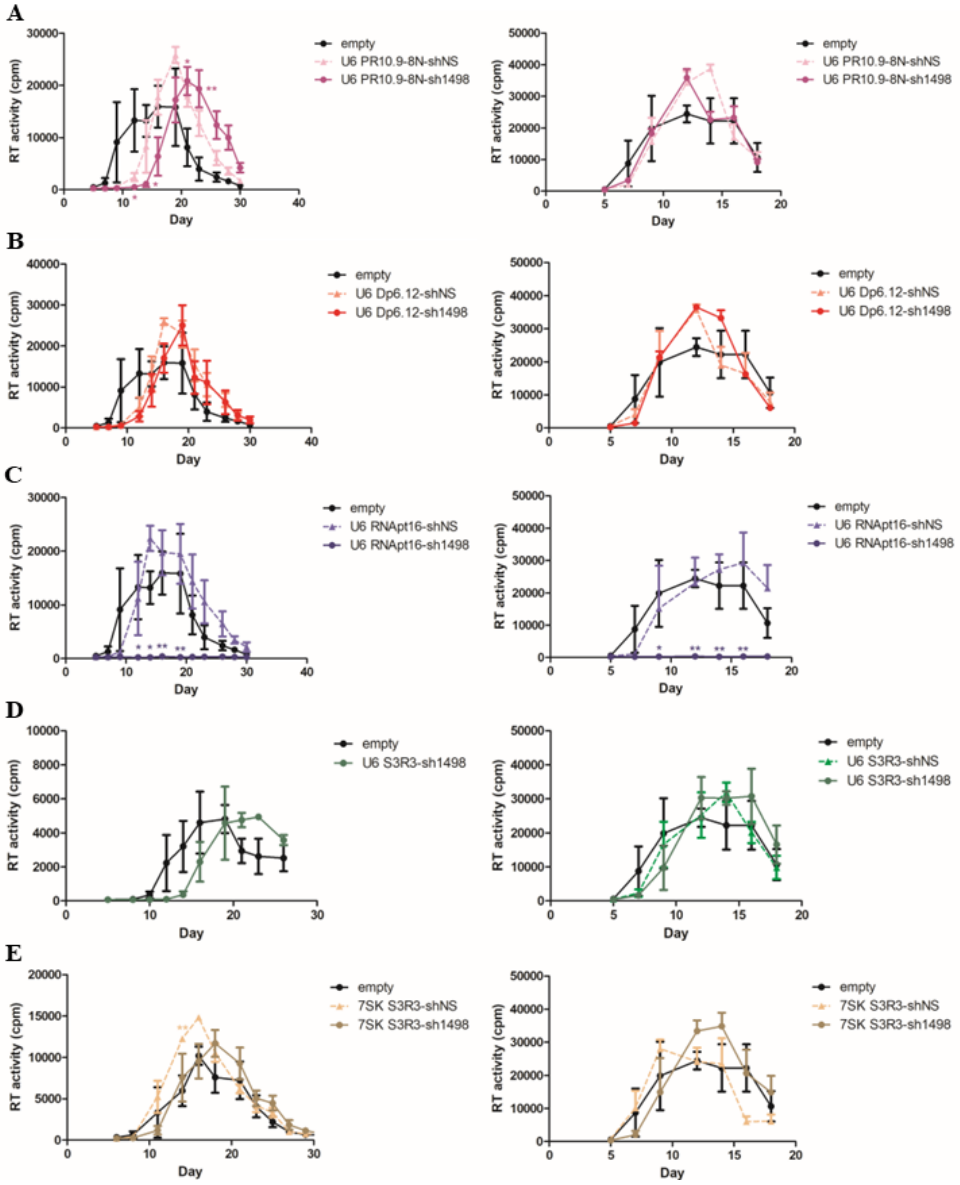

Figure 3.4 Aptamers-sh1498 chimeras driven by U6 or 7SK promoter inhibit HIV replication through the action of the shRNA. The aptamers PR10.9-8N (A), Dp6.12 (B), RNApt16 (C) and S3R3 (D & E) were incorporated within the terminal loop of shNS and sh1498 to generate aptamer-shRNA chimeras. SupT1 cells were transduced with HIV-7-EGFP lentiviral vectors expressing different aptamer-shRNA chimeras from the U6 or 7SK promoters and infected with HIV-1 NL4-3 at 1750 cpm/mL. The mean RT activity (cpm) was measured in culture supernatants at various days post infection. Each data point represents the mean +/- SEM from three infections (n=3). A two-way ANOVA with Bonferroni test was used to compare means to means of empty vector transduced cells. Significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) is shown for those data points that were significantly different from the empty vector transduced controls. Left and right panels are independent experiments performed in similar conditions.

3.4.3 shRNAs conjugated to large aptamers are not toxic in SupT1 cells

Following cell sorting of transduced SupT1 cells, competitive growth assays were performed by mixing GFP positive sorted transduced SupT1 cells with GFP negative sorted non-transduced SupT1 cells. The ratio of GFP negative to GFP positive cells was then measured over time to evaluate if a growth advantage existed for one of the two cell populations. Cells transduced with RNApt16-sh1498 showed large growth defects where % of GFP positive cells decreased over time to less than 10% (Figure 3.5A and D). Since a serious deficit in cell growth was observed in these cell cultures, it is likely that the absence of viral replication seen in U6 driven RNApt16-sh1498 is not entirely due to the chimeric molecule targeting HIV but that a loss of healthy cells from cytotoxic effects prevented proper establishment of an infection at the start of the experiment. We previously reported this same phenomenon when sh1498 was expressed alone from the U6 and 7SK promoters[5] (Chapter 2, Figure 2.6D). Additionally, cells expressing U6 driven RNApt16-shNS displayed some growth defects in one independent experiment (Figure 3.5A). However, the loss of healthy cells due to cytotoxic effects was not so severe as to prevent the establishment of infection as viral replication was detected in these cell cultures in an infection assay (Figure 3.4C). Unexpectedly, all other aptamer-shRNA chimeric cultures remained close to 50% GFP throughout the competitive growth assays (Figure 3.5A-D). It is important to note that the aptamer RNApt16 is an extremely small aptamer (Figure 3.3), which is similar in both molecular length and three-dimensional structure to the canonical shRNA loop. Comparatively, the aptamers S3R3, PR10.9-8N and Dp6.12 are all large aptamers which form three-dimensional structure that are largely dissimilar to the canonical shRNA loop. This suggests that the growth impairments associated to shRNAs when expressed from the U6 or 7SK promoters do not take place if the shRNA is

conjugated to large aptamers, but can still occur when the shRNA is conjugated to small aptamers which resemble the canonical shRNA loop.

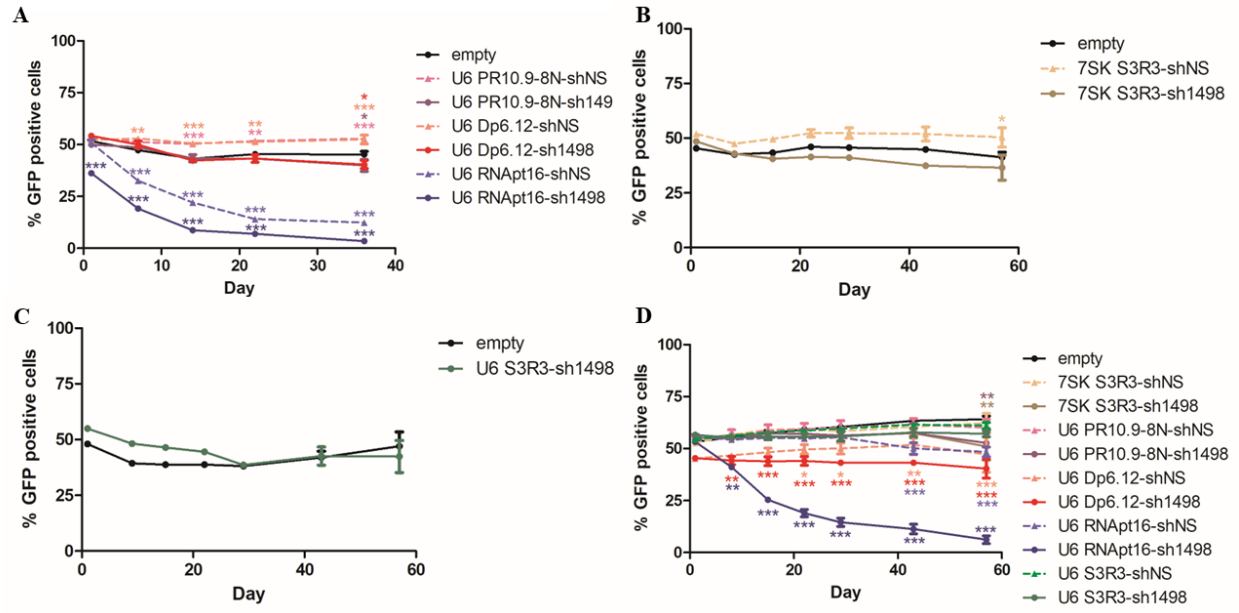

Figure 3.5 Growth defects are not detected when shRNAs are conjugated to large aptamers. (A-D) Panel D is a repeated independent experiment of panels A-C. Transduced SupT1 cells were mixed with untransduced SupT1 cells and the percentage of GFP positive cells was measured at various days post-mixing. Each data point represents the mean +/- SEM from three experiments (n=3). A two-way ANOVA with Bonferroni test was used to compare means to means of empty vector transduced cells. Significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) is shown for those data points that were significantly different from the empty vector transduced controls.

3.4.4 Guide strands from shRNAs are less abundant in cells expressing aptamer-shRNA chimeras compared to cells expressing shRNAs with canonical loops.

We hypothesized that the difference in growth kinetics resulting from the cellular expression of sh1498 alone compared to aptamer-sh1498 chimeras involving large aptamers may be due in part to differences in the total amount of guide strands present in the cell. This could be the result of differences in the total expression level of canonical shRNAs compared to aptamer-shRNA chimeras. Alternatively, expression levels may be similar for canonical shRNAs and aptamer-

shRNA chimeras but the efficacy of shRNA processing during biogenesis could differ. Both of these scenarios would result in differential levels of circulating shRNA guide strands in the cell. To determine if this is the case, we transfected cells with the HIV-7-EGFP plasmid expressing sh1498 or S3R3-sh1498 from the U6 or 7SK promoter and recovered the RNA to measure the expected shRNA guide strands by Northern blot. We observed that sh1498 guide strands were present in higher quantities when cells expressed sh1498 in its conical form compared to when cells expressed S3R3-sh1498 (Figure 3.6). It is therefore likely that the high abundance of sh1498 guide strands in cells expressing canonical sh1498 contributes to the observed growth defects during its expression from the U6 and 7SK promoters.

Figure 3.6 sh1498 guide strands are more abundant in cells expressing canonical sh1498 than in cells expressing S3R3-sh1498. RNA harvested from 293T cells transfected with either sh1498 or S3R3-sh1498 expression vectors was migrated in a 15% polyacrylamide-urea gel. shRNA guide strands were detected with a ^{32}P-labeled RNA probe.

3.5 Discussion

As HIV treatments first became available it was quickly discovered that multiple inhibitors of HIV must be used in combination to lower the levels of viremia in infected individuals and avoid the occurrence of resistant virus[2]. Efforts have been made to establish and evaluate gene therapies utilizing multiple anti-HIV genes, but it is unclear whether these therapies can control the infection in the absence of cART[16,34]. An important determinant to how well a gene therapy can inhibit HIV

replication relates to the optimization of the expression or molecular design of antiviral RNAs. We have previously shown that the promoter selected to drive the expression of anti-HIV genes influences the inhibitory capabilities of the product molecule as well as their potential to elicit cytotoxicity[5]. Specifically, shRNAs expressed from the U6 and 7SK promoters are generally more potent than when the H1 promoter is used to drive expression because of differences in transcriptional activity. However, this higher transcriptional activity associated to the U6 and 7SK promoters can also cause cytotoxicity during shRNA expression.

As the combination of different antiviral RNA molecules is necessary to establish an optimal anti-HIV gene therapy, we explored the possibility of combining aptamers to shRNAs that have HIV inhibitory capabilities. The evaluation of previously identified aptamers expressed from the 7SK RNA Pol III promoter either alone or with flanking ribozymes showed only minimal effectiveness except for the RNApt16 LTR aptamer (Figure 3.2). This suggests that aptamers expressed from the RNA Pol III promoters can only inhibit HIV if their molecular target is present in the nucleus. This is in conformity with previous assays where aptamers with cytoplasmic targets were expressed from an RNA Pol II promoter and could inhibit viral replication[17,18].

In contrast, shRNAs require expression from RNA Pol III promoters to inhibit HIV and we have previously shown for sh1498 a correlation between high antiviral potency and associated cytotoxicity[5]. We initially set out to incorporate several different aptamers to replace the loop of sh1498 and establish a combination gene therapy from a single transcript to determine how HIV replication would be affected by the expression of these aptamer-sh1498 chimeras compared to the sh1498 expressed alone. Only slight inhibitory effects from the expression of the different aptamer-sh1498 chimeras was seen in our inhibitory assays, except for cells expressing U6 driven RNApt16-sh1498 where no viral replication was detected (Figure 3.4). We previously observed

that viral replication also did not occur when sh1498 was expressed on its own from the U6 and 7SK promoters, but that a depletion of healthy cells early in the infection due severe cytotoxicity likely played a role in the absence of replication instead of the antiviral effects of the shRNA being solely responsible[5]. Our results in this study showed an identical pattern where the U6 driven RNApt16-sh1498 also caused severe growth defects in our competitive growth assays. As the other aptamer-sh1498 chimeras did not show growth defects when expressed from the U6 and 7SK promoters in our competitive growth assays, it is difficult to compare their inhibitory effects to U6 and 7SK driven sh1498. Indeed, our results showed that these chimeric molecules did not dramatically inhibit viral replication as was the case of sh1498 expressed alone, which could be explained by the absence of cytotoxicity providing a population of healthy cells to productively support viral replication. Nevertheless, it is likely that sh1498 is more effective at inhibiting viral replication in its canonical form than when it is expressed as an aptamer-shRNA chimera. Quantifying the difference in inhibitory effects between these two sh1498 molecular designs would be useful and could be addressed directly by co-transfection experiments using the different constructs.

Incorporating the aptamers within shNS allowed us to evaluate the antiviral effects of the aptamers when expressed as aptamer-shRNA chimeras. With the exception of 7SK driven S3R3-shNS, all aptamer-shNS chimeras had slight inhibitory effects in our inhibitory assays, showing that the presence of aptamers alone can inhibit viral replication (Figure 3.4). Surprisingly, the inhibitory capabilities of aptamer-sh1498 and aptamer-shNS chimeras were quite similar. The Dp6.12-shRNA chimeras provided nearly identical inhibitory levels of HIV replication whereas the PR10.9-8N-shRNA and S3R3-shRNA chimeras were all slightly more effective when the aptamers were incorporated within sh1498 (Figure 3.4). Large differences in inhibitory effects

were observed between RNApt16-shNS and RNApt16-sh1498 but this is likely due to the cytotoxic effects of RNApt16-sh1498, as described previously. The most striking and unexpected feature from these assays is the absence of cytotoxicity in our competitive growth assays from U6 or 7SK driven PR10.9-8N-sh1498, Dp6.12-sh1498 and S3R3-sh1498 as we previously showed that sh1498 expressed alone from these same promoters causes high toxicity[5]. Possible mechanisms of shRNA-mediated cytotoxicity include off-target binding to cellular RNAs[35], saturation of the RNAi pathway[36,37]and eliciting the innate immune system through activation of host RNA sensors[38-41]. The alternative shRNA loop sequences in the aptamer-sh1498 chimeras may lead to differences in Dicer cleavage patterns and consequently generating varying guide and/or passenger strands species with different cytotoxic potential. It is also possible that the aptamer-sh1498 chimeras are not efficiently loaded into RISC after Dicer cleavage or that their cleavage by Dicer is reduced. Both of these scenarios would lead to a decrease in the total guide and passenger strands being produced. Finally, gene expression levels may be lower in the case of aptamer-sh1498 chimeras compared to the canonical sh1498. While PR10.9-8N-sh1498, Dp6.12-sh1498 and S3R3-sh1498 were not cytotoxic when expressed from the U6 or 7SK promoters, U6 driven RNApt16-sh1498 was associated with severe growth defects in our competitive growth assays (Figure 3.5A and D). Additionally, U6 driven RNApt16-shNS showed growth defects in our first competitive growth assay (Figure 3.5A) but not in our second competitive growth assay (Figure 3.5D). This assay will be repeated to confirm whether cytotoxicity is elicited from this construct. The aptamer RNApt16 is an extremely small aptamer measuring 16 nucleotides in length and possesses a three-dimensional structure similar to the canonical shRNA loop[19] (Figure 3.3). It is therefore likely processed quite similarly to the canonical sh1498 which would explain why cells expressed either U6 driven RNApt16-sh1498 or sh1498 alone exhibit similar growth

defects in our assays. In contrast, PR10.9-8N, Dp6.12, and S3R3 are all large aptamers which, when incorporated as the loop of sh1498, could have an impact on the processing and/or gene expression of the chimeric molecules compared to the canonical sh1498. Another explanation for the absence of toxicity could be that aptamer-shRNA chimeras involving large aptamers are too large to pass through the exportin-5 dependent shRNA nuclear export pathway and are retained in the nucleus.

To clarify some of these possibilities, we tested whether different levels of sh1498 guide strands were present in cells expressing the S3R3-sh1498 chimera compared to cells expressing canonical sh1498 by measuring the relative abundance of sh1498 guide strands by Northern blot. Our quantification showed that the sh1498 guide strands were more abundant in cells expressing the canonical sh1498 than when cells expressed the S3R3-sh1498 chimera. It is therefore likely that the guide strands themselves possess the same cytotoxic potential regardless of if they originate from the canonical sh1498 or the S3R3-sh1498 chimera but that the higher amount present in cells expressing the canonical sh1498 is responsible for at least some of the growth defects in these cells. However, this experiment cannot differentiate between a lower total gene expression level of S3R3-sh1498 or whether the chimeric molecule is being processed in a suboptimal manner after transcription. In both cases, the total amount of sh1498 guide strands generated from S3R3-sh1498 expression would be lower than during expression of canonical sh1498. Transcription rates are largely a consequence of the promoter driving expression and we have shown similar gene expression levels when different shRNA sequences are expressed[5]. We therefore hypothesize that the processing of S3R3-sh1498 to generate sh1498 guide strands is less efficient than the processing of the canonical sh1498 which results in observable differences in the total amount of guide strands in the Northern blot. Another limitation of our experiment is that it cannot elucidate

whether the guide and passenger strand species are identical after Dicer cleavage of S3R3-sh1498 and canonical sh1498. RNA sequencing experiments could address this issue and evaluate whether the alternative loop of aptamer-sh1498 chimeras influences Dicer cleavage patterns to generate RNA species. Evaluating why the guide strand of sh1498 is cytotoxic is another issue that should be addressed by testing for the possibility of induced innate immunity pathways, saturation of the RNAi pathway and off-target effects generated by the molecule.

This study shows the importance of investigating alternative shRNA molecular designs to optimize their use as antivirals in gene therapy. While it is important to maximize the antiviral effects, reducing cytotoxicity is equally important. Our results show that incorporating large aptamers as the loop of an shRNA effectively prevents the cytotoxicity observed when the shRNA is expressed in its canonical form. It will be beneficial to apply this molecular design to other shRNAs in future studies to identify an aptamer-shRNA chimera which can strongly inhibit viral replication with no cytotoxic effects.

3.6 Materials and methods.

3.6.1 Cell culture

HEK293T cells were grown in Dulbecco's modified Eagle's medium (HyClone, Logan, UT) supplemented with 10% fetal bovine serum (FBS, HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies, Carlsbad, CA, USA). SupT1 cells were grown in Roswell Park Memorial Institute Medium 1640 (HyClone), supplemented with 10% heat-inactivated FBS (HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies).

3.6.2 Vector construction

The DNA inserts coding for the aptamers with different flanking secondary structures as well as those coding for aptamer-shRNA chimeras were generated either by annealing complementary oligonucleotides or were synthesized by overlapping PCR as previously described[4,10,24]. The primers used for each construct are shown in Tables 3.1 and 3.2. These DNA inserts were then digested with BbsI and ligated into BbsI-digested psiRNA-7SKGFP::Zeo or into BbsI-digested psiRNA-U6GFP::Zeo. The constructs PR10.1-8E with flanking hairpins, PR10.1-8A with flanking hairpins and PR10.9-8N with flanking hairpins were ligated into EcoRI and ApaI-digested RNApt16_hp-psiRNA-7SKGFP::Zeo. To include the antiviral genes into lentiviral vectors, the expression cassettes of aptamer-shRNA chimeras were amplified using forward primer 5′-TATGCGGCCGCAGGGATTTTGGTCATGTTCTTAATCGATACTA-3′ and reverse primer 5′-GTAACGCCTGCAGGTTAATTAAGTCTAGAAGCTTTTCCAA-3′ to include the restriction sites NotI and XbaI. The synthesized DNA fragments were then digested with NotI and XbaI and ligated into NotI and XbaI-digested lentiviral transfer vector HIV-7-EGFP (donated by Dr. J. Rossi)[33].

Table 3.1 Primers used to generate constructs by annealing complementary oligonucleotides.

Primer name	Primer sequence 5'-3'
RNApt16_F	ACCTCCCCCGGCAAGGAGGGGTTT
RNApt16_R	CAAAAAACCCCTCCTTGCCGGGGG
RNApt16_hp_F	ACCTCGTGCTCGCTTCGGCAGCACATATACATGAATTC CCCCGGCAAGGAGGGGGGGCCCAGAGCGGACTTCGGT CCGCTTT
RNApt16_hp_R	CAAAAAAGCGGACCGAAGTCCGCTCTGGGCCCCCCCT CCTTGCCGGGGGAATTCATGTATATGTGCTGCCGAAGCGAGC ACG
RNApt16_HDV_F	ACCTCCCCCGGCAAGGAGGGGGGCCGGCATGGTCCCAGCCTC CTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAAT GGGACTTT

RNApt16_HDV_R	CAAAAAAGTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGCCCCCCTCCTTGCCGGGGG
PR10.1-8E_F	ACCTCCTTAAGTGTAACTTCTCATAACATCCGGAGGCTTTTACTTCCGGGGACCTTTT
PR10.1-8E_R	CAAAAAAAGGTCCCCGGAAGTAAAAGCCTCCGGATGTTATGAGAAGTTACACTTAAGG
PR10.1-8E_hp_F	AATTCCTTAAGTGTAACTTCTCATAACATCCGGAGGCTTTTACTTCCGGGGACCTGGGCC
PR10.1-8E_hp_R	CAGGTCCCCGGAAGTAAAAGCCTCCGGATGTTATGAGAAGTTACACTTAAGG
PR10.1-8A_F	ACCTCCTTAAGTGTAACTTCTCGTAATTCCCAAGGCTTTTACCTCGGGGTCCTTTT
PR10.1-8A_R	CAAAAAAAGGACCCCGAGGTAAAAGCCTTGGGAATTACGAGAAGTTACACTTAAGG
PR10.1-8A_hp_F	AATTCCTTAAGTGTAACTTCTCGTAATTCCCAAGGCTTTTACCTCGGGGTCCTGGGCC
PR10.1-8A_hp_R	CAGGACCCCGAGGTAAAAGCCTTGGGAATTACGAGAAGTTACACTTAAGG
PR10.9-8N_F	ACCTCTTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTTTT
PR10.9-8N_R	CAAAAAAACCAGGGCGAGGTCTCTGAACTCGAAGCCGTTATCTTACCTTAGGTCAAG
PR10.9-8N_hp_F	AATTCTTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTGGGCC
PR10.9-8N_hp_R	CACCAGGGCGAGGTCTCTGAACTCGAAGCCGTTATCTTACCTTAGGTCAAG
RNApt16-sh1498_F	ACCTCGCAGGAACTACTAGTACCCTTCCCCGGCAAGGAGGGGAAGGGTACTAGTAGTTCCTGCTTT
RNApt16-sh1498_R	CAAAAAAGCAGGAACTACTAGTACCCTTCCCCTCCTTGCCGGGGAAGGGTACTAGTAGTTCCTGCG
RNApt16-shNS_F	ACCTCGTACCGCACGTCATTCGTATTCCCCGGCAAGGAGGGGGATACGAATGACGTGCGGTACTTT
RNApt16-shNS_R	CAAAAAAGTACCGCACGTCATTCGTATCCCCCTCCTTGCCGGGGAATACGAATGACGTGCGGTACG

PR10.9-8N-sh1498_F	ACCTCGCAGGAACTACTAGTACCCTTTTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTAAGGGTACTAGTAGTTCCTGCTTT
PR10.9-8N-sh1498_R	CAAAAAAGCAGGAACTACTAGTACCCTTACCAGGGCGAGGTCTCTGAACTCGAAGCCGTTATCTTACCTTAGGTCAAAAGGGTACTAGTAGTTCCTGCG
PR10.9-8N-shNS_F	ACCTCGTACCGCACGTCATTCGTATTTTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTGATACGAATGACGTGCGGTACTTT
PR10.9-8N-shNS_R	CAAAAAAGTACCGCACGTCATTCGTATCACCAGGGCGAGGTCTCTGAACTCGAAGCCGTTATCTTACCTTAGGTCAAAATACGAATGACGTGCGGTACG

Table 3.2 Primers used to synthesize constructs by overlapping PCR.

Primer name	Primer sequence 5'-3'
RNApt16_HH_HDV full insert: GCGGCCGCGCCGGGGCTGATGAGTCCGTGAGGACGAAACGGTACCCGGTACCGTCCCCCGGCAAGGAGGGGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTT	
RNApt16_HH_HDV_F	GCGGCCGCGCCGGGGCTGATGAGTCCGTGAGGACG AAACGGTACCCGGTACCGTCCCCCGGCAAGGAGGGG GGCCGGCATG
RNApt16_HH_HDV_R	AAAGTCCCATTCGCCATGCCGAAGCATGTTGCCC AGCCGGCGCCAGCGAGGAGGCTGGGACCATGCC GGCCCCCCTCCTTG
RNApt16_HH_HDV_Flanking Forward	GAGTTCACGGAAGACCGACCTCGCGGCCGCGCCGGG
RNApt16_HH_HDV_Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTCCCATTCGCCA
PR10.1-8E_HDV full insert: CTTAAGTGTAACTTCTCATAACATCCGGAGGCTTTTACTTCCGGGGACCTGG	

CCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTT	
PR10.1-8E_HDV_F	CTTAAGTGTAACTTCTCATAACATCCGGAGGCTTTTACTTCCGGGGACCTGGCCGGCATGGTCCCAGCCT
PR10.1-8E_HDV_F	AAAGTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGC
PR10.1-8E_HDV_Flanking Forward	GAGTTCACGGAAGACCGACCTCCTTAAGTGTAACTT
PR10.1-8E_HDV_Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTCCCATTCGCCA
PR10.1-8A_HDV full insert: CTTAAGTGTAACTTCTCGTAATTCCCAAGGCTTTTACCTCGGGGTCCTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTT	
PR10.1-8A_HDV_F	CTTAAGTGTAACTTCTCGTAATTCCCAAGGCTTTTACCTCGGGGTCCTGGCCGGCATGGTCCCAGCCTCC
PR10.1-8A_HDV_R	AAAGTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGC
PR10.1-8A_HDV_Flanking Forward	GAGTTCACGGAAGACCGACCTCCTTAAGTGTAACTT
PR10.1-8A_HDV_Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTCCCATTCGCCA
PR10.9-8N_HDV full insert: TTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTT	
PR10.9-8N_HDV_F	TTGACCTAAGGTAAGATAACGGCTTCGAGTTCAGAGACCTCGCCCTGGTGGCCGGCATGGTCCCAGCCTC

PR10.9-8N_HDV_R	AAAGTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGC
PR10.9-8N_HDV_Flanking Forward	GAGTTCACGGAAGACCGACCTCTTGACCTAAGGTAA
PR10.9-8N_HDV_Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTCCCATTCGCCA
DP6.12 full insert: GGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGTTT	
DP6.12_F	GGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTAC
DP6.12_R	AAACGTATGACTTCAGGCTAGCACTAGAGCCAAGTTTCTGTAAGCAAGCCGCTGCGAAGTAGGCGAGTGGGGTTTTTTT
DP6.12_Fanking Forward	GAGTTCACGGAAGACCGACCTCGGACAGCAAGCGTA
DP6.12_Fanking Reverse	TGATGCTATGAAGACTCCAAAAAAAACGTATGACTTCAG
DP6.12_hp full insert: GTGCTCGCTTCGGCAGCACATATACATGAATTCGGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGGGGCCCAGAGCGGACTTCGGTCCGCTTT	
DP6.12_hp_F1	GTGCTCGCTTCGGCAGCACATATACATGAATTCGGACAGCAAGCGTACATCTAAGACGGC
DP6.12_hp_F2	GAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTG
DP6.12_hp_R1	GGTTTTTTTTGCCGCAACTCGAACAACCGCTTCGACTTAAGCCGTCTTAGATGTACGCTT

DP6.12_hp_R2	AAAGCGGACCGAAGTCCGCTCTGGGCCCCGTATGACTTCAGGCTAGCACTAGAGCCAAGTTTCTGTAAGCAAGC
DP6.12_hp Flanking Forward	GAGTTCACGGAAGACCGACCTCGTGCTCGCTTCGGC
DP6.12_hp Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGCGGACCGAAGTC
DP6.12_HDV full insert: **GGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGGGCCGGCATGGTCCCAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTT**	
DP6.12_HDV_F1	GGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAA
DP6.12_HDV_F2	TTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGGGCCGGCATGGTCCCAGCCTC
DP6.12_HDV_R1	AGCCAAGTTTCTGTAAGCAAGCCGCTGCGAAGTAGGCGAGTGGGGTTTTTTTTGCCGCAACTCGA
DP6.12_HDV_R2	AAAGTCCCATTCGCCATGCCGAAGCATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGC
DP6.12_HDV Flanking Forward	GAGTTCACGGAAGACCGACCTCGGACAGCAAGCGTA
DP6.12_HDV Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTCCCATTCGCCA
DP6.12-sh1498 full insert: **GCAGGAACTACTAGTACCCTTGGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGAAGGGTACTAGTAGTTCCTGCTTT**	
DP6.12-sh1498_F	GCAGGAACTACTAGTACCCTTGGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTA
DP6.12-sh1498_R	AAAGCAGGAACTACTAGTACCCTTCGTATGACTTCAGGCTAGCACTAGAGCCAAGTTTCTGTAAGCAAGCCGCTGCGAAGTAGGCGAGTGGGGTTTTTTT

DP6.12-sh1498 Flanking Forward	GAGTTCACGGAAGACCGACCTCGCAGGAACTACTAG
DP6.12-sh1498 Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGCAGGAACTACTA
DP6.12-shNS full insert: GTACCGCACGTCATTCGTATTGGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTACTTCGCAGCGGCTTGCTTACAGAAACTTGGCTCTAGTGCTAGCCTGAAGTCATACGGATACGAATGACGTGCGGTACTTT	
DP6.12-shNS_F	GTACCGCACGTCATTCGTATTGGACAGCAAGCGTACATCTAAGACGGCTTAAGTCGAAGCGGTTGTTCGAGTTGCGGCAAAAAAAACCCCACTCGCCTA
DP6.12-shNS_R	AAAGTACCGCACGTCATTCGTATCCGTATGACTTCAGGCTAGCACTAGAGCCAAGTTTCTGTAAGCAAGCCGCTGCGAAGTAGGCGAGTGGGGTTTTTT
DP6.12-shNS Flanking Forward	GAGTTCACGGAAGACCGACCTCGTACCGCACGTCAT
DP6.12-shNS Flanking Reverse	TGATGCTATGAAGACTCCAAAAAAAAGTACCGCACGTCA

3.6.3 RT assay

HEK293T cells were seeded in 96-well plates at a concentration of 2.5 x 10^5 - 3 x 10^5 cells/mL 24 h prior to co-transfections in a volume of 150 µl. Co-transfections were established with 50 ng HIV-1 molecular clone pNL4-3 along with 250 ng of aptamer-expressing plasmid using TransIT-LT1 (Mirus Bio, Madison, WI, USA) according to the manufacturer's protocol. Supernatants were harvested 48 h after transfection for measurement of RT activity. HIV RT activity in supernatants was measured using a radioactive RT assay as previously described[14,42]. Briefly, 2.5 or 5 µL of supernatant was incubated for 2 h at 37°C with 25 or 50 µL of RT cocktail containing a poly(A)

template (Roche, Basel, Switzerland), an oligo(dT) primer (Life Technologies) and [^{32}P] deoxythymidine triphosphate (dTTP; 3,000 Ci/mmol; Perkin Elmer, Waltham, MA, USA). The poly dT RT product was then detected by spotting 5 μL of the reaction mixture onto diethylaminoethyl (DEAE) filter mats (Perkin Elmer) or a positively charged nylon membrane (Hybond-N$^+$, Amersham Biosciences, Little Chalfont, UK) and unincorporated [^{32}P] dTTP were removed with five washes using 2x saline sodium citrate (SSC) buffer (20× SSC buffer: 3 M NaCl and 0.3 M sodium citrate) along with two washes using 95% ethanol. Counts per minute (cpm) were measured on a microplate scintillation counter (MicroBeta TriLux; Perkin Elmer). The amount of HIV RT enzyme in the supernatants is proportional to the cpm readout.

3.6.4 Lentivirus production

HEK293T cells were seeded at 2.75 x 10^5 cells/mL 24 h prior to co-transfections in T75 flasks in a volume of 20 mL. Co-transfections were performed with 10 μg HIV-7-EGFP plasmid, 3.4 μg plasmid expressing vesicular stomatitis virus G protein (from Dr. J. Rossi) and 10 or 50 μg packaging plasmid psPAX2 (Addgene, number 12260) using PEI (3 μL/μg of DNA) (Polysciences). Supernatants were harvested 48 h after transfection, where lentiviral particles were isolated and concentrated using Lenti-X (Clontech) according to the manufacturer's protocol.

3.6.5 SupT1 T cell transduction, infection, and competitive growth

A serial dilution of isolated lentiviral particles was used to transduce SupT1 cells, and the percentage of GFP-positive cells was then measured to determine lentivirus titers. These lentiviruses were then transduced at an MOI of 1 into 5 ml SupT1 cell cultures containing 6 x 10^5 – 1 x 10^6 cells, using 8 μg/mL Polybrene (Sigma-Aldrich, St. Louis, MO, USA). Cells were sorted 72 h after transduction for GFP expression with a FACSAria Fusion cell sorter (BD Biosciences, Franklin Lakes, NJ, USA). The gating is shown in Figures 3.7-3.10. One set of sorted cells were

used for HIV infection and plated in 96-well round-bottom plates at 2 x 10^4 cells/well in a volume of 100 µL. Twenty-four hours after sorting, the cells were infected with HIV-1 NL4-3 (1,750 cpm/mL, determined using the HIV RT assay) and supernatants were harvested on various days post infection to measure RT activity and establish infection kinetics. A second set of sorted cells were used to establish a competitive growth and plated in 96-well round-bottom plates at 2 x 10^4 cells/well in a volume of 100 µL. These cells were immediately mixed with an equal number of untransduced SupT1 cells that had been passed through the flow cytometer without sorting, bringing the total volume in each well to 200 µL with 4 x 10^4 cells/well. Half of the cell populations were collected at various days post mixing to measure the percentage of GFP-positive cells using a LSRFortessa flow cytometer (BD Biosciences).

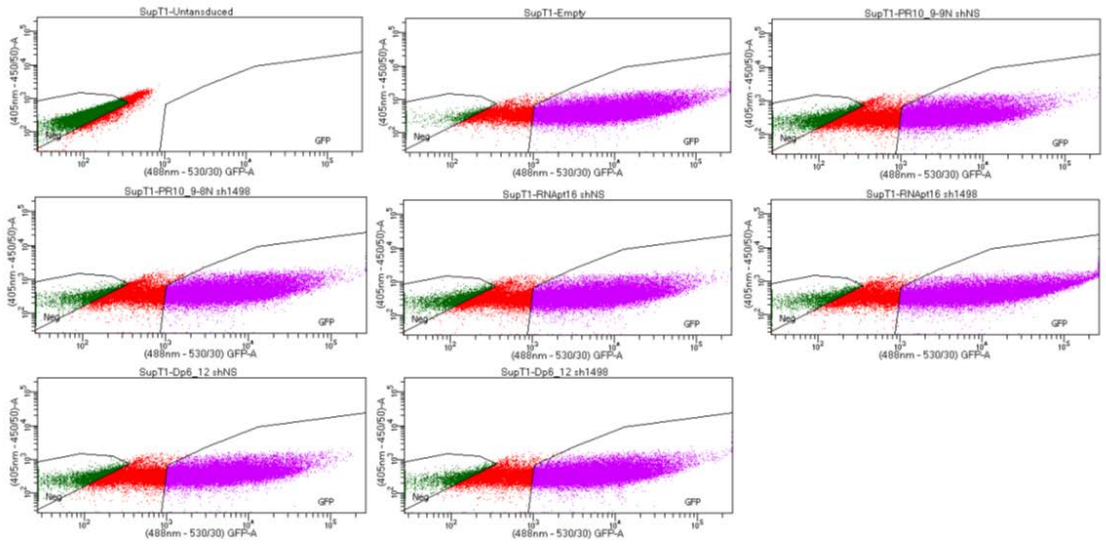

Figure 3.7 FACS plots of cell cultures expressing U6 driven aptamer-shRNA chimeras. LVs encoding the selected aptamer-shRNA chimera genes were used to transduce SupT1. Transduced cell cultures were sorted by GFP expression using FACS. Populations were previously gated for single lymphocytes. Purple populations in the GFP gate were collected and used to establish the infection and competitive growth shown in Figures 3.4 and 3.5.

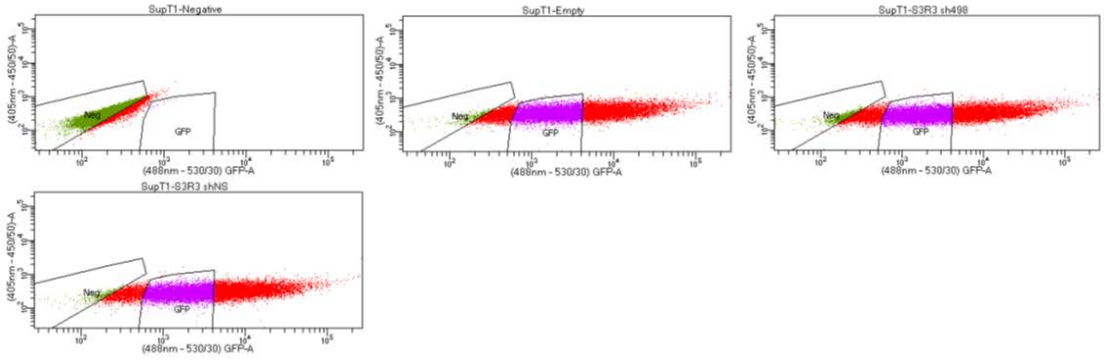

Figure 3.8 FACS plots of cell cultures expressing 7SK driven aptamer-shRNA chimeras. LVs encoding the selected aptamer-shRNA chimera genes were used to transduce SupT1. Transduced cell cultures were sorted by GFP expression using FACS. Populations were previously gated for single lymphocytes. Purple populations in the GFP gate were collected and used to establish the infection and competitive growth shown in Figures 3.4 and 3.5.

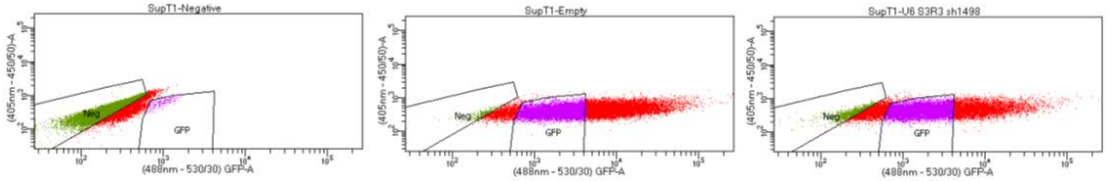

Figure 3.9 FACS plots of cell cultures expressing U6 driven aptamer-shRNA chimeras. LVs encoding the selected aptamer-shRNA chimera genes were used to transduce SupT1. Transduced cell cultures were sorted by GFP expression using FACS. Populations were previously gated for single lymphocytes. Purple populations in the GFP gate were collected and used to establish the infection and competitive growth shown in Figures 3.4 and 3.5.

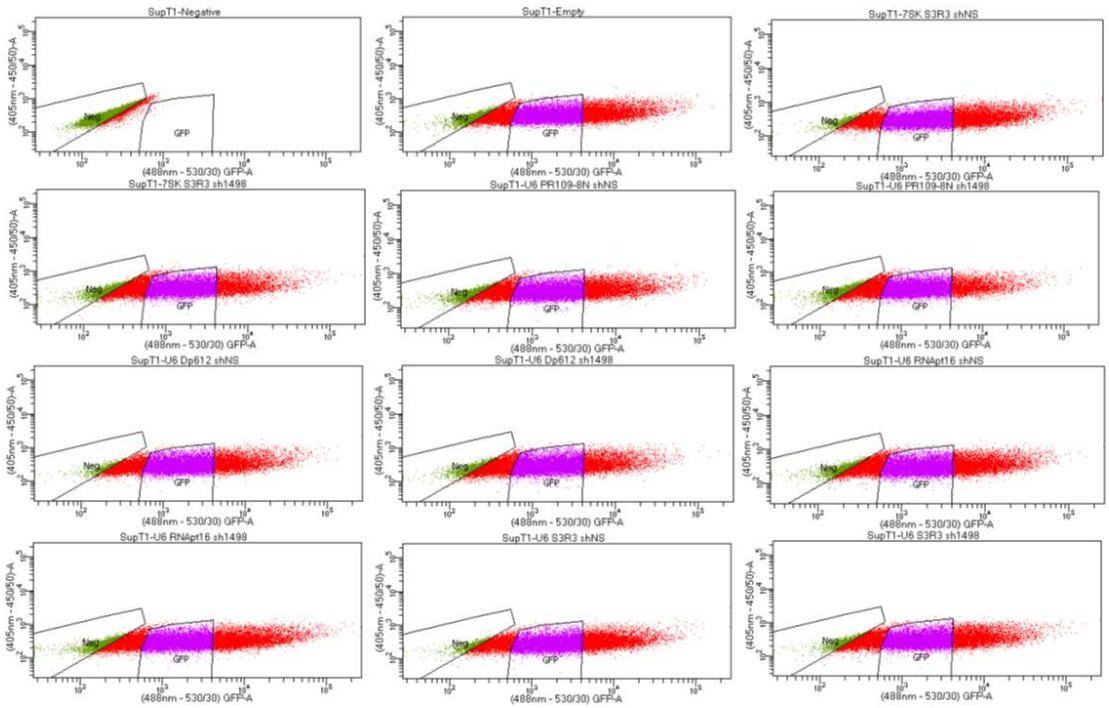

Figure 3.10 FACS plots of cell cultures expressing U6 and 7SK driven aptamer-shRNA chimeras. LVs encoding the selected aptamer-shRNA chimera genes were used to transduce SupT1. Transduced cell cultures were sorted by GFP expression using FACS. Populations were previously gated for single lymphocytes. Purple populations in the GFP gate were collected and used to establish the infection and competitive growth shown in Figures 3.4 and 3.5.

3.6.6 Northern blot

HEK293T cells were seeded at 3.5 x 10^6 cells/dish 24h prior to transfections in 10 mL tissue cultures dishes. Transfections were performed with 5 μg of plasmid using TransIT-LT1 (Mirus Bio, Madison, WI, USA) according to the manufacturer's protocol. 48h post-transfection, cell lysates were collected using TRIzol reagent (Life Technologies) and total RNA was recovered by phenol chloroform extraction followed by cleanup with a miRNeasy mini kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. A total of 15 μg of the extracted RNA was mixed with an equal volume of 2× gel loading buffer and resolved in a 15% polyacrylamide-urea gel as previously described[24]. The RNA was transferred from the gel to a positively charged nylon membrane (Hybond-N^+, Amersham Biosciences, Little Chalfont, UK) using a semi-dry

electroblotter (25 min, 4°C, 20 V). The transferred RNA was chemically crosslinked with 1-Ethyl-3-3(3-dimethylaminopropyl) carbodiimide (Sigma-Aldrich, St. Louis, MO, USA) and 1-Methylimidazole (Thermo-Fisher, Waltham, MA, USA) and incubating at 60°C for 80 min. Prehybridization buffer consisting of 6× SSC, 2× Denhardt's solution, and 0.1% SDS was used to prehybridize the membrane. Identical buffer composition was used to hybridize the ^{32}P-labeled RNA probes. Probe hybridization was done overnight at 37°C and the membrane was subsequently washed for 15 min in wash buffer #1 (2× SSC), 15 min in wash buffer #2 (1× SSC), and 15 min in wash buffer #3 (0.1× SSC), all at 37°C. Finally, radiographic films were exposed to visualize the Northern bands.

3.6.7 Probe labeling

^{32}P-labeled RNA probes were generated by *in vitro* transcription using the HiScribe T7 kit (NEB) according to the manufacturer's instructions with annealed template DNA and [α-32P] cytidine triphosphate (CTP, 800 Ci/mmol; Perkin Elmer). To eliminate the template DNA after in vitro transcription, the reaction mixture was treated with DNase I (New England Biolabs, Ipswich, MA, USA). The radioactive RNA probes were then purified with ProbeQuant G50 Micro Columns (GE Healthcare, Little Chalfont, UK) to remove unincorporated nt. The probe sequences with the T7 promoter sequence underlined were as follows:

sh1498-antisense: 5′-GGGTACTAGTAGTTCCTGCCTATAGTGAGTCGTATTAATTTC-3′

3.7 References.

1. Laskey, S.B. & Siliciano, R.F. A mechanistic theory to explain the efficacy of antiretroviral therapy. *Nat Rev Microbiol* **12**, 772-780 (2014).
2. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).
3. Scarborough, R.J. & Gatignol, A. RNA Interference Therapies for an HIV-1 Functional Cure. *Viruses* **10**, 8 (2018).
4. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
5. Goguen, R.P., *et al.* Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK, and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021).
6. Lange, M.J., Nguyen, P.D.M., Callaway, M.K., Johnson, M.C. & Burke, D.H. RNA-protein interactions govern antiviral specificity and encapsidation of broad spectrum anti-HIV reverse transcriptase aptamers. *Nucleic Acids Res* **45**, 6087-6097 (2017).
7. Pang, K.M., Castanotto, D., Li, H., Scherer, L. & Rossi, J.J. Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. *Nucleic Acids Res* **46**, e6 (2018).
8. Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. *Mol Ther Nucleic Acids* **18**, 815-830 (2019).
9. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**, 421-433 (2015).
10. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).
11. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol Ther* **14**, 883-892 (2006).
12. Centlivre, M., *et al.* Preclinical in vivo evaluation of the safety of a multi-shRNA-based gene therapy against HIV-1. *Mol Ther Nucleic Acids* **2**, e120 (2013).
13. McIntyre, G.J., *et al.* 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* **6**, 55 (2009).
14. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. & Gatignol, A. Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp* **115**, e54486 (2016).
15. Lee, N.S., *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20**, 500-505 (2002).
16. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).

17. Duclair, S., Gautam, A., Ellington, A. & Prasad, V.R. High-affinity RNA Aptamers Against the HIV-1 Protease Inhibit Both In Vitro Protease Activity and Late Events of Viral Replication. *Mol Ther Nucleic Acids* **4**, e228 (2015).
18. Ramalingam, D., *et al.* RNA aptamers directed to human immunodeficiency virus type 1 Gag polyprotein bind to the matrix and nucleocapsid domains and inhibit virus production. *J Virol* **85**, 305-314 (2011).
19. Sanchez-Luque, F.J., Stich, M., Manrubia, S., Briones, C. & Berzal-Herranz, A. Efficient HIV-1 inhibition by a 16 nt-long RNA aptamer designed by combining in vitro selection and in silico optimisation strategies. *Sci Rep* **4**, 6242 (2014).
20. Gulick, R.M., *et al.* Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med* **337**, 734-739 (1997).
21. Hammer, S.M., *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* **337**, 725-733 (1997).
22. Chung, J., *et al.* Optimized lentiviral vectors for HIV gene therapy: multiplexed expression of small RNAs and inclusion of MGMT(P140K) drug resistance gene. *Mol Ther* **22**, 952-963 (2014).
23. Lambeth, L.S. & Smith, C.A. Short hairpin RNA-mediated gene silencing. *Methods Mol Biol* **942**, 205-232 (2013).
24. Goguen, R.P., Gatignol, A. & Scarborough, R.J. Cloning and Detection of Aptamer-Ribozyme Conjugations. *Methods Mol Biol* **2167**, 253-267 (2021).
25. Ma, H., *et al.* Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids* **3**, e161 (2014).
26. Paul, C.P., *et al.* Localized expression of small RNA inhibitors in human cells. *Mol Ther* **7**, 237-247 (2003).
27. Herrera-Carrillo, E., Gao, Z. & Berkhout, B. Influence of a 3' Terminal Ribozyme on AgoshRNA Biogenesis and Activity. *Mol Ther Nucleic Acids* **16**, 452-462 (2019).
28. Gao, Z., Herrera-Carrillo, E. & Berkhout, B. Improvement of the CRISPR-Cpf1 system with ribozyme-processed crRNA. *RNA Biol* **15**, 1458-1467 (2018).
29. Benedict, C.M., Pan, W., Loy, S.E. & Clawson, G.A. Triple ribozyme-mediated down-regulation of the retinoblastoma gene. *Carcinogenesis* **19**, 1223-1230 (1998).
30. An, D.S., *et al.* Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors. *Mol Ther* **14**, 494-504 (2006).
31. Shimizu, S., *et al.* Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**, 8 (2009).
32. Makinen, P.I., *et al.* Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* **8**, 433-441 (2006).
33. Yam, P.Y., *et al.* Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol Ther* **5**, 479-484 (2002).
34. Wolstein, O., *et al.* Preclinical safety and efficacy of an anti-HIV-1 lentiviral vector containing a short hairpin RNA to CCR5 and the C46 fusion inhibitor. *Mol Ther Methods Clin Dev* **1**, 11 (2014).

35. Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci U S A* **112**, E4007-4016 (2015).
36. Yi, R., Doehle, B.P., Qin, Y., Macara, I.G. & Cullen, B.R. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA* **11**, 220-226 (2005).
37. Grimm, D., *et al.* Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *J Clin Invest* **120**, 3106-3119 (2010).
38. Fish, R.J. & Kruithof, E.K. Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. *BMC Mol Biol* **5**, 9 (2004).
39. Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**, 263-264 (2003).
40. Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H. & Williams, B.R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**, 834-839 (2003).
41. Singhania, R., Khairuddin, N., Clarke, D. & McMillan, N.A. RNA interference for the treatment of papillomavirus disease. *Open Virol J* **6**, 204-215 (2012).
42. Scarborough, R.J., Levesque, M.V., Perreault, J.P. & Gatignol, A. Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* **1103**, 31-43 (2014).

Chapter IV

Anti-HIV aptamer-shRNA chimeras expressed from the U6 and 7SK promoters strongly inhibit HIV replication in lymphocytic cells

This chapter is adapted from the following manuscript: Goguen, R.P., Malard, C.M.G., Chen, M.J., Intrevado, I., Daher, A., Rossi, J. J., Gatignol, A., Scarborough, R.J. In Preparation.

Note: Some shRNA names have been changed from the figures and text because of pending IP discussions; Supplementary data are not included because of IP considerations.

4.1 Preface.

The body of work shown in chapter III focused on optimizing our anti-HIV gene therapy by exploring various molecular designs. The purpose of this was to eliminate the cytotoxicity observed in chapter II caused by the expression of shRNAs from the U6 and 7SK promoters. The incorporation of aptamer within the terminal loop of shRNAs to generate aptamer-shRNA chimeras successfully lead to a method to avoid shRNA-mediated cytotoxicity. Similar to the lower levels of shRNA expression from the H1 promoter observed in chapter II, the circulating levels of shRNA guide strand were shown to be the factor influencing the cytotoxic potential of the anti-HIV gene therapy. Specifically, expression of aptamer-shRNA chimeras led to lower amounts of guide strands compared to canonical shRNAs and this eliminates the cytotoxic effects elicited by the expression of canonical shRNAs.

Although a method to avoid shRNA-mediated cytotoxicity was identified in chapter III, none of the characterized aptamer-shRNA chimeras could act as potent inhibitors of HIV. Therefore, this chapter focuses on identifying an aptamer-shRNA chimera that can potently inhibit HIV without eliciting cytotoxicity. Furthermore, only a single anti-HIV shRNA was evaluated in chapters II and III. As multiple anti-HIV RNAs must be used during gene therapy to avoid the emergence of resistance, this chapter also focuses on identifying additional potent anti-HIV shRNAs for eventual use in combination gene therapy.

The results presented in this chapter allowed us to characterize additional anti-HIV shRNAs for eventual use in combination gene therapy. We expressed these shRNAs from the U6, 7SK and H1 promoters as was done in chapter II and observed once again that anti-HIV shRNAs were more potent when expressed from the U6 and 7SK promoters. However, cytotoxicity was not elicited in the case of one shRNA expressed from the U6 promoter, reinforcing our hypothesis that the

nucleotide identity of shRNAs influences whether expression from the U6, 7SK or H1 promoter leads to maximal shRNA activity without cytotoxicity. Additionally, the results from this chapter served to identify aptamer-shRNA chimeras that could potently inhibit HIV without cytotoxic effects. Overall, this work will help to establishing a combination gene therapy which avoids the emergence of viral resistance.

CONFLICT OF INTEREST:

The authors declare no competing interests.

ACKNOWLEDGEMENTS

We would like to thank Christian Young (Lady Davis Institute flow cytometry manager) for assistance in cell sorting. This work was supported by grant PJT-148704 awarded to AG and RJS from the Canadian Institutes of Health Research (CIHR). RPG is supported by a Doctoral fellowship from the Fonds de Recherche du Québec-Santé (FRQS). CMGM and MJC were supported by MSc scholarships from the FRQS.

4.2 Abstract.

Short hairpin RNAs (shRNAs) expressed in cells by gene therapy are effective inhibitors of HIV-1 replication but quickly lose their antiviral efficacy when used alone due to the development of viral resistance. The use of multiple shRNAs can be used to prevent the emergence of resistant virus. Although clinical trials have been conducted using combination gene therapy, optimization of anti-HIV RNAs is needed to identify better combinations. We compared the effects of several shRNAs identified in previous screens expressed from the H1 promoter and show that three with highly conserved target sites were superior in inhibiting HIV replication. We then compared the efficacy and toxicity of these three shRNAs when expressed from the stronger 7SK and U6 promoters. All three shRNAs strongly inhibited HIV replication when expressed from these two promoters on a lentiviral vector in lymphocytic cells but exhibited varying degrees of cytotoxicity observed by growth defect. This toxicity could be alleviated by replacing the loop sequences of the shRNAs with an anti-HIV aptamer. These results are important for the establishment of optimized anti-HIV gene therapies.

4.3 Introduction.

Infection by the human immunodeficiency virus type 1 (HIV) can be controlled by the administration of combination antiretroviral therapies (cART) to delay the progression to AIDS. However, this strategy is not curative which consequently results in a lifelong treatment[1]. As such, the need for continuous chronic administration of medication places financial pressures on patients and health care systems while also being associated to undesirable side effects. Gene therapy with antiviral molecules could provide a functional cure towards HIV infection where cART would no longer be needed[2,3]. Various antiviral RNAs have been developed to inhibit HIV replication by gene therapy including short hairpin RNAs (shRNAs), aptamers, ribozymes and U1 interference RNAs[4-8]. As is the case with cART, a combination of inhibitors will certainly be necessary to prevent the occurrence of resistance and clinical trials have been conducted using multiple anti-HIV RNAs[9-11].

RNA interference (RNAi) is an endogenous cellular post-transcriptional gene regulation pathway first discovered in *Caenorhabditis elegans*[12]. This pathway has since been identified in mammalian cells and has been implicated in regulating cell growth, metabolism and other biological processes[13,14]. RNAi has been used to target various mRNAs involved in numerous diseases, including HIV infection[4,5,15]. Such treatment of HIV infection by RNAi relies on the expression of anti-HIV shRNAs by gene therapy to disrupt the expression of viral genes or cellular genes utilized by the virus for its replication. Gene silencing from shRNAs occurs through their incorporation into the RNA-induced silencing complex (RISC) after processing into small interfering RNAs (siRNAs) by the RNase III endonuclease Dicer to remove the loop sequence[3]. The guide strand of the siRNA directs, by nucleotide complementarity, the Argonaute2 (Ago2) protein present in the RISC towards an RNA target which results in cleavage of the target[15]. Thus, anti-HIV shRNAs

are specifically designed so that the guide strand of the molecule is complementary to RNA species that are critical for the completion of the viral replication cycle.

Although shRNAs are able to potently inhibit HIV replication, the virus can develop resistance to any single shRNA. Mutations within the target site, resulting in nucleotide mismatches, as well as mutations outside of the target site, resulting in the formation of RNA secondary structures surrounding the target site, can abolish the antiviral activity of anti-HIV shRNAs[16-18]. It has therefore been proposed that the usage of multiple shRNAs targeting different conserved viral sites in combination will be required to avoid the occurrence of resistance[10,11]. This approach is similar to cART, which relies on the administration of at least three antiretroviral drugs with different viral targets to control viremia while avoiding the emergence of resistance[1,19,20].

In an effort to identify highly active anti-HIV shRNAs to use in combination, several studies have screened a number of shRNAs targeting conserved regions of different HIV-1 genomes for antiviral activity. In one study, a total of 170 HIV-1 genomes were examined to generate 86 different shRNAs targeting highly conserved sequences[10]. A total of 1 in 4 of the shRNAs were found to be effective inhibitors of HIV, targeting eight different viral genome regions. Another study also used sequence conservation as the primary criterion for target site identification and generated 96 different shRNAs against 22 distinct regions of the viral genome[21]. Of those shRNAs, 65 were found to be highly active inhibitors of HIV. A third study identified target sites by selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemical probing for RNA secondary structures and used the information to generate a set of 26 shRNAs[22]. Evaluation of inhibitory capabilities showed that 14 out of the 26 shRNAs were highly potent inhibitors of HIV.

Another class of antiviral functional RNAs are aptamers, which are generated by systematic evolution of ligands by exponential enrichment (SELEX)[23]. These molecules possess binding

specificities for protein or RNA targets which results in an inactivation of the target upon aptamer binding[2]. Binding affinity is a consequence of the specific three-dimensional structure that is adopted through intramolecular folding of these single stranded RNA aptamers. Recently, aptamers have been included within the loop of shRNAs in an attempt to express multiple anti-HIV molecules from a single RNA transcript[7]. As part of the biogenesis of shRNAs includes cleavage of the stem-loop structure by Dicer, separation of the aptamer from the shRNA will occur through Dicer cleavage following nuclear export of the aptamer-shRNA chimera. Careful design of such aptamer-shRNA chimeras which target different viral elements could avoid the occurrence of resistance in the same manner as employing multiple shRNAs in parallel.

Expression of anti-HIV shRNAs for gene therapy has primarily been done with the type 3 RNA polymerase (Pol) III promoters U6, 7SK and H1. Post-transcriptional modifications (5' cap and poly(A) tail) which could interfere with RNA functionality are not included in transcripts produced from these promoters and the transcriptional start site is generally well defined to allow for expression of shRNAs with reliable sequences[5,24,25]. Appropriate promoter selection is imperative to designing an effective anti-HIV gene therapy using shRNAs as the antiviral potency and potential to elicit cytotoxicity of the molecules has been reported to be affected by the promoter used to drive their expression[5]. Specifically, studies have reported that shRNAs may offer more potent gene silencing when expressed by the U6 promoter compared to the H1 promoter, but that expression by the U6 promoter is also associated to cytotoxic effects[26,27]. We have previously identified an shRNA (sh1498) targeting the *gag* region of HIV RNA, and evaluated its effects when expressed from the U6, 7SK and H1 promoters[4,5]. We found that sh1498 potency is maximized when its expression is driven by the U6 and 7SK promoters but that expression from these promoters can also lead to cytotoxicity. In contrast, one study has reported no cytotoxicity

associated to other shRNAs expressed from the U6 promoter[28]. Due to the inconsistency of the literature, it is unlikely that a single promoter is universally the prime choice for expression of all anti-HIV shRNAs. Rather, each shRNA should be investigated individually when expressed from the different type 3 RNA Pol III promoters to identify the optimal promoter in a case-by-case fashion.

In this study, we identified 3 highly active anti-HIV shRNAs and expressed these from the H1, 7SK and U6 promoters. Although antiviral effects were maximized when the shRNAs were expressed from the U6 and 7SK promoters, expression from these two promoters led to severe growth defects for two of the three shRNAs. We also included an aptamer within the loop of the 3 highly potent shRNAs to generate aptamer-shRNA chimeras and expressed these from the 7SK and U6 promoters. Interestingly, severe growth defects were not detected when any of the chimeric molecules were expressed from these two promoters, along with viral replication being strongly restricted by the aptamer-shRNA chimeras. These results suggests that both the expression method as well as the molecular design of shRNAs are important determinants of antiviral potency and potential to elicit cytotoxicity.

4.4 Results.

4.4.1 Selection of top shRNAs from literature screens

Target site conservation of anti-HIV shRNAs identified in previous large shRNA screens was evaluated using the Los Alamos National Laboratory (LANL) HIV Sequence Database[10,21,22] (Supplementary Table S1). Selection of the top shRNA candidates from these screens was determined according to target site conservation, potency in previous HIV inhibitory assays, and

inclusion in clinical trials and/or pre-clinical animal studies. An shRNA targeting the psi packaging sequence (sh688) and another targeting the pol gene (sh4749) were also selected[29,30]. Additionally, we identified highly conserved regions of the LTR and RRE to design two novel shRNAs, sh582 (LTR-specific) and sh7797 (RRE-specific). We also included an shRNA targeting the *tat/rev* region of HIV RNA previously tested in an anti-HIV combination gene therapy clinical trial[9], denoted as sh5983, along with a Gag-specific sh1498 that our lab has previously developed[4]. An alternative Gag-specific shRNA has been demonstrated in the literature to negatively impact T-cell growth[31]. We therefore selected this shRNA as a positive control relating to the evaluation of cytotoxic effects. Target site conservation of all selected shRNAs was determined using the LANL database across all HIV-1 strains, as well as across clade B and C strains and is shown in Supplementary Table S2.

4.4.2 Three shRNAs expressed from the H1 promoter are particularly effective at inhibiting HIV replication

A total of 23 shRNAs were selected after target site conservation analysis and were included within an H1 promoter-driven expression plasmid (psiRNA-H1GFP::Zeo) to evaluate their effects on HIV production. A negative control non-sense shRNA (shNS) that does not target HIV RNA was also expressed from the H1 promoter. Co-transfections were performed in HEK 293T cells using various doses of each shRNA expression plasmids along with the HIV-1 molecular clone pNL4-3. Relative HIV production was estimated by measuring the HIV RT activity present in cell supernatants harvested from the co-transfections. All data were normalized to the RT activity measured in co-transfections of pNL4-3 with the empty vector (psiRNA-H1GFP::Zeo). All shRNAs showed dose-dependent inhibitory effects, except for shG5 and sh7797, which were not

able to inhibit HIV production (Figure 4.1). The most effective inhibitors from this screen were sh516, shL1, sh1498, shP2, shS3, shP28, shT140, and sh5983.

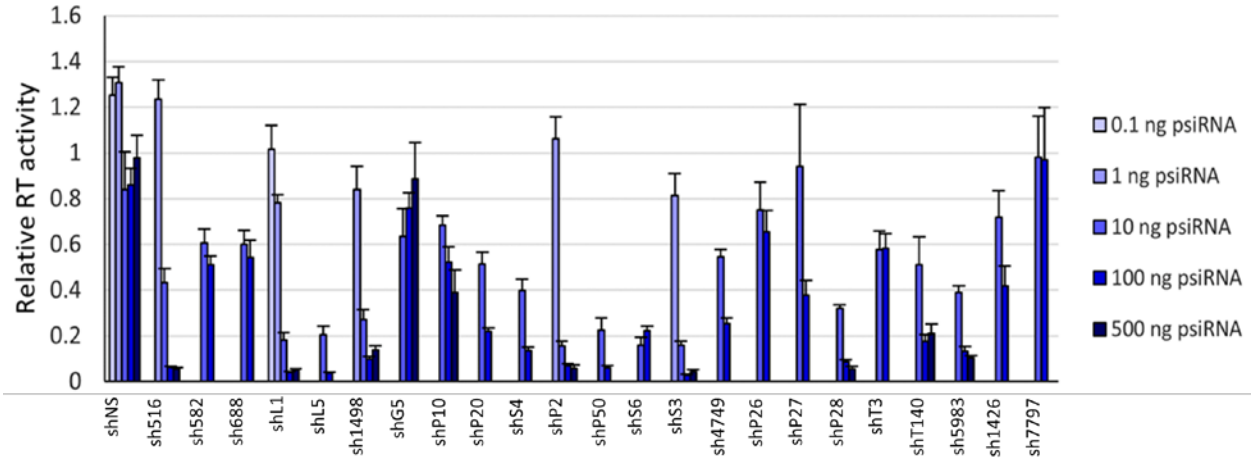

Figure 4.1 shRNA inhibition of HIV production from co-transfection of HEK293T cells. HEK 293T cells were co-transfected with the HIV-1 molecular clone pNL4-3 along with varying amounts of different shRNA expression plasmids. Supernatants were collected 48 h post-transfection and virus production was estimated by measuring HIV RT activity. Data are expressed as a percentage of RT activity in cells cotransfected with the empty expression plasmid. Each data point represents the mean +/- standard error mean (SEM) from two independent experiments with two replicates (n = 4).

Based on the screening of the inhibitory capabilities of shRNAs within the HIV production assay, we selected 10 shRNAs to characterize in a human T lymphocytic cell line (SupT1 cell line). Although shP10 and shG5 were not strong inhibitors in our primary screen, we included these molecules in our selection as these shRNAs were selected for evaluation in preclinical mouse studies along with shP28 and shT140[32]. We cloned the 10 selected shRNA genes as well as the negative control shNS gene into lentiviral vectors (HIV-7-GFP[33]) and generated lentiviruses (LVs) containing each shRNA gene. These LVs were transduced into SupT1 cells, and the cells were sorted by GFP expression to select those with properly integrated lentiviral vectors. Collected GFP positive cells were then infected with HIV-1 NL4-3, where RT activity was measured over time

to evaluate the inhibitory effects of the different anti-HIV shRNAs. In two independent experiments shL1, shP2 and shS3 were the most effective at inhibiting HIV replication (Figure 4.2A and B).

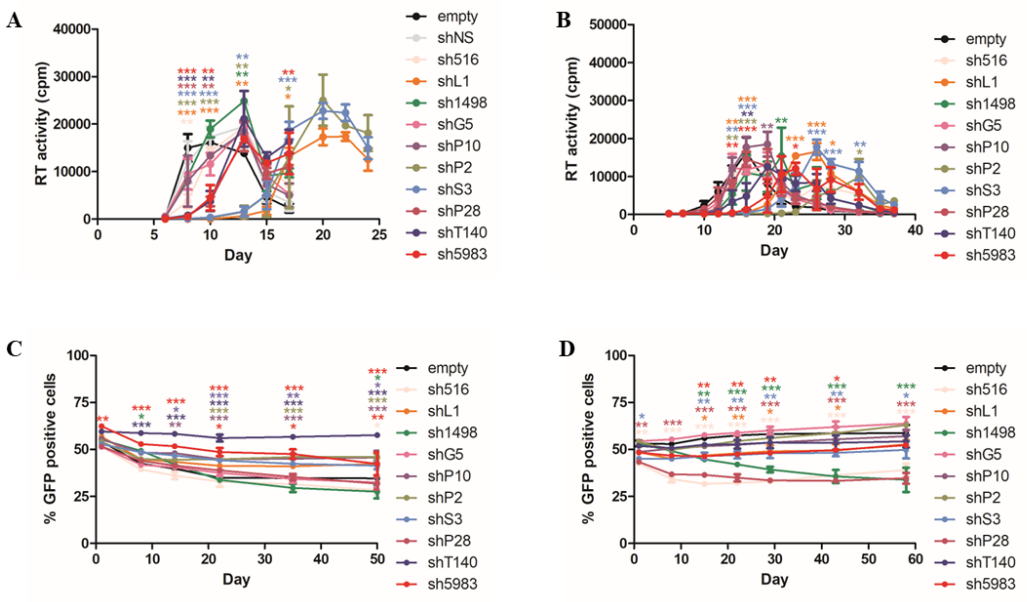

Figure 4.2 Three shRNAs expressed from the H1 promoter strongly restrict HIV replication and cells expressing shRNAs from the H1 promoter do not exhibit growth defects. (A,B) Panels A and B are independent experiments. Lentiviruses generated from HIV-7-GFP lentiviral vectors containing the different shRNA genes were transduced into SupT1 cells. These transduced SupT1 cells were sorted for GFP expression and subsequently infected with HIV-1 NL4-3. Viral kinetics were established by measuring mean RT activity (cpm) in culture supernatants at various days post infection. Each data point represents the mean ± SEM from three infections (n=3). (C,D) Panels C and D are independent experiments. Following cell sorting by GFP expression of transduced SupT1 cells, transduced GFP positive SupT1 cells were mixed with untransduced GFP negative SupT1 cells. The percentage of GFP positive cells within the cultures was measured at various days post mixing to identify a growth advantage for either GFP positive or GFP negative cells. Each data point represents the mean ± SEM from three experiments (n=3). A two-way ANOVA with Bonferroni test was used to compare means to means of empty vector transduced cells. Significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) is shown for those data points that were significantly different from the empty vector transduced controls.

Competitive growth assays were used to evaluate whether expression of any of the shRNAs results in growth defects in the transduced SupT1 cells. This was done by creating cell cultures consisting of 50% GFP positive sorted transduced SupT1 cells and 50% GFP negative sorted non-transduced SupT1 cells. The percentage of GFP-positive cells was then measured over time. None of the shRNA transduced cells displayed major growth defects compared to empty vector cells with only sh1498 having a consistent decrease in GFP positive cells over time (Figure 4.2C and D). Based on this data, we can conclude that there are no major defects in cell growth in response to expression of the different shRNAs from the H1 promoter.

4.4.3 shRNAs are more potent when expressed from the U6 and 7SK promoters but can be cytotoxic

Since the U6 and 7SK promoters have higher transcriptional efficiencies compared to the H1 promoter[5], we next investigated whether shRNAs expressed from these promoters would more potently inhibit viral production. The genes coding for shP2, shL1 and shS3 were therefore cloned within plasmids containing the 7SK or U6 promoter (psiRNA-7SKGFP::Zeo, or psiRNA-U6GFP::Zeo). Various amounts of each generated plasmid construct were cotransfected with HIV-1 pNL4-3 in HEK 293T cells. RT activity in the cell supernatants was then measured to estimate viral production. Each cpm value was normalized to the corresponding empty vector; psiRNA-7SKGFP::Zeo, psiRNA-U6GFP::Zeo or psiRNA-H1GFP::Zeo (Figure 4.3). Dose-dependent decreases in viral production were observed for all shRNA-promoter constructs. At the lowest dose, all shRNAs showed more potent inhibition of viral production when expressed from the 7SK and U6 promoters compared to when expressed from the H1 promoter. The shL1 has similar inhibitory effects when expressed from the 7SK and U6 promoters while both shP2 and shS3 were most potent when expressed from the U6 promoter.

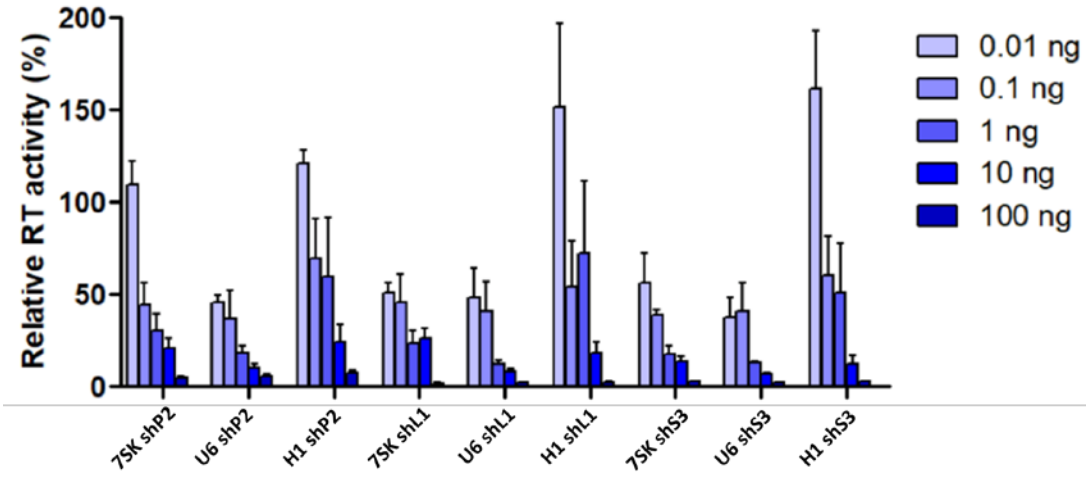

Figure 4.3 shRNAs expressed from the U6, 7SK and H1 promoters have differential inhibitory capabilities on HIV production. HEK 293T cells were co-transfected with 100 ng of the HIV-1 molecular clone pNL4-3 along with 0.01, 0.1, 1, 10 or 100 ng of 7SK, U6 and H1 driven shRNA expression plasmids. Supernatants were collected 48 h post-transfection and virus production was estimated by measuring HIV RT activity. Data are expressed as a percentage of RT activity in cells cotransfected with the empty expression plasmid. Each data point represents the mean +/- standard error mean (SEM) from one to three independent experiments with two replicates (n = 2-6).

We next investigated each construct's efficacy in long-term cultures by transducing SupT1 cells with LVs carrying the different promoter-shRNA gene constructs. Following cell sorting by GFP expression, the transduced cells were infected with HIV-1 NL4-3. Viral replication was slightly inhibited in cells expressing shP2, shL1 and shS3 from the H1 promoter (Figure 4.4A and B). Additionally, viral replication was not detected in cells expressing the shRNAs from the U6 and 7SK promoter, except for U6 driven shS3 which supported viral replication at a late time point in one infection assay. Overall, the shRNAs were most potent when expressed from the U6 and 7SK promoters within our infection assay. Competitive growth assays were also established after cell sorting to assess the cytotoxic potential of each expression construct. All shRNAs expressed from the H1 promoter maintained steady GFP levels throughout the assays, (Figure 4.4C and D). In contrast, all cultures containing shRNAs expressed from the 7SK or U6 promoter had reductions

in the GFP positivity rate over time, except for U6 and 7SK driven shP2 which maintained steady GFP levels. The reductions in the GFP positivity rate throughout the assays indicate that shL1 and shS3 expressed from the U6 and 7SK promoters cause growth defects in transduced cells. These results collectively show that shRNAs expressed from the 7SK and U6 promoters are more potent than when expressed from the H1 promoter but in the case of two out of three shRNAs, expression from these two promoters causes growth defects.

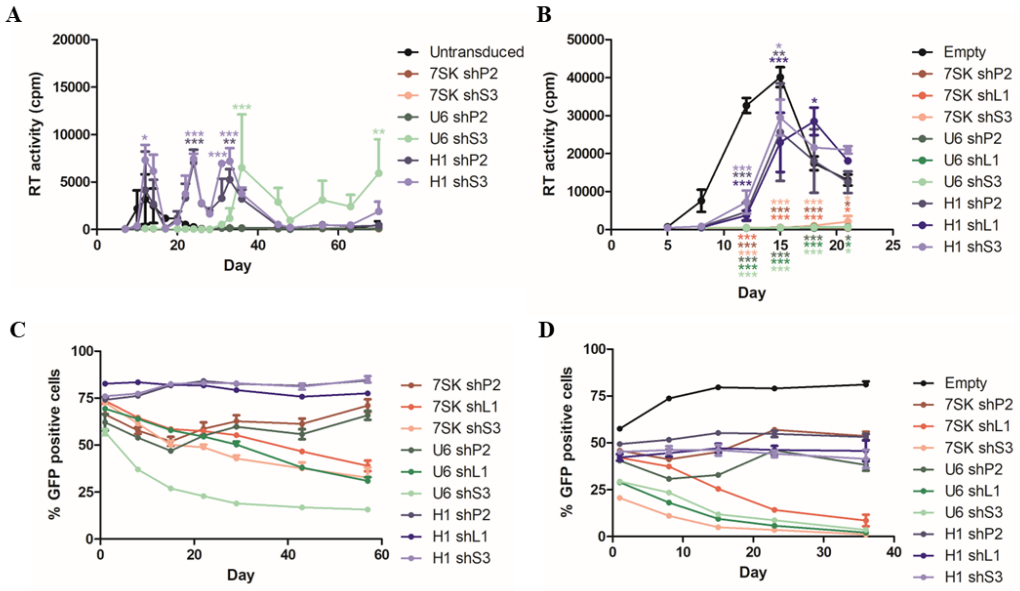

Figure 4.4 The U6 and 7SK promoters produce potent shRNA but can exhibit growth defects in cells. (A,B) Panels A and B are independent experiments. SupT1 cells were transduced with lentiviruses expressing the different shRNAs from the H1, 7SK or U6 promoter. These transduced cells were sorted for GFP expression and subsequently infected with HIV-1 NL4-3 at 1750 cpm/mL. Viral kinetics were established by measuring mean RT activity (cpm) in culture supernatants at various days post infection. Each data point represents the mean ± SEM from three infections (n=3). (C,D) Panels C and D are independent experiments. Following cell sorting by GFP expression of transduced SupT1 cells, transduced GFP positive SupT1 cells were mixed with untransduced GFP negative SupT1 cells. The percentage of GFP positive cells within the cultures was measured at various days post mixing. Each data point represents the mean ± SEM from three experiments (n=3). A two-way ANOVA with Bonferroni test was used for the infection assays to compare means to means of empty vector transduced cells or untransduced cells. Significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) is shown for those data points that were significantly different from the empty vector transduced controls.

4.4.4 shRNAs conjugated to aptamers are strong inhibitors of HIV replication and are not cytotoxic

We next wanted to investigate whether the novel molecular design reported in the literature of incorporating aptamers within the terminal loop of shRNAs could improve their inhibition of HIV replication. We therefore included the aptamer S3R3 (integrase aptamer)[7] within the loop of shP2, shL1 and shS3 as well as within the negative control, shNS. We generated LVs including each of these aptamer-shRNA chimera genes, using both 7SK and U6 promoters to drive their expression. Transduced SupT1 cells were sorted to select for cells harboring the integrated lentiviral vector and HIV RT was measured at various time points after infection with HIV-1 NL4-3. Cells expressing S3R3-shNS from either the 7SK or U6 promoter could not inhibit HIV replication, indicating that the expression of the S3R3 integrase aptamer alone could not act as a potent inhibitor of HIV (Figure 4.5A). In contrast, HIV replication was strongly inhibited in cells expressing S3R3-shL1 as well as S3R3-shS3 from the 7SK promoter whereas replication was moderately inhibited from the expression of 7SK driven S3R3-shP2 as well as U6 driven S3R3-shP2, S3R3-shS3 and S3R3-shL1 (Figure 4.5B-D). Competitive growth assays were also performed with these aptamer-shRNA chimeras. No growth defects were detected in the different cell cultures as they remained close to 50% GFP except for S3R3-L1 expressed from the U6 promoter where GFP decreased to approximately 30% in cell cultures within the first week of the assay (Figure 4.4E and F). The absence of major growth defects with the U6 and 7SK driven S3R3-shS3 and S3R3-shL1 is in sharp contrast to when the shRNAs are expressed alone from the U6 and 7SK promoters. This suggests that the complexing of a large aptamer to the terminal loop of these shRNAs effectively eliminates shRNA-mediated cytotoxicity.

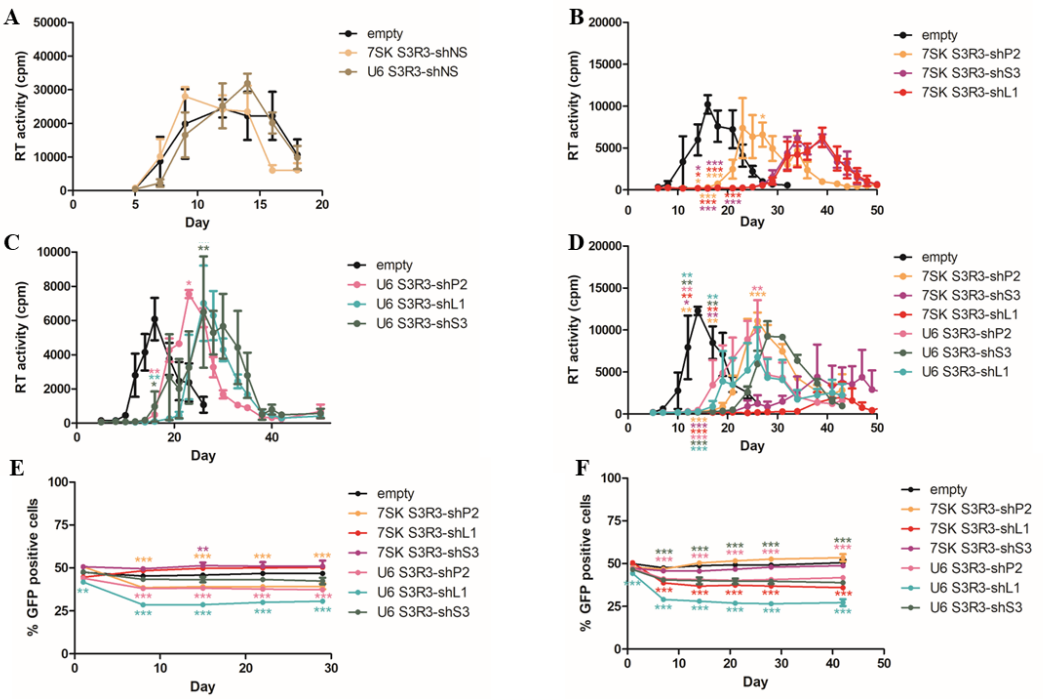

Figure 4.5 HIV replication is restricted in SupT1 cells expressing S3R3-shRNA chimeric molecules and cell growth is not compromised. (A,B,C,D) Panel D is a repeated independent experiment of panels B and C. SupT1 cells were transduced with HIV-7-EGFP lentiviral vectors expressing S3R3-shRNA chimeras from the U6 or 7SK promoters and infected with HIV-1 NL4-3 at 1750 cpm/mL. The mean RT activity (cpm) was measured in culture supernatants at various days post infection. Each data point represents the mean +/- SEM from three infections (n=3). (B) Panels E and F are independent experiments. Transduced SupT1 cells were mixed with untransduced SupT1 cells and the percentage of GFP positive cells was measured at various days post-mixing. Each data point represents the mean +/- SEM from three experiments (n=3). A two-way ANOVA with Bonferroni test was used to compare means to means of empty vector transduced cells. Significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) is shown for those data points that were significantly different from the empty vector transduced controls.

4.5 Discussion.

The use of multiple antiviral RNAs in parallel will be necessary to avoid the emergence of resistance during gene therapy to treat HIV infection. The development of viral resistance in response to anti-HIV shRNAs has been reported in several studies[16-18]. With this in mind, we have re-assessed the conservation of shRNA targets sites. The shRNAs that we screened in this study

included shRNAs developed by other groups[10,21,22,29,30], the sh1498 that we developed previously in our lab[4] and the sh5983 previously used in a clinical trial[9]. Analysis of target site conservation is important to reduce the occurrence resistance, as a high nucleotide conservation often correlates with critical genomic regions through which mutations will have severe impacts on virus fitness. Therefore, using these highly conserved areas as target sites for shRNAs can increase the genetic barrier to resistance as well as broaden the activity of the molecules across different viral strains and clades. The 23 different shRNAs screened in this study were selected according to target site conservation as well as their antiviral potency reported in the literature. Target site conservation for all shRNAs was at least 50% for LANL database genomes and at least 60% for clades B and C, with the exception of sh5983 which displayed poor target site conservation across the LANL sequences as well as sequences from clades B and C (Supplementary Table S2). While we included sh5983 because it was tested in an HIV gene therapy clinical trial[9], its poor target site conservation highlights the importance of selecting shRNAs with highly conserved target sites for future HIV gene therapy clinical trials. This will decrease the likelihood of resistance occurring and ensure antiviral efficacy across diverse viral strains and clades. It is worth noting that the shP2 and shS3 possess overlapping target sites in the HIV-1 genome. Although both of these shRNAs are used throughout this study, we are wary of using them simultaneously in a future combination gene therapy due the risk of facilitating the development of viral resistance.

After selection of anti-HIV shRNAs with highly conserved target sites, we assessed the antiviral efficacy of these molecules and their ability to elicit cytotoxic effects. Within co-transfection experiments in HEK 293T cells, our results were similar to the original shRNA screens by ter Brake *et al.*[10], McIntyre *et al.*[21], and Low *et al*[22]. However, differing results were seen in response to expression of the shG5 and shP10. These two shRNAs were shown to be potent HIV inhibitors

in the ter Brake screen but showed little to no inhibitory effects in our assay (Figure 4.1). This disparity in results could be due to the fact that we used the pNL4-3 molecular clone in our experiments whereas the ter Brake screen used the pLAI molecular clone. Although the target site of these two shRNAs is conserved between the pNL4-3 and pLAI molecular clones, the surrounding sequences are not as there are multiple single nucleotide polymorphisms along with a 36-nucleotide deletion upstream of the shP10 target site in the pNL4-3 genome. It is possible that this variability in the surrounding sequences leads to the formation of nearby RNA secondary structures which could occlude the accessibility of the shRNAs towards their target sites in the NL4-3 RNA.

Our results relating to the inhibition of viral replication after HIV challenge in SupT1 cells stably expressing the 10 selected shRNAs from the H1 promoter correspond with our results obtained during co-transfection experiments. In both cases, the the shL1, shP2 and shS3 were the strongest inhibitors of HIV while shG5 and shP10 had no effect on viral replication (Figures 4.1 and 4.2A,B). The only inconsistency in our data between co-transfection and infection experiments relates to sh516 and sh1498 which were each able to inhibit HIV production (Figure 4.1) but did not inhibit replication (Figure 4.2A and B). We have previously reported this phenomenon in the case of sh1498 expressed from the H1 promoter and shown that low levels of gene expression are responsible for the inability of the molecule to inhibit viral replication[5]. This problem can be resolved by expressing shRNAs from the 7SK and U6 promoters, which offer higher levels of gene expression.

We have shown through the data of our competitive growth assays that there are no major growth defects in SupT1 cells expressing the different shRNAs from the H1 promoter but that there was a slight decrease in the percentage of GFP positive cells expressing the sh1498 over time (Figure

4.2C and D). This was expected as we have previously reported this phenomenon and shown that sh1498 can cause growth defects that are at least partially sequence dependent[5]. However, the absence of cytotoxicity during expression of shG5 was unexpected as studies have previously characterized this molecule as having negative effects on cell growth[31]. Compared to this other study, we used an alternative loop sequence which could potentially lead to differences in shG5 processing. Variations in Dicer cleavage of the shRNA may result in differences in guide and/or passenger strand sequences, which would consequently impact the potential to elicit cytotoxicity. Additionally, we used an H1 promoter derived from the expression plasmid psiRNA while the other study used one derived from the expression plasmid pSuper[34]. The H1 promoter was used to drive gene expression in both cases, but the sequences of the H1 promoter differ slightly between the psiRNA and pSuper plasmids and this could have contributed to differences in processing or expression of shG5 with impacts on cytotoxic effects.

In an attempt to further increase the antiviral potency of shL1, shP2 and shS3, we expressed these three shRNAs from the U6 and 7SK promoters instead of by the H1 promoter. We have previously shown that expressing the sh1498 from these two promoters increases the inhibitory effect of the molecule due to higher expression levels compared to the H1 promoter[5]. However, we have also demonstrated that the higher levels of gene expression from the U6 and 7SK promoters can also lead to cytotoxic effects. Our results from both HIV inhibition assays as well as competitive growth assays with cells expressing shL1, shP2 and shS3 from the U6 and 7SK promoters are consistent with our previous investigations involving the sh1498. In each case, shRNA potency is greater when they are expressed from these two promoters instead of by the H1 promoter (Figure 4.4A and B). Additionally, some association to growth defects was once again seen when the shL1 and shS3 were expressed from the U6 and 7SK promoters (Figure 4.4C and D). Interestingly, the

cytotoxic effects from U6 and 7SK driven shL1 and shS3 were less dramatic than the cytotoxic effects that we previously reported during expression of sh1498 from these same promoters. Additionally, U6 and 7SK driven shP2 was not cytotoxic. This reinforces our proposition that shRNA-mediated cytotoxicity is at least partially sequence dependent. Cytotoxicity in response to shRNA expression can occur from off-target binding to cellular RNAs[35], saturation of the RNAi pathway[36,37] and by eliciting the innate immune system through activation of host RNA sensors[38-41]. Further studies will be needed to determine the exact mechanism leading to the growth defects, but it stands to reason that the specific sequence of the shRNA is a determinant in the potential to elicit cytotoxicity. It is worth noting that while we have not specifically measured the gene expression levels of shL1, shP2 and shS3 expressed from the different RNA Pol III promoters, the consistency in the results between these shRNAs and our previously reported study regarding the expression of sh1498 leads us to speculate that the U6 and 7SK promoters are also more transcriptionally active than the H1 promoter when expressing shL1, shP2 and shS3.

We decided to use the aptamer-shRNA molecular design reported in the literature[7] to assess whether we could once again increase the antiviral potency of the shRNAs. Although viral replication was restricted when the S3R3 integrase aptamer was incorporated within the terminal loop of shL1, shP2 and shS3, the inhibition of viral replication was greater when these three shRNAs were expressed alone from the U6 and 7SK promoters. Specifically, expression of the shRNAs alone could completely suppress viral replication in our infection assays whereas replication always eventually occurred in cells expressing the different aptamer-shRNA chimeras (Figures 4.4A,B and 4.5B-D). Interestingly, viral replication was not inhibited when cells expressed the S3R3-shNS from either the U6 or 7SK promoter (Figure 4.5A). As the shNS serves as a negative control that does not target HIV RNA, this indicates that the inhibitory capability of

the other aptamer-shRNA chimeras is strictly mediated by the shRNA component of the molecule and that the aptamer cannot act as a potent inhibitor of HIV. Comparatively, a previous study by Pang *et al.*[7] showed that the S3R3 aptamer could inhibit HIV when it was expressed as an aptamer-shRNA chimera with a negative shRNA control (shLuc). When generating the stable cell lines which express the aptamer-shRNA chimeras, this study performed cell sorting twice at 14-day intervals and selected the brightest cells in each instance. In our study, we only performed cell sorting once and omitted cells that expressed high levels of GFP as we wanted to select cells which received a single lentiviral integration event (Supplementary Figures S6-15). Due to these differences in cell sorting parameters, the stable cell lines expressing the aptamer-shRNA chimeras in our study certainly expressed lower levels of the therapeutic molecules compared to the cell lines generated by Pang *et al*. It is therefore likely that S3R3-shNS was not expressed at sufficiently high levels in our cells for the S3R3 aptamer to inhibit HIV replication.

Surprisingly, the observed growth defects associated to the expression of shL1 and shS3 from the 7SK and U6 promoters (Figure 4.4) were absent when they were expressed as aptamer-shRNA chimeras from the same promoters (Figure 4.5). Alternative shRNA loop sequences seem to be the common variable that could explain these differences in cytotoxicity results as well as our divergent cytotoxicity results relating to the shG5 compared to what is reported in the literature[31]. It is possible that differences in the sequences of the loop leads to variations in Dicer cleavage of the shRNA which could result in differences in the guide and/or passenger strand species that are generated. Alternatively, changes in the shRNA loop sequences could translate to differences in gene expression levels and as we have demonstrated in our study investigating the expression of sh1498[5], transcriptional activity does affect the potential to elicit cytotoxicity. Further

investigations will determine the mechanism leading to the absence of growth defects during expression of aptamer-shRNA chimeras.

Altogether, this study provides a comprehensive screening of shRNAs previously reported in the literature and highlights the importance of optimizing their use either by evaluating various promoters for their expression or utilizing alternative shRNA structures. As the goal of HIV gene therapy is to achieve a functional cure, it is indispensable to avoid the emergence of viral resistance in response to such a therapy. Our results are therefore valuable to achieving this goal as we demonstrate that different shRNAs are effective inhibitors of viral replication which could pave the way to the establishment of an effective anti-HIV combination gene therapy. The employment of additional classes of functional anti-HIV RNAs may also be beneficial within future studies seeking to put into place a combination gene therapy against HIV infection.

4.6 Materials and methods.

4.6.1 Cell culture

HEK293T cells were grown in Dulbecco's modified Eagle's medium (HyClone, Logan, UT) supplemented with 10% fetal bovine serum (FBS, HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies, Carlsbad, CA, USA). SupT1 cells were grown in Roswell Park Memorial Institute Medium 1640 (HyClone), supplemented with 10% heat-inactivated FBS (HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies).

4.6.2 Vector construction

Genes coding for the different shRNAs and aptamer-shRNA chimeras were generated by annealing complementary oligonucleotides or by overlapping PCR as previously described[4,15]. The

primers used are shown in Supplementary Table S3 and S4. These DNA fragments were digested with BbsI and ligated into psiRNA-hH1SKGFP::Zeo, psiRNA-7SKGFP::Zeo and psiRNA-U6GFP::Zeo. To include the anti-HIV genes into lentiviral vectors, both the associated promoter and antiviral gene was amplified by PCR out of the psiRNA expression plasmid using forward primer 5′-TATGCGGCCGCAGGGATTTTGGTCATGTTCTTAATCGATACTA-3′ and reverse primer 5′-GTAACGCCTGCAGGTTAATTAAGTCTAGAAGCTTTTCCAA-3′. The amplified DNA fragments were then digested with NotI and XbaI and ligated into NotI and XbaI-digested lentiviral transfer vector HIV-7-EGFP[33].

4.6.3 Transfection and RT assay

HEK 293T cells were seeded in 96 well plates using 150 µL of 2.5-2.75 × 10^5 cells/mL per well, 24 h prior to co-transfection. Co-transfections were performed using 20-25 µL plasmid solutions at the indicated concentrations along with the HIV-1 molecular clone pNL4-3 and using TransIT-LT1 (Mirus Bio, Madison, WI, USA) according to the manufacturer's protocol. Cell supernatants were harvested 48 h post-transfection and RT activity was measured to evaluate virus production as previously described[42,43]. Briefly, 2.5 or 5 µL of supernatant was added to 12.5 or 25 µL of non-radioactive cocktail (60 mM TrisCl, 75 mM KCl, 5mM MgCl2, 1.04 mM EDTA, 1% NP-40) and subsequently incubated with 12.5 or 25 µL of RT cocktail [60 mM TrisCl, 75 mM KCl, 5mM MgCl2, 1.04 mM EDTA, 10 µg/mL polyA (Roche, Basel, Switzerland), 0.33 µg/mL oligo dT (Life Technologies), 8 65 mM DTT, 0.05 miC/mL α^{32}P-TTP (3,000 Ci/mmol; Perkin Elmer, Waltham, MA, USA)] for 2 h at 37°C. Once the incubation was complete, 5 µL of the radioactive solution was spotted onto diethylaminoethyl (DEAE) filter mats (Perkin Elmer) or a positively charged nylon membrane (Hybond-N^+, Amersham Biosciences, Little Chalfont, UK) to detect the

resulting poly dT RT product. Five washes were then performed using 2× saline sodium citrate (SSC) buffer (20× SSC buffer: 3 M NaCl and 0.3 M sodium citrate) along with two washes using 95% ethanol to remove unincorporated [^{32}P] dTTP. Counts per minute (cpm) were measured using a microplate scintillation counter (MicroBeta TriLux; Perkin Elmer). The amount of HIV RT enzyme in the supernatants is proportional to the cpm readout.

4.6.4 Lentivirus production

HEK 293T were seeded in T75 flasks or cell culture dishes 24 h prior to co-transfections to generate LVs. Co-transfections were performed with the HIV-7-EGFP plasmid, a plasmid expressing vesicular stomatitis virus G protein and the packaging plasmid psPAX2 (Addgene, number 12260) using PEI (3 µL/µg of DNA) (Polysciences). Cell culture media was changed 3-5 h post-transfection. Supernatants were harvested 48 h after transfection, and lentiviral particles were isolated and concentrated using Lenti-X (Clontech) according to the manufacturer's protocol.

4.6.5 Infection and competitive growth

Lentivirus titers were determined by measuring the percentage of GFP-positive SupT1 cells, which were transduced with a serial dilution of lentiviral particles (range of 1 in 4 to 1 in 2048). To generate a T cell line stably expressing each antiviral gene, SupT1 cells were transduced with lentiviruses at an MOI of 1 using 8 µg/mL Polybrene (Sigma-Aldrich, St. Louis, MO, USA). Cells were sorted 72-96 h after transduction for GFP expression with a FACSAria Fusion cell sorter (BD Biosciences, Franklin Lakes, NJ, USA). The gating is shown in Supplementary Figures S6-15. To establish an HIV infection, sorted cells were seeded in 96-well round-bottom plates at 2 x 10^4 cells/well in a volume of 100 µL and infected with HIV-1 NL4-3 (1,750 cpm/mL, determined using the HIV RT assay) 24 h post-sorting. Infection kinetics was determined by measuring the RT activity of supernatants harvested at various days post-infection. To establish a competitive

growth assay, GFP positive sorted cells were seeded in 96-well round-bottom plates at either 1 x 10^4 cells/well or 2 x 10^4 cells/well in a volume of 100 μL and mixed with an equal amount of GFP negative untransduced cells that had been passed through the flow cytometer without sorting. The percentage of GFP-positive cells was measured over time using a LSRFortessa flow cytometer (BD Biosciences).

4.7 References.

1. Pau, A.K. & George, J.M. Antiretroviral therapy: current drugs. *Infect Dis Clin North Am* **28**, 371-402 (2014).
2. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).
3. Scarborough, R.J. & Gatignol, A. RNA Interference Therapies for an HIV-1 Functional Cure. *Viruses* **10**, 8 (2018).
4. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
5. Goguen, R.P., *et al.* Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK, and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021).
6. Lange, M.J., Nguyen, P.D.M., Callaway, M.K., Johnson, M.C. & Burke, D.H. RNA-protein interactions govern antiviral specificity and encapsidation of broad spectrum anti-HIV reverse transcriptase aptamers. *Nucleic Acids Res* **45**, 6087-6097 (2017).
7. Pang, K.M., Castanotto, D., Li, H., Scherer, L. & Rossi, J.J. Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. *Nucleic Acids Res* **46**, e6 (2018).
8. Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. *Mol Ther Nucleic Acids* **18**, 815-830 (2019).
9. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).
10. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol Ther* **14**, 883-892 (2006).
11. Chung, J., *et al.* Optimized lentiviral vectors for HIV gene therapy: multiplexed expression of small RNAs and inclusion of MGMT(P140K) drug resistance gene. *Mol Ther* **22**, 952-963 (2014).
12. Fire, A., *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811 (1998).

13. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**, 421-433 (2015).
14. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
15. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).
16. Das, A.T., *et al.* Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J Virol* **78**, 2601-2605 (2004).
17. Westerhout, E.M., Ooms, M., Vink, M., Das, A.T. & Berkhout, B. HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res* **33**, 796-804 (2005).
18. Boden, D., Pusch, O., Lee, F., Tucker, L. & Ramratnam, B. Human immunodeficiency virus type 1 escape from RNA interference. *J Virol* **77**, 11531-11535 (2003).
19. Gulick, R.M., *et al.* Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med* **337**, 734-739 (1997).
20. Hammer, S.M., *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* **337**, 725-733 (1997).
21. McIntyre, G.J., *et al.* 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* **6**, 55 (2009).
22. Low, J.T., *et al.* SHAPE-directed discovery of potent shRNA inhibitors of HIV-1. *Mol Ther* **20**, 820-828 (2012).
23. Zhou, J. & Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nat Rev Drug Discov* **16**, 440 (2017).
24. Ma, H., *et al.* Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Mol Ther Nucleic Acids* **3**, e161 (2014).
25. Paul, C.P., *et al.* Localized expression of small RNA inhibitors in human cells. *Mol Ther* **7**, 237-247 (2003).
26. An, D.S., *et al.* Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors. *Mol Ther* **14**, 494-504 (2006).
27. Shimizu, S., *et al.* Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**, 8 (2009).
28. Makinen, P.I., *et al.* Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* **8**, 433-441 (2006).
29. Gondai, T., Yamaguchi, K., Miyano-Kurosaki, N., Habu, Y. & Takaku, H. Short-hairpin RNAs synthesized by T7 phage polymerase do not induce interferon. *Nucleic Acids Res* **36**, e18 (2008).
30. Naito, Y., *et al.* Optimal design and validation of antiviral siRNA for targeting HIV-1. *Retrovirology* **4**, 80 (2007).
31. Herrera-Carrillo, E., Harwig, A. & Berkhout, B. Silencing of HIV-1 by AgoshRNA molecules. *Gene Ther* **24**, 453-461 (2017).

32. Centlivre, M., *et al.* Preclinical in vivo evaluation of the safety of a multi-shRNA-based gene therapy against HIV-1. *Mol Ther Nucleic Acids* **2**, e120 (2013).
33. Yam, P.Y., *et al.* Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol Ther* **5**, 479-484 (2002).
34. Brummelkamp, T.R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550-553 (2002).
35. Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci U S A* **112**, E4007-4016 (2015).
36. Yi, R., Doehle, B.P., Qin, Y., Macara, I.G. & Cullen, B.R. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA* **11**, 220-226 (2005).
37. Grimm, D., *et al.* Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *J Clin Invest* **120**, 3106-3119 (2010).
38. Fish, R.J. & Kruithof, E.K. Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. *BMC Mol Biol* **5**, 9 (2004).
39. Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**, 263-264 (2003).
40. Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H. & Williams, B.R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**, 834-839 (2003).
41. Singhania, R., Khairuddin, N., Clarke, D. & McMillan, N.A. RNA interference for the treatment of papillomavirus disease. *Open Virol J* **6**, 204-215 (2012).
42. Scarborough, R.J., Levesque, M.V., Perreault, J.P. & Gatignol, A. Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* **1103**, 31-43 (2014).
43. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. & Gatignol, A. Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp* **115**, e54486 (2016).

Chapter V

Discussion

Note: Some shRNA names have been changed from the figures and text because of pending IP discussions.

5.1 Selection of promoters for use in single and combination gene therapy

Functional anti-HIV-1 RNAs have been expressed by both the RNA Pol II[1-3] and Pol III[4-7] promoters to act as inhibitors of HIV-1. The RNA Pol II promoters have typically been used to express antiviral RNAs which have inhibitory mechanisms taking place in the cytoplasm. This is because transcripts generated from the RNA Pol II promoters are modified by post-transcriptional 5' capping and 3' polyadenylation, allowing for their nuclear export[8-12]. These extraneous elements that are added to transcripts produced from the RNA Pol II promoters have the potential to disrupt the activity of functional anti-HIV-1 RNAs. Specifically, the three-dimensional structure of aptamers may become distorted due to the presence of post-transcriptional modifications, and this could negatively impact their ability to bind to their designated target. In the case of shRNAs, the presence of extraneous elements will interfere with the ability of shRNAs to associate with cytosolic Dicer and TRBP for processing into double stranded siRNAs. Although miRNAs are driven by RNA Pol II promoters, where the removal of the 5' cap and 3' polyadenylation structures from pri-miRNAs is carried out through the actions of Drosha and DGCR8, shRNA biogenesis does not include processing steps with these proteins and therefore the removal of the resulting post-transcriptional modifications in shRNAs cannot occur[13-15]. Based on this information, we decided to focus our efforts on solely optimizing the expression of antiviral RNAs from the RNA Pol III promoters as transcripts originating from these promoters do not undergo 5' capping or 3' polyadenylation and are therefore more appropriate for the expression of functional anti-HIV-1 RNAs than the RNA Pol II promoters.

As reports in the literature are inconsistent regarding which RNA Pol III type 3 promoters (U6, H1 or 7SK) is capable of producing the most potent anti-HIV-1 RNA molecules with the lowest cytotoxic potential, it is likely that the sequence of the generated transcripts influences these

two parameters. We therefore decided to express multiple different antiviral RNAs from the U6, H1 and 7SK promoters to individually optimize the expression of each molecule by maximizing their inhibitory effects while avoiding cytotoxicity. Comparisons of all three promoters against each other was done as most other studies have only compared two RNA Pol III type 3 promoters at a time[16-18]. Apart from optimizing the expression of each anti-HIV-1 RNA, the U6, H1 and 7SK promoters were also evaluated as each will need to be utilized when designing a combination gene therapy. If a single promoter is used in such a therapy, the ensuing repeat sequences of the expression cassettes can lead to recombination between the transcriptional units, resulting in deletions in the antiviral genes during lentivirus production[19,20]. To avoid this, the multiple antiviral RNAs which will be employed in a combination gene therapy will be required to each be driven by different promoters.

Compared to the conflicting information in the literature regarding the RNA Pol III type 3 promoters, our results are generally consistent when evaluating the antiviral potency of different anti-HIV-1 shRNAs expressed from either the U6, H1 or 7SK promoters. We expressed five different shRNAs (shL1, shS3, shP2, sh5983 and sh1498) from each of these three promoters. In each case, the U6 and 7SK promoters produced more potent shRNAs compared to when the molecules were expressed from the H1 promoter (chapters II and IV). We have evaluated the expression profile of sh1498, sh5983 and shNS and showed that gene expression levels are higher when the shRNAs are driven from the U6 and 7SK promoters compared to the H1 promoter, suggesting that these differences in transcriptional activity are responsible for the observed differences in antiviral potency. Several other studies comparing the U6 and H1 promoters to express anti-HIV-1 shRNAs have also reported similar results[16,18,21]. Divergent results were reported in one study showing that shRNAs are more potent when expressed from the H1 promoter

compared to the U6 and 7SK promoters[19], and another study reporting similar antiviral effects when shRNAs were expressed from the three promoters[20]. These conflicting observations could be explained by differences in experimental design. We have characterized the activity level of shRNAs expressed from the different promoters at various doses with HIV-1 production assays, whereas the antiviral potency of the shRNAs was compared at only one dose in some of the other studies. The single dose used in these studies may have been higher than the dosage needed for complete suppression of viral production, and therefore no differences in activity level would be detected in this instance. Based on our results, the U6 and 7SK promoters generally express more potent shRNAs due to higher levels of gene expression. However, we advise against using this as a one-size-fits all rule when designing an anti-HIV-1 gene therapy due to the inconsistency of the literature. Therefore, when seeking to use shRNAs against HIV-1, it is imperative to individually optimize each molecule by expressing them from the U6, H1 and 7SK promoters to evaluate which produces the most potent therapeutic shRNA in each case.

Regarding the cytotoxic potential of shRNAs expressed from the U6, H1 and 7SK promoters, our results were once again generally consistent. Our competitive growth assays showed more dramatic decreases in the GFP positive rate over time in cultures expressing shRNAs from the U6 and 7SK promoters compared to when expression was driven by the H1 promoter. This growth advantage for GFP negative cells in the culture signals that some cytotoxicity was elicited in GFP positive cells, causing growth defects. When the different shRNAs were expressed from the H1 promoter, growth defects were not detected except for sh1498 where cultures showed a slight decrease in the GFP positivity rate over time. From this data, there seems to be a general trend in that growth defects are mostly associated to shRNAs expressed from the U6 and 7SK promoters while cytotoxicity is minimized when expression is driven by the H1 promoter.

However, the sequence identity of the shRNA appears to also be a determinant influencing the severity of the growth defects in cells. This is exemplified by the fact that regardless of the promoter that was used for expression, cultures expressing sh1498 showed more dramatic decreases in the GFP positivity rate compared to cultures expressing shNS, shL1, shS3 or shP2 (Chapter II and Chapter IV). While shRNA expression from the U6 and 7SK promoters can be associated with cytotoxic effects, this does not always seem to be the case as shP2 was not cytotoxic when expressed from these two promoters (Chapter IV, Figure 4.4C and D). Additionally, we have incorporated different large aptamers within the terminal loop of shRNAs and found that growth defects were alleviated in cultures expressing these aptamer-shRNA chimeras from the U6 and 7SK promoters (Chapter IV, Figure 4.5E and F). As the total nucleotide identity differs between canonical shRNAs and aptamer-shRNA chimeras, this reinforces our argument that the sequence identity is a determinant on the potential to elicit cytotoxicity. We therefore advise once again to individually optimize each molecule when designing an anti-HIV-1 gene therapy by expressing them from the U6, H1 and 7SK promoters to evaluate whether growth defects are present. It is important not to assume that the U6 and 7SK promoters will cause cytotoxicity in every case, but that growth defects can potentially be remedied by alternative shRNA molecular designs.

Further optimization of the expression strategy of functional RNAs to inhibit HIV-1 replication could focus on investigating the usage of RNA Pol III type 3 promoters from other species. Various studies have investigated the induction of RNAi through the expression of shRNAs by fish, murine, bovine, human and chicken derived promoters[22-25]. This could allow for the employment of additional promoters for combination gene therapy. Increasing the selection of promoters when designing these therapies would be extremely desirable to avoid the occurrence

of recombination from repeats sequences in the different expression cassettes. The nucleotide sequences of the RNA Pol III type 3 promoters from varying species should be sufficiently divergent to prevent the deletion of antiviral genes during lentiviral vector production as a result of recombination. Combination gene therapies could thus express different antiviral genes from the same promoter, albeit originating from different species. For example, a researcher may seek to minimize cytotoxicity by expressing each therapeutic gene by the H1 promoter. This would not be feasible if only the human H1 promoter is available but is possible if we consider the H1 promoter derived from alternative species. One major drawback is that shRNA gene knockdown is less potent when expression is mediated by non-human promoters in human cells[26]. It is unclear whether this difference in gene knockdown effectiveness is due to varying gene expression levels between human and non-human promoters. If a lower transcriptional activity is responsible for this difference in shRNA effectiveness, then it may be that the non-human promoters are also less likely to elicit cytotoxicity. This could therefore potentially offset the cytotoxicity of shRNAs expressed by the human U6 promoter, where a slight decrease in shRNA activity from a non-human U6 promoter would be a worthy tradeoff to ensure that an anti-HIV-1 gene therapy is safe in patient cells.

5.2 Processing of shRNAs and shRNA-aptamer chimeras

Designing an effective anti-HIV-1 gene therapy that is not cytotoxic is of upmost importance. Regardless of its potency in inhibiting HIV-1, an anti-HIV-1 RNA which causes adverse effects in cells and within *in vivo* models will never reach clinical development. Therefore, our discovery of the replacement of the canonical shRNA loop by large aptamers to eliminate shRNA-mediated cytotoxicity is indispensable in designing an effective inhibitor of HIV-1 that does not induce cytotoxicity. Equally as important as this reported effect is understanding the

underlying mechanism in the cell which is causing cytotoxicity from the expression of canonical shRNAs, but not from aptamer-shRNA chimeras. Our investigations using the different RNA Pol III promoters to express various shRNAs have all indicated that the U6 and 7SK promoters can cause shRNA mediated cytotoxicity due to the high transcriptional activity of these two promoters[6]. It therefore stands to reason that a high cellular prevalence of shRNA and its associated post-cleavage products have the potential to elicit cytotoxicity. With this in mind, we showed that the sh1498 guide strand was more abundant in cells expressing canonical sh1498 compared to those expressing S3R3-sh1498 (Chapter III, Figure 3.6). As cytotoxicity is detectable in cells expressing canonical shRNAs but not aptamer-shRNA chimeras, this result supports our statement that high cellular levels of shRNA and its metabolites can induce cytotoxicity. It is worth noting that we have evaluated the transcriptional activity of the RNA Pol III promoters when expressing three different shRNAs (sh1498, sh5983 and shNS) and found that expression levels were high from the U6 and 7SK promoter in each case (Chapter II, Figure 2.5). Due to the uniformity of these results, we suspect that the expression levels of canonical shRNAs and aptamer-shRNA chimeras to actually be quite similar. Instead, it is likely that the aptamer-shRNA chimeras are deficient in their ability to be processed. This could be due to suboptimal cleavage by Dicer or loading of the aptamer-shRNA chimeras into RISC may be inefficient. Interestingly, this presumed deficiency in shRNA processing does not render the aptamer-shRNA chimeras ineffective as antivirals. Slight inhibition of viral replication was seen when cells expressed PR10.9-8N-sh1498, Dp6.12-sh1498 and S3R3-sh1498 (Chapter III, Figure 3.4A,B,D and E) while moderate to strong inhibition was seen from the expression of S3R3-shP2, S3R3-shL1 and S3R3-S3 (Chapter IV, Figure 4.5B-D). Comparatively, the antiviral potency of canonical sh1498, shP2, shL1 and shS3 was greater than

the aptamer-shRNA chimeras, but the cytotoxicity associated to the canonical sh1498, shL1 and shS3 may not warrant their usage in future gene therapies (Chapter IV).

Although we have taken the first step into elucidating the mechanism explaining the absence in cytotoxicity from the expression of aptamer-shRNA chimeras, further details must be resolved. Firstly, although we hypothesize that shRNA processing of the aptamer-shRNA chimeras is compromised, we do not possess complete experimental proof to support this. We are currently working to optimize our Northern blot to detect unprocessed shRNA and verify if larger quantities are found in cells expressing S3R3-sh1498 compared to cells expressing canonical sh1498. Alternatively, RNA sequencing could also detect the prevalence of unprocessed shRNA or aptamer-shRNA chimera present in cells expressing these molecules. Additionally, Dicer cleavage patterns may differ between canonical shRNAs and aptamer-shRNA chimeras regardless of cleavage efficiency. This could result in different guide and/or passenger strand species being generated with differential cytotoxic potential. Again, RNA sequencing experiments would be useful as the results could elucidate the nucleotide identity of the Dicer cleavage products from canonical shRNAs and aptamer-shRNA chimeras to determine whether the incorporation of large aptamers in the terminal loop of shRNAs is influencing Dicer cleavage patterns.

While we can detect growth defects in SupT1 cells expressing shRNAs, it is not clear as to how the shRNAs or their cleavage product are causing adverse cellular effects. From the literature, we know that shRNA mediated cytotoxicity can be caused by off-target binding to cellular RNAs[27], saturation of the RNAi pathway[28,29], as well as eliciting the innate immune system through activation of host RNA sensors[30-33]. Although microarray experiments have been performed when designing the sh1498 to ensure that there were minimal off-target effects on human RNAs[4], further investigations may be necessary to confirm that this shRNA is not

mediating cytotoxicity by binding to cellular RNAs. It would also be worthwhile to overexpress cellular Dicer or other proteins included in the RISC by introducing corresponding expression plasmids. If the overexpression of these proteins alleviates the shRNA induced cytotoxicity, this would indicate that the observed growth defects are caused by saturation of the RNAi pathway. Finally, Western blots could be performed to assess whether proteins involved in the innate immune response to dsRNAs, such as JAK and STAT proteins, are being activated from the expression of shRNAs to cause cytotoxic effects. Another point that is unclear is whether the observed growth defects of GFP positive cells in our experiments are due to a true decrease in the rate of cell growth or whether a portion of cells expressing anti-HIV-1 RNAs are undergoing cell death. If cell death is responsible for the observed growth defects in our competitive growth assays, it would be valuable to assess whether this is occurring by apoptosis or necrosis. These two mechanisms of cell death can be differentiated by their associated cellular metabolites and therefore molecular markers unique to apoptosis or necrosis can be detected by fluorescence-activated cell sorting (FACS) to determine the route of cell death[34,35].

Our investigations relating to the mechanism of shRNA mediated cytotoxicity have thus far been carried out in HEK293T cells. While these cells are ideal for transfection of shRNA expression vectors for quick assessment by Northern blot, they are not a relevant cell type in the context of anti-HIV-1 gene therapy. The lymphocytic cell line SupT1 is more relevant towards our overall gene therapy project but experiments involving their transduction by shRNA coding lentiviral vectors are laborious due to the need for cell sorting to isolate cells with successfully integrated lentiviral vectors. We have therefore used HEK293T cells in our experiments to generate preliminary data as to the prevalence of the sh1498 guide strand and to optimize our Northern blot conditions. So far, cell sorting conditions after lentiviral transduction of SupT1 cells

has yielded sufficient RNA to perform Northern blot experiments in order to evaluate whether our results pertaining to HEK293T cells are similar in the case of SupT1 cells.

5.3 Perspectives for future combination gene therapies to prevent viral resistance

Astounding progress has been achieved since the start of the HIV-1 pandemic in developing therapeutics which can inhibit the replicatory actions of the virus to improve the quality of life of infected individuals as well as increasing their life span[36]. However, the battle is not over as these antiretroviral therapies are not curative and must be administered chronically to maintain low levels of viremia in infected individuals[37]. Modifying patient HSCs using gene therapy by introducing antiviral genes within the chromosomes could allow for control of the infection following a single transplant of modified resistant cells[38]. As is the case with current antiretroviral therapies, the development of viral resistance is a continuous threat to the effectiveness of inhibitors. Therefore, the use in combination of multiple genes coding for antiviral RNAs targeting different conserved sites will be necessary to avoid the occurrence of resistance during gene therapy[39,40]. Using multiple antiviral RNAs which each target different elements of the viral replication cycle will make it exceedingly difficult for the virus to escape such therapies as several beneficial mutations must be acquired simultaneously to render each therapeutic agent inert.

Clinical trials have been conducted to assess the feasibility and safety of gene therapy to treat HIV-1 infection[41,42]. Due to ethical concerns, clinical trials infusing gene modified cells must also include the infusion of unmanipulated cells as the toxicity associated to the transduced cells is unknown. While this is important to ensure the safety of patients, it is troublesome when attempting to evaluate whether the expression of antiviral genes can effectively inhibit HIV-1 replication *in vivo*. Therefore, the focus has rather been on the tolerance to transgene expression in patients as well as characterizing the length of time that the therapeutic gene product is

detectable post infusion. Clinical trials have documented drops in the expression of antiviral RNAs following infusion in patients but have not reported any major adverse events from the therapies[41,42]. The presence of HIV-1, and thus the interruption of cART, is likely needed as a selection factor to maintain levels of therapeutic molecules over time but this poses another ethical dilemma as cessation of antiretroviral therapy is associated with increased morbidity and mortality[43]. The conservative regulations surrounding gene therapy clinical trials have certainly influenced their effectiveness which makes it difficult to determine whether these therapies can control HIV-1 infection alone. However, it is necessary to err on the side of caution as looser restrictions in previous clinical trials have led to patient mortality[44]. Currently, gene therapies to treat HIV-1 infection must be demonstrably safe and feasible in early phase clinical trials, which has been accomplished in at least two instances. These therapies must then be evaluated to determine whether they can inhibit viral replication indefinitely without the help of cART, as well as not causing the emergence of viral resistance. Future clinical trials may want to enroll participants who have not yet initiated cART, so that treatment interruption and its associated risks are avoided.

Although other research groups have established combination gene therapies that have progressed to clinical trials, their limited effectiveness demonstrates that there is a need for more potent anti-HIV-1 RNAs. Our goal was therefore to optimize individual antiviral RNAs by exploring varying expression strategies and molecular designs to maximize antiviral potency while avoiding cytotoxicity before establishing a combination gene therapy. Our results show that modifying these two parameters strongly impacts the inhibitory capabilities and cytotoxic potential of anti-HIV-1 RNAs. While we could achieve notable suppression of viral replication without cytotoxicity when using optimized anti-HIV-1 RNAs, expression of these optimized molecules

alone consistently led to eventual productive replication of HIV-1 in our experiments. However, when we evaluated double molecule therapy by expressing a U1i RNA along with an shRNA or shRNA-chimera, there were instances where viral replication did not occur at any time point in our assays (Appendix). Unfortunately, our most effective double molecule therapies were not infallible, with at least one out of three infection samples supporting replication cycles in our experiments. We theorize that replication in these cases was due to the development of viral resistance towards the combination gene therapies. Puzzlingly, sequencing of the therapeutic target sites within the viral genome of samples which supported replication did not reveal mutations in these genomic regions. As we did not sequence the entirety of the HIV-1 RNA in these samples, viral resistance may instead have been conferred by mutations outside of the target site which we could not detect in our experiments. Nevertheless, our preliminary work in establishing double molecule therapies with optimized anti-HIV-1 RNAs showed a greater level of inhibition of viral replication compared to single molecule therapy.

While our preliminary results using combination gene therapy are encouraging, we are actively testing additional molecule combinations. The goal is to identify either a double or triple molecule gene therapy which can inhibit viral replication indefinitely in cell culture before evaluating such a therapy *in vivo*. The potential for adverse effects from toxicity will be rigorously tested, with special considerations depending on the molecules selected to establish the combination therapy. For example, saturation of the RNAi pathway will be monitored if multiple shRNAs are used so as not to dysregulate cellular gene expression. In addition to testing our therapy *in vivo*, the ability for cellular differentiation of transduced $CD34^+$ cells will also be evaluated *in vitro* to determine whether myeloid and lymphoid cells can be produced which maintain expression of the antiviral molecules. This is routinely done in preclinical studies, and

other research groups have successfully transduced HSCs to express anti-HIV-1 RNAs without adverse effects on cellular differentiation[45]. Murine models will likely serve as the starting point to assess the effectiveness of our therapy *in vivo*, where results supporting its safe and effective use could allow for the commencement of clinical trials with HIV-1 positive participants. As is the case with other clinical trials that have been conducted to treat HIV-1 infection by gene therapy[41,42], the early phase trials of our therapy will need to abide by conservative regulations to ensure the safety of the enrolled patients. Specifically, HIV-1 positive patients who suffer from lymphoma or leukemia will be selected to allow for an ethically acceptable setting where myeloablative therapy is performed prior to HSC engraftment. Additionally, non-transduced HSCs will need to be mixed with gene modified HSCs during the engraftment process and all patients will likely be maintained on cART throughout the clinical trial.

5.4 Conclusions

A functional cure for HIV-1 infection is needed to eliminate the need for cART. Gene therapy has the potential to provide such a functional cure and while clinical trials using gene therapy to treat HIV-1 infection have been conducted, their effectiveness in controlling the infection is limited. We therefore focused on optimizing individual anti-HIV-1 RNAs *in vitro* with an eventual goal of designing a combination gene therapy that is safe and effective at inhibiting viral replication. Our research into such optimization showed that various shRNAs are more potent and expressed at higher levels when transcription is driven by the U6 and 7SK promoters rather than the less transcriptionally active H1 promoter. Additionally, this high transcriptional activity can lead to cytotoxicity when shRNAs are expressed from the U6 and 7SK promoters. These results are corroborated by other investigators. However, we advise against assuming that this is the case in every instance as we have shown that at least one shRNA is not cytotoxic when

expressed from the U6 and 7SK promoter and there has been a report in the literature of shRNAs being more potent when expressed by the H1 promoter rather than the U6 and 7SK promoter.

Promoter selection is not the sole parameter for the optimization of anti-HIV-1 RNAs. We also found that the molecular design of antivirals was an important determinant towards their inhibitory capabilities and their potential to elicit cytotoxicity. Indeed, our research has demonstrated that replacing the canonical loop of shRNAs by large aptamers resolves the cytotoxicity that is elicited by shRNAs expressed alone from the U6 and 7SK promoters. Our data suggests that this is due to a lower abundance of shRNA Dicer cleavage products in cells expressing aptamer-shRNA chimeras compared to cells expressing canonical shRNAs. We hypothesize that this effect is a consequence of a deficiency in Dicer cleavage of the aptamer-shRNA chimeras. Regardless of the mechanism, three of the aptamer-shRNA chimeras assessed in the project were able to strongly inhibit viral replication without inducing cytotoxicity.

Overall, this project provided valuable information contributing towards our eventual goal in designing a safe and effective anti-HIV-1 combination gene therapy. Proper evaluation of the different RNA Pol III promoters was necessary as each of the U6, 7SK and H1 promoter will be required in a combination gene therapy to avoid recombination between the different transcriptional units. Additionally, our investigation in alternative shRNA molecular designs has led to the discovery of a method to resolve shRNA mediated cytotoxicity which is critical to ensuring the safety of a future combination gene therapy. Overall, our results have contributed to a better understanding of how the expression strategy and molecular design of anti-HIV-1 RNAs can influence their antiviral capabilities and cytotoxic potential.

References.

1. Ramalingam, D., *et al.* RNA aptamers directed to human immunodeficiency virus type 1 Gag polyprotein bind to the matrix and nucleocapsid domains and inhibit virus production. *J Virol* **85**, 305-314 (2011).
2. Duclair, S., Gautam, A., Ellington, A. & Prasad, V.R. High-affinity RNA Aptamers Against the HIV-1 Protease Inhibit Both In Vitro Protease Activity and Late Events of Viral Replication. *Mol Ther Nucleic Acids* **4**, e228 (2015).
3. Lange, M.J., Nguyen, P.D.M., Callaway, M.K., Johnson, M.C. & Burke, D.H. RNA-protein interactions govern antiviral specificity and encapsidation of broad spectrum anti-HIV reverse transcriptase aptamers. *Nucleic Acids Res* **45**, 6087-6097 (2017).
4. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
5. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).
6. Goguen, R.P., *et al.* Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK, and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021).
7. Li, M.J., *et al.* 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**, 900-909 (2005).
8. Good, P.D., *et al.* Expression of small, therapeutic RNAs in human cell nuclei. *Gene Ther* **4**, 45-54 (1997).
9. Egloff, S. & Murphy, S. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**, 280-288 (2008).
10. Lewis, J.D. & Izaurralde, E. The role of the cap structure in RNA processing and nuclear export. *Eur J Biochem* **247**, 461-469 (1997).
11. Banerjee, A.K. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol Rev* **44**, 175-205 (1980).
12. Coller, J.M., Gray, N.K. & Wickens, M.P. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes Dev* **12**, 3226-3235 (1998).
13. Lee, Y., *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060 (2004).
14. Lambeth, L.S. & Smith, C.A. Short hairpin RNA-mediated gene silencing. *Methods Mol Biol* **942**, 205-232 (2013).
15. Garcia-Lopez, J., Brieno-Enriquez, M.A. & Del Mazo, J. MicroRNA biogenesis and variability. *Biomol Concepts* **4**, 367-380 (2013).
16. An, D.S., *et al.* Optimization and functional effects of stable short hairpin RNA expression in primary human lymphocytes via lentiviral vectors. *Mol Ther* **14**, 494-504 (2006).
17. Shimizu, S., *et al.* Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther* **7**, 8 (2009).

18. Makinen, P.I., *et al.* Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* **8**, 433-441 (2006).
19. Spanevello, F., *et al.* Development of Lentiviral Vectors Simultaneously Expressing Multiple siRNAs Against CCR5, vif and tat/rev Genes for an HIV-1 Gene Therapy Approach. *Mol Ther Nucleic Acids* **5**, e312 (2016).
20. ter Brake, O., *et al.* Lentiviral Vector Design for Multiple shRNA Expression and Durable HIV-1 Inhibition. *Mol Ther* **16**, 557-564 (2008).
21. Sun, C.P., *et al.* Studies of efficacy and liver toxicity related to adeno-associated virus-mediated RNA interference. *Hum Gene Ther* **24**, 739-750 (2013).
22. Zenke, K. & Kim, K.H. Novel fugu U6 promoter driven shRNA expression vector for efficient vector based RNAi in fish cell lines. *Biochem Biophys Res Commun* **371**, 480-483 (2008).
23. Lambeth, L.S., Wise, T.G., Moore, R.J., Muralitharan, M.S. & Doran, T.J. Comparison of bovine RNA polymerase III promoters for short hairpin RNA expression. *Anim Genet* **37**, 369-372 (2006).
24. Wise, T.G., *et al.* Characterization and comparison of chicken U6 promoters for the expression of short hairpin RNAs. *Anim Biotechnol* **18**, 153-162 (2007).
25. Lambeth, L.S., *et al.* Characterisation and application of a bovine U6 promoter for expression of short hairpin RNAs. *BMC Biotechnol* **5**, 13 (2005).
26. Roelz, R., Pilz, I.H., Mutschler, M. & Pahl, H.L. Of mice and men: human RNA polymerase III promoter U6 is more efficient than its murine homologue for shRNA expression from a lentiviral vector in both human and murine progenitor cells. *Exp Hematol* **38**, 792-797 (2010).
27. Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci U S A* **112**, E4007-4016 (2015).
28. Yi, R., Doehle, B.P., Qin, Y., Macara, I.G. & Cullen, B.R. Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA* **11**, 220-226 (2005).
29. Grimm, D., *et al.* Argonaute proteins are key determinants of RNAi efficacy, toxicity, and persistence in the adult mouse liver. *J Clin Invest* **120**, 3106-3119 (2010).
30. Fish, R.J. & Kruithof, E.K. Short-term cytotoxic effects and long-term instability of RNAi delivered using lentiviral vectors. *BMC Mol Biol* **5**, 9 (2004).
31. Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.L. & Iggo, R. Induction of an interferon response by RNAi vectors in mammalian cells. *Nat Genet* **34**, 263-264 (2003).
32. Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H. & Williams, B.R. Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* **5**, 834-839 (2003).
33. Singhania, R., Khairuddin, N., Clarke, D. & McMillan, N.A. RNA interference for the treatment of papillomavirus disease. *Open Virol J* **6**, 204-215 (2012).
34. Ryan, J., Montero, J., Rocco, J. & Letai, A. iBH3: simple, fixable BH3 profiling to determine apoptotic priming in primary tissue by flow cytometry. *Biol Chem* **397**, 671-678 (2016).
35. Rys, R.N., *et al.* Apoptotic Blocks in Primary Non-Hodgkin B Cell Lymphomas Identified by BH3 Profiling. *Cancers (Basel)* **13**(2021).
36. Pau, A.K. & George, J.M. Antiretroviral therapy: current drugs. *Infect Dis Clin North Am* **28**, 371-402 (2014).

37. Hawkins, T. Understanding and managing the adverse effects of antiretroviral therapy. *Antiviral Res* **85**, 201-209 (2010).
38. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).
39. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol Ther* **14**, 883-892 (2006).
40. Chung, J., *et al.* Optimized lentiviral vectors for HIV gene therapy: multiplexed expression of small RNAs and inclusion of MGMT(P140K) drug resistance gene. *Mol Ther* **22**, 952-963 (2014).
41. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).
42. Mitsuyasu, R.T., *et al.* Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nature medicine* **15**, 285-292 (2009).
43. Strategies for Management of Antiretroviral Therapy Study, G., *et al.* CD4+ count-guided interruption of antiretroviral treatment. *N Engl J Med* **355**, 2283-2296 (2006).
44. Sibbald, B. Death but one unintended consequence of gene-therapy trial. *CMAJ* **164**, 1612 (2001).
45. Wolstein, O., *et al.* Preclinical safety and efficacy of an anti-HIV-1 lentiviral vector containing a short hairpin RNA to CCR5 and the C46 fusion inhibitor. *Mol Ther Methods Clin Dev* **1**, 11 (2014).

Chapter VI

Appendix

Identification of highly active combinations of anti-HIV-1 RNA molecules

Note: Some shRNA names have been changed from the figures and text because of pending IP discussions; Supplementary Figures are not included because of IP considerations.

6.1 Abstract.

Combination gene therapy based on the expression of antiviral RNAs has the potential of giving rise to a functional cure for HIV-1 infection by providing a long-term therapy, which is currently not the case for cART. Advancements have been made in developing such a combination gene therapy, with clinical trials demonstrating that these therapies are safe for PLWH. As our goal is to establish an optimal combination, we first attempted to optimize the expression and molecular design of various anti-HIV-1 RNAs. We show that extending the recognition domain at the 5' end of a U1i RNA molecule leads to increased antiviral potency without any signs of cytotoxicity. These optimized U1i RNAs were expressed simultaneously with optimized shRNAs to generate double molecule gene therapies. Among these combinations, we identified four combinations, which prevented viral replication for at least one out of three infection samples. Additionally, we report that growth defects can be detected in the case of certain double molecule combination therapies. These results suggest that the utilization of two antiviral RNAs can prevent HIV-1 replication with no associated cytotoxicity. Additional combinations must be explored to completely suppress viral replication and identify an optimal combination.

6.2 Introduction.

Combination antiretroviral therapy (cART) is extremely effective in delaying the progression to acquired immunodeficiency syndrome (AIDS) from human immunodeficiency virus type 1 (HIV) infection. However, these therapies are not curative, are associated with undesirable side effects and impose financial constraints on both patients and health care systems[1]. Therefore, a curative strategy which could resolve these issues is highly desirable. Genetic manipulation of patient cells to render them HIV resistant could be used as a functionally curative strategy to solve the issues associated to cART[2,3]. This process is known as gene therapy and various studies have shown that the expression of functional anti-HIV-1 RNA molecules within cells can inhibit viral replication[4-8].

An approach to achieving a curative strategy for HIV infection has eluded scientists for decades, with cART being able to control but not eliminate the infection and vaccine trials being largely ineffective due to the genetic diversity of the virus[9,10]. The idea of gene therapy to treat HIV infection has gained traction through three different patients and likely two others being cured from HIV-1 after receiving hematopoietic stem cell (HSC) transplants from HIV resistant donors following myeloablative therapy to treat their acute myeloid leukemia or Hodgkin's lymphoma[11]. In each of these cases, HIV-1 resistance was mediated by a homozygous 32 base pair deletion genotype in the chemokine receptor 5 gene (CCR5Δ32/Δ32) within the donor HSCs[11]. The first success using this approach took place in 2007 with Timothy Brown (the "Berlin Patient") while the second success took place in 2016 with Adam Castillejo (the "London Patient")[12,13]. A third case is the recently confirmed "Düsseldorf" patient[14]. A fourth case occurred in a woman transfused with cord blood from a donor having the CCR5Δ32 mutation[15] and a fifth case was recently announced as the "City of Hope patient", a 66 years old man who also had an HSC

transplant from a CCR5Δ32 donor[16]. An apparent elimination of the virus was observed in both patients following their transplant as HIV-1 was undetectable within their tissues and bodily fluids in the absence of cART. These cases have served as a "proof of cure", showing that transplantation of genetically resistant cells can control HIV-1 infection and provide the means for a functional cure. Scaling up this exact procedure to treat all HIV-1 infected individuals is unfeasible due to the scarcity of individuals containing the CCR5Δ32/Δ32 genotype and the risk of graft versus host disease (GVHD) in the recipient[17,18]. Therefore, modifying patient cells ex-vivo by gene therapy to express functional anti-HIV-1 RNAs could provide a functional cure for infected individuals to mimic the cases of the "Berlin, London, Düsseldorf and City of Hope patients", as well as the cord-blood transfused woman.

Various classes of functional anti-HIV-1 RNA molecules have been developed to inhibit HIV-1 replication by gene therapy, including short hairpin RNAs (shRNAs), aptamers and U1 interference RNAs[3,8]. shRNAs function by co-opting elements of the RNAi pathway to mediate gene silencing by target mRNA cleavage. This cleavage occurs by the actions of Ago2, which is directed to a particular mRNA species through nucleotide complementarity of the shRNA guide strand with the target mRNA[19]. We have previously shown that an shRNA targeting the *gag* region of HIV-1, called sh1498, is more potent when expressed from the RNA polymerase (Pol) III promoters U6 and 7SK but that the higher expression levels from these promoters also leads to cytotoxicity[4,5]. Additionally, we have identified three shRNAs (shL1, shS3 and shP2) which are more potent inhibitors than sh1498 and have shown that the antiviral capabilities of these three shRNAs is also maximized when they are expressed from the U6 and 7SK promoters (Chapter IV). U1i RNAs are reengineered U1 small nuclear RNA (U1 snRNA) which can be directed to bind to HIV-1 transcripts and cause recruitment of spliceosomal components which will inhibit

viral replication[8]. Binding of U1i RNAs to either 5'ss or downstream of 3'ss will result in an enhancement of splicing of viral transcripts which leads to the absence of full length viral transcripts being available for packaging of new virions[20]. Binding of U1i RNAs to the polyadenylation signal of viral transcripts is also possible, effectively inhibiting polyadenylation[21]. We have previously shown that an anti-HIV-1 U1i RNA which enhances splicing, called U1-T6, is an effective inhibitor of HIV-1[8]. Finally, aptamers are RNA molecules which adopt a specific three-dimensional structure to allow them to bind to a specific target and render it inert[2]. These molecules are generated by systematic evolution of ligands by exponential enrichment (SELEX) using an RNA library with randomized sequences[22]. Recently, aptamers have been incorporated within the terminal loop of shRNAs[7]. Processing of these aptamer-shRNA chimeras by cellular Dicer causes cleavage of the stem-loop and thus separation of the aptamer from the shRNA. We have previously shown that aptamer-shRNA chimeras effectively prevent the cytotoxic effects that are seen when shRNAs are expressed on their own from the U6 and 7SK promoters[5] (Chapters III and IV).

The emergence of viral resistance is a constant threat to the effectiveness of current antiretroviral drugs and the administration of double or triple drug therapies is required to avoid the occurrence of resistance[1,23,24]. Similarly, resistance to functional anti-HIV-1 RNA molecules can occur, causing an abolishment of antiviral activity. This resistance is driven primarily by point mutations within the target site of the antiviral RNAs but can also be due to mutations outside of the target site, resulting in the formation of RNA secondary structures which impede accessibility to the target sequence[25-27]. The solution to this dilemma is identical to cART, where administrating multiple antiviral molecules in parallel will prevent the emergence of viral resistance by increasing the genetic barrier to resistance[28,29]. The various classes of anti-HIV-1 RNAs therefore offer

diverse mechanisms to inhibit viral replication and can be used in combination to avoid viral resistance. With this in mind, we describe here the optimization of the U1 T6 by extending the length of its recognition domain and use this optimized molecule to establish various combination gene therapies. Specifically, we observed that extending the recognition domain of the U1 T6 at the 5' end allows the molecule to inhibit HIV-1 replication without causing cytotoxic effect whereas recognition domain extensions at the 3' end were associated with growth defects. We also show that double molecule combination therapies expressing these 5' extended U1 T6 molecules along with different anti-HIV-1 shRNAs or aptamer-shRNA chimeras can prevent viral replication from occurring.

6.3 Results.

6.3.1 Cells transduced with U1 T6 5' end length variants restrict HIV-1 replication

We have previously screened different U1i RNAs and identified the splicing enhancer U1-T6 as the most effective design[8]. To improve its specificity, we extended its 10-nucleotide recognition site on the 3' or 5' by either 3 or 6 nucleotides and tested these length variants against HIV-1 replication while evaluating effects on cell growth. Measuring the effects against HIV-1 replication was done by challenging U1i RNA-expressing SupT1 cells with HIV NL4-3. The U1i RNA-expressing SupT1 cells were generated by transduction with lentiviruses carrying a particular antiviral gene, followed by cell sorting by GFP to isolate cells with properly integrated lentiviral vectors. The negative control consisted of U1 WT and cells transduced with lentiviral vectors coding for this molecule had similar viral replication kinetics as cells transduced with the empty lentiviral vector negative control. Comparatively, expression of U1 T6 and all its length variants inhibited HIV-1 replication compared to the negative control empty vector, with the strongest

inhibition being caused by the 5' end length variants (Figure 6.1A and B). Effects on cell growth were measured by establishing competitive growth assays after cell sorting of lentiviral transduced SupT1 cells. GFP positive sorted transduced SupT1 cells were mixed with an equal number of GFP negative sorted non-transduced SupT1 cells and the ratio of GFP positive cells was measured over time. In contrast to our previous results where U1-T6 was not associated to cytotoxicity[8], U1-T6 transduced cells had slight growth defects in the competitive growth assay (Figure 6.1C and D). The 3' end length variant of the U1 T6 molecule also conferred some degree of growth defect, with the largest decrease in the % of GFP positive cells being associated to the 3-nucleotide extension. In contrast, the cultures of both 5' end length variants remained close to 50% GFP, showing a lack of cytotoxicity which makes them good candidates for use in combination therapy.

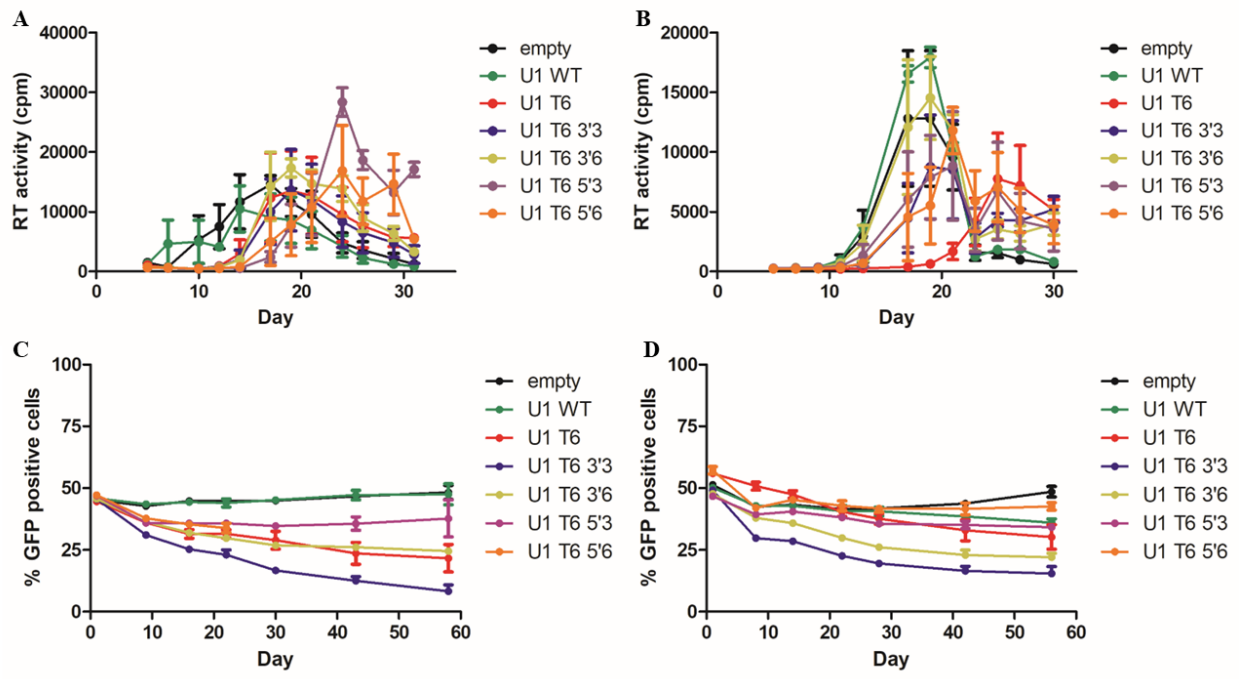

Figure 6.1 HIV-1 replication is slightly restricted in SupT1 cells expressing U1 T6 5' end length variants and cell growth is not compromised when expressing 5' end length variants. (A,B) Panels A and B are repeat experiments. SupT1 cells were transduced with HIV-7-EGFP

lentiviral vectors expressing U1 T6 length variants and infected with HIV-1 NL4-3 at 1750 cpm/mL. The mean RT activity (cpm) was measured in culture supernatants at various days post infection. Each data point represents the mean +/- SEM from three infections (n=3). (C,D) Panels C and D are repeat experiments. Transduced SupT1 cells were mixed with untransduced SupT1 cells and the percentage of GFP positive cells was measured at various days post-mixing. Each data point represents the mean +/- SEM from three experiments (n=3).

6.3.2 Combination therapy with U1 T6 5' end length variants and shRNAs can prevent HIV-1 infection

As the 5' end length variants of U1 T6 could restrict HIV-1 replication with no negative effects on cell growth, we selected these to use in combination inhibitory assays along with sh1498[4], shP2 and shS3 expressed from the H1 promoter, PR10.9-8N-sh1498 (Chapter III) expressed from the U6 promoter, as well as S3R3-shP2, S3R3-shL1 and S3R3-shS3 (Chapter IV) expressed from the 7SK promoter. These combination therapies were established by introducing a multiple cloning site (MCS) within our lentiviral vectors to allow for the cloning of multiple anti-HIV-1 expression cassettes within the vector. Transduction of the generated lentiviruses within SupT1 cells allows for simultaneous expression of multiple antiviral molecules within the cells. When looking at the curve generated from the average of three infections, HIV-1 replication seems to only be strongly inhibited in cells expressing the combinations U1 T6 5'3 with H1 driven shP2 and U1 T6 5'6 with 7SK driven S3R3-shL1 (Figure 6.2A-C). However, there are four combinations where viral replication does not occur in some cases. This can be seen by individually representing the data of each infection within the inhibitory assay for these four combinations (Figure 6.2D-G). Specifically, only one well supported viral replication for the combinations U1 T6 5'3 with H1 driven shP2 and U1 T6 5'6 with H1 driven shP2 (Figure 6.2D and E). In the case of the combinations U1 T6 5'3 with H1 driven sh1498 and U1 T6 5'6 with U6 driven PR10.9-8N-sh1498, only two wells supported viral replication (Figure 6.2F and G). Competitive growth assays were

also put into place for the various combination therapies. Within our first assay, only the combinations U1 T6 5'3 with U6 driven PR10.9-8N-sh1498 and U1 T6 5'6 with U6 driven PR10.9-8N-sh1498 cell cultures had similar percentages of GFP positive cells as the negative control empty vector cell cultures (Figure 6.2H and I). The cultures in each of the other combinations displayed lower percentages of GFP positive cells compared to the negative control empty cell cultures and therefore growth defects are likely occurring when cells express these various combinations of anti-HIV-1 molecules. However, as the % GFP increased over time in negative control empty cell cultures due to the sorted GFP-negative SupT1 cells being unhealthy following cell sorting, it is difficult to rely upon these results to make definitive conclusions on the extent of cytotoxicity that is elicited by these combinations. Within our second assay, the negative control empty cells cultures maintained an approximate 50% GFP positive rate throughout the assay and the results are therefore more reliable (Figure 6.2J). Each of the combinations within this assay showed initial signs of cytotoxicity as the percentage of GFP positive cells decreased during the first week of the assay, with the most dramatic decrease in GFP occurring for the combination U1 T6 5'6 with 7SK driven S3R3-shL1.

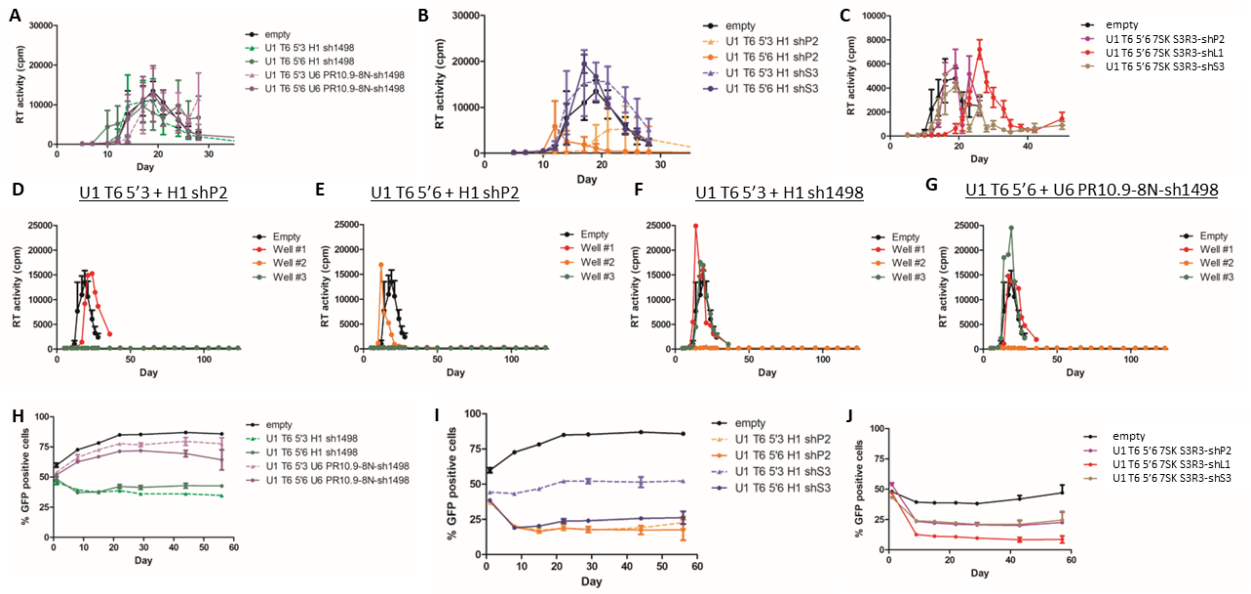

Figure 6.2 HIV-1 infection can be prevented in SupT1 cells expressing both U1 T6 5' end length variants and shRNAs. (A,B,C) SupT1 cells were transduced with HIV-7-EGFP lentiviral vectors expressing different combinations of U1 T6 5' end length variants with shRNAs or aptamer-shRNA conjugations and infected with HIV-1 NL4-3 at 1750 cpm/mL. The mean RT activity (cpm) was measured in culture supernatants at various days post infection. Each data point represents the mean +/- SEM from three infections (n=3). (D,E,F,G) Infections from four different combinations are plotted individually instead of a curve being generated from the average of three infections, at least one well from each of the four combinations does not support replication after challenge with HIV-1 NL4-3. (H, I, J) Transduced SupT1 cells were mixed with untransduced SupT1 cells and the percentage of GFP positive cells was measured at various days post-mixing. Each data point represents the mean +/- SEM from three experiments (n=3).

6.4 Discussion.

Although combination gene therapy has the potential to eliminate the need for cART, their effectiveness in the absence of cART has yet to be demonstrated[32,33]. Therefore, we have previously optimized anti-HIV RNAs by evaluating various expression strategies and molecular designs to maximize inhibitory capabilities while avoiding cytotoxicity[5] (Chapters III and IV). These investigations have shown that the U6 and 7SK promoter are more transcriptionally active than the H1 promoter, which causes shRNAs to be more potent when expressed from the U6 and 7SK promoters. We have also shown that this high rate of transcription can cause cytotoxicity in

the case of U6 and 7SK driven shRNAs, but that incorporating large aptamers within the terminal loop of the shRNAs effectively eliminates the adverse effects observed from canonical shRNAs.

While we have performed extensive work to optimize shRNAs for use in gene therapy, we also wanted to optimize the molecular design of anti-HIV-1 U1i RNAs. We have previously shown that extending the recognition domain of the U1 T6 can overcome a loss of antiviral effect resulting from mismatches in the target site[8]. While this is encouraging for its usage across multiple HIV-1 strains, we investigated here whether extensions of the recognition domain could improve its specificity. Our results within infection assays show that U1 T6 and all its length variants could inhibit HIV-1 replication compared to negative controls. However, only extensions of the recognition site at the 5' end led to further increases in inhibitory effects compared to the native U1 T6. In addition, the native U1 T6 as well as the 3' end length variants were associated with slight defects in cell growth within our competitive growth assays whereas no cytotoxicity was detected in cells expressing the two 5' end length variants. It is possible that the sequences of the native U1 T6 and its 3' end length extensions cause off-target binding of cellular RNAs that do not occur in the case of the 5' end length variants. Although our previous work did not show any growth defects in cells expressing the native U1 T6[8], variability in experimental manipulations could have led to the growth defects detected in our results as these defects were minimal. Since the 5' end length variants of the U1 T6 offered higher antiviral potency along with a lower potential to elicit cytotoxicity, these molecules are considered to be optimal for use with other anti-HIV-1 RNAs for the establishment of combination gene therapy.

We have taken together all of our previous optimization work and applied it towards establishing an effective combination gene therapy. We chose two different classes of functional anti-HIV-1 RNA molecules, the U1i RNAs and shRNAs, in order to decrease the risk of viral resistance

towards one specific class of antiviral RNAs. Using two different classes of anti-HIV-1 RNAs also decreases the likelihood of saturating any one endogenous cellular pathway. The U1 T6 5'3 and U1 T6 5'6 was chosen as these length variants provide maximal inhibition of virus replication and do not elicit cytotoxicity. The H1 promoter was chosen to express non-chimeric shRNAs, as expression from this promoter is less likely to induce cytotoxicity compared to when expression is driven by the U6 or 7SK promoters[5]. However, we chose the U6 and 7SK promoters to express the aptamer-shRNA chimeras because we previously observed that shRNA-mediated cytotoxicity from the U6 and 7SK promoter does not occur when the loop of shRNAs are replaced by large aptamers (Chapters III and IV).

When determining the inhibitory capabilities of U1 T6, U1 T6 3'3, U1 T6 3'6, U1 T6 5'3 and U1 T6 5'6 expressed alone, the replication kinetics were similar across each of the three individual infections that were put into place for the inhibitory assays of each molecule. This was not the case for some of our combination therapies where there was a large disparity in the replication kinetics across each infection sample in the assays. This was the case for the combinations U1 T6 5'3 with H1 driven shP2, U1 T6 5'6 with H1 driven shP2, U1 T6 5'3 with H1 driven sh1498, as well U1 T6 5'6 with U6 driven PR10.9-8N-sh1498, and we therefore decided to represent this data by plotting each infection individually since plotting the average of the three infections does not properly fit our data. The difference in replication kinetics between the individual infections for these combinations relates to whether replication is detected or not. The combinations U1 T6 5'3 with H1 driven shP2 and U1 T6 5'6 with H1 driven shP2 caused the greatest inhibition of viral replication where no replication was detected for two out of three samples while the combinations U1 T6 5'3 with H1 driven sh1498 and U1 T6 5'6 with U6 driven PR10.9-8N-sh1498 also inhibited viral replication but to a lesser extent, with no replication being detected for one out of three

samples (Figure 6.2C-F). Unexpectedly, the samples which do support viral replication have similar replication kinetics to cells transduced with the negative control empty vector. Within our previous work, antiviral RNAs were capable of delaying the time point at which viral replication occurred, as is seen when the U1 T6 and its length variants are expressed alone. One explanation to the difference in results with our previous work is that when one antiviral RNA is utilized, the virus may attain a threshold level of viral RNA production where some transcripts can escape the inhibitory actions of the therapeutic RNAs due to the overwhelming number of viral transcripts and then go on to complete the replication steps necessary to generate new virions. It is possible that employing two anti-HIV-1 RNAs increases this threshold level of viral RNA production for escape by overwhelming numbers to an unreachable point. If this is the case, the replication detected in some of the samples of the combinations U1 T6 5'3 with H1 driven shP2, U1 T6 5'6 with H1 driven shP2, U1 T6 5'3 with H1 driven sh1498 and U1 T6 5'6 with U6 driven PR10.9-8N-sh1498 may be a consequence of the development of viral resistance to the combination therapies. However, we have attempted to determine if this was the case by sequencing the target sites of the different combinations within the viral RNA of the samples that supported viral replication and did not find evidence of mutations within these various target sites (data not shown). While this data does not support our hypothesis for the development of resistance in samples which had detectable levels of viral replication, it is possible that mutations outside of the target site occurred which conferred viral resistance to the combination therapies. It will therefore be helpful to sequence the entire viral RNA of the samples which supported viral replication to assess whether resistance emerged from mutations in other genomic regions.

Growth defects are detected within competitive growth assays by comparing the percentage of GFP in cultures over time to the negative empty control which should be maintained at 50%.

However, the negative control did not remain at 50% in one of our competitive growth assays involving different combination therapies. Instead, the percentage of GFP increased over time which signals that the GFP negative cells were unhealthy following cell sorting. Due to this, the results from this assay are not completely reliable. However, the data within this competitive growth assay showed that only the combinations U1 T6 5'3 with U6 driven PR10.9-8N-sh1498 and U1 T6 5'6 with U6 driven PR10.9-8N-sh1498 had similar percentages of GFP positive cells through time as the negative control empty vector. Therefore, we suspect that all other combinations evaluated within this assay likely elicited at least minimal cytotoxic effects. Our second competitive growth assay did not have an issue with the viability of the GFP negative cells. While the cell cultures expressing each of the combinations within this assay had decreasing percentages of GFP, this decrease only occurred within the first week of the assay. It is possible that when establishing this assay, the GFP positive cells of the different combinations were stressed compared to the GFP negative cells following cell sorting. Therefore, an initial growth advantage would be conferred towards the GFP negative cells until the GFP positive cells were stable within the cell culture conditions which resulted in an initial decrease in the percentage of GFP in the overall culture.

This study allowed for the optimization of a U1i RNA by exploring extended lengths of the recognition domain and showed that optimized functional anti-HIV-1 RNAs are able to prevent the occurrence of HIV-1 replication within cell cultures. These results highlight the importance of properly optimizing each antiviral molecule individually to allow for the establishment of combination therapies which can prevent viral replication rather than only delaying its occurrence. However, this study also demonstrates that double molecule therapy is not sufficient in every case to completely prevent replication. It will be valuable to explore additional double or triple molecule

therapies with the goal of identifying an anti-HIV combination gene therapy which can completely suppress viral replication indefinitely.

6.5 Materials and methods.

6.5.1 Cell culture

HEK293T cells were grown in Dulbecco's modified Eagle's medium (HyClone, Logan, UT) supplemented with 10% fetal bovine serum (FBS, HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies, Carlsbad, CA, USA). SupT1 cells were grown in Roswell Park Memorial Institute Medium 1640 (HyClone), supplemented with 10% heat-inactivated FBS (HyClone), 50 μg/mL streptomycin, and 50 U/mL penicillin (Life Technologies).

6.5.2 Vector construction

A MCS containing the restriction sites Bsp1191, SgsI, PacI, MunI, SfaAI, and MluI was generated by annealing complementary oligonucleotides and included between the NotI and XbaI restriction sites previously present within the lentiviral transfer vector HIV-7-EGFP[34]. To include the U1i RNA genes within the MCS containing lentiviral transfer vector, both the U1 promoter and the U1i RNA gene was amplified by PCR out of the HIV-1 U1 snRNAs UBC plasmids generated in our previous study[8] using forward primer 5'-ATTAGCGGCCGCTTGCTCCTTACACAG-3' and reverse primer 5'-GATGAGGCGCGCCGCTTTACACTTTATG-3'. The amplified DNA fragments were then digested with NotI and SgsI to be ligated into NotI and SgsI-digested, MCS containing, lentiviral transfer vector. To establish double molecule combination gene therapies with our lentiviral vectors, both the associated promoter and the anti-HIV-1 gene of interest was amplified out of the psiRNA expression plasmid described previously (Chapters III and IV) using forward primer 5'-TATGGCGCGCCAGGGATTTTGGTCATGTTCTTAATCGATACTA-3'

and reverse primer 5'-GCGCAATTGGTTATGTAACGCCTGCAGGTTAATTAAGTCTAGA-3'. The double molecule combinations which include 7SK driven S3R3-shP2, S3R3-shS3 and S3R3-shL1 were established by amplifying both the 7SK promoter and antiviral gene out of the HIV-7-EGFP plasmid which did not contain an MCS using the forward primer 5'-TATGGCGCGCCAGGGATTTTGGTCATGTTCTTAATCGATACTA-3' and reverse primer 5'- GCGCAATTGCGAATTCCTGCAGCCCGGGGGATCCATCTCTAGA-3'. The various amplified DNA fragments were then digested with SgsI and MunI to be ligated into SgsI and MunI-digested lentiviral transfer vector already containing the MCS as well as the gene coding for either the U1 T6 5'3 or U1 T6 5'6.

6.5.3 Lentivirus production

HEK 293T cells were seeded 24 h prior to co-transfections in T75 flasks at 2.75 x 10^5 cells/mL within a volume of 20 mL. LVs were produced by co-transfecting 9 μg of the HIV-7-EGFP transfer vector, 3.4 μg plasmid expressing vesicular stomatitis virus G protein (from Dr. J. Rossi) and 10, 50 or 65 μg packaging plasmid psPAX2 (Addgene, number 12260) using PEI (3 μL/μg of DNA) (Polysciences). Supernatants were harvested 48 h after transfection, where lentiviral particles were isolated and concentrated using Lenti-X (Clontech) according to the manufacturer's protocol.

6.5.4 Infection and competitive growth

SupT1 cells were transduced with a serial dilution of lentivirus particles (range of 1 in 4 to 1 in 2048) in order to determine lentivirus titers by measuring the percentage of GFP-positive cells after transduction. Once the lentivirus titers were determined, SupT1 cells were transduced at an MOI of 1 using 8 μg/mL Polybrene (Sigma-Aldrich, St. Louis, MO, USA) to generate T cell lines stably expressing antiviral genes. Cells were sorted 72-96 h after transduction for GFP expression (Gating shown in Supplementary Figures S1-4) with a FACSAria Fusion cell sorter (BD

Biosciences, Franklin Lakes, NJ, USA). Sorted cells were seeded in 96-well round-bottom plates at 2 x 10^4 cells/well in a volume of 100 µL and challenged with HIV-1 NL4-3 (1750 cpm/mL, determined using the HIV-1 RT assay) 24 h post-sorting. The supernatant was periodically harvested and replaced with fresh media, where the RT activity within the collected supernatants was measured to establish the infection kinetics. Sorted cells were also seeded in 96-well round-bottom plates at 2 x 10^4 cells/well in a volume of 100 µL to be used within competitive growth assays. These cells were immediately mixed with an equal number of untransduced SupT1 cells that had been passed through the flow cytometer without sorting, bringing the total volume in each well to 200 µL with a concentration of 2 x 10^4 cells/well. The percentage of GFP-positive cells within the cultures was measured over time using a LSRFortessa flow cytometer (BD Biosciences).

6.5.5 RT assay

Infection kinetics were determined by measuring RT activity as previously described[35,36]. Briefly, 2.5 or 5 µL of supernatant was added to 12.5 or 25 µL of non-radioactive cocktail (60 mM TrisCl, 75 mM KCl, 5mM MgCl2, 1.04 mM EDTA, 1% NP-40) and subsequently incubated with 12.5 or 25 µL of RT cocktail [60 mM TrisCl, 75 mM KCl, 5mM MgCl2, 1.04 mM EDTA, 10 µg/mL polyA (Roche, Basel, Switzerland), 0.33 µg/mL oligo dT (Life Technologies), 8 65 mM DTT, 0.05 miC/mL α32P-TTP (3,000 Ci/mmol; Perkin Elmer, Waltham, MA, USA)] for 2 h at 37°C. A total of 5 µL of the radioactive solution was then spotted onto diethylaminoethyl (DEAE) filter mats (Perkin Elmer) or a positively charged nylon membrane (Hybond-N^+, Amersham Biosciences, Little Chalfont, UK) to detect the resulting poly dT RT product while unincorporated [^{32}P] dTTP were removed by performing five washes using 2x saline sodium citrate (SSC) buffer

(20× SSC buffer: 3 M NaCl and 0.3 M sodium citrate) along with two washes using 95% ethanol. Counts per minute (cpm) were measured using a microplate scintillation counter (MicroBeta TriLux; Perkin Elmer). The amount of HIV-1 RT enzyme in the supernatants is proportional to the cpm readout

6.6 References.

1. Ndung'u, T., McCune, J.M. & Deeks, S.G. Why and where an HIV cure is needed and how it might be achieved. *Nature* **576**, 397-405 (2019).
2. Goguen, R.P., Malard, C.M., Scarborough, R.J. & Gatignol, A. Small RNAs to treat human immunodeficiency virus type 1 infection by gene therapy. *Curr Opin Virol* **38**, 10-20 (2019).
3. Scarborough, R.J. & Gatignol, A. RNA Interference Therapies for an HIV-1 Functional Cure. *Viruses* **10**, 8 (2018).
4. Scarborough, R.J., *et al.* A Conserved Target Site in HIV-1 Gag RNA is Accessible to Inhibition by Both an HDV Ribozyme and a Short Hairpin RNA. *Mol Ther Nucleic Acids* **3**, e178 (2014).
5. Goguen, R.P., *et al.* Efficacy, accumulation, and transcriptional profile of anti-HIV shRNAs expressed from human U6, 7SK, and H1 promoters. *Mol Ther Nucleic Acids* **23**, 1020-1034 (2021).
6. Lange, M.J., Nguyen, P.D.M., Callaway, M.K., Johnson, M.C. & Burke, D.H. RNA-protein interactions govern antiviral specificity and encapsidation of broad spectrum anti-HIV reverse transcriptase aptamers. *Nucleic Acids Res* **45**, 6087-6097 (2017).
7. Pang, K.M., Castanotto, D., Li, H., Scherer, L. & Rossi, J.J. Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. *Nucleic Acids Res* **46**, e6 (2018).
8. Del Corpo, O., *et al.* A U1i RNA that Enhances HIV-1 RNA Splicing with an Elongated Recognition Domain Is an Optimal Candidate for Combination HIV-1 Gene Therapy. *Mol Ther Nucleic Acids* **18**, 815-830 (2019).
9. Le Douce, V., *et al.* Achieving a cure for HIV infection: do we have reasons to be optimistic? *J Antimicrob Chemother* **67**, 1063-1074 (2012).
10. Gray, G.E., Laher, F., Lazarus, E., Ensoli, B. & Corey, L. Approaches to preventative and therapeutic HIV vaccines. *Curr Opin Virol* **17**, 104-109 (2016).
11. Scarborough, R.J., Goguen, R.P. & Gatignol, A. A second patient cured of HIV infection: hopes and limitations. *Virologie (Montrouge)* **23**, 1-4 (2019).
12. Hutter, G., *et al.* Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* **360**, 692-698 (2009).
13. Gupta, R.K., *et al.* HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* **568**, 244-248 (2019).

14. Jensen, B.O., *et al.* In-depth virological and immunological characterization of HIV-1 cure after CCR5Delta32/Delta32 allogeneic hematopoietic stem cell transplantation. *Nature medicine* **29**, 583-587 (2023).
15. Hsu, J.M., *et al.* HIV-1 remission with CCR5Δ32Δ32 haplo-cord transplant in a U.S. woman: IMPAACT P1107. [Abstract 65]. In Special Issue: Abstracts from the 2022 Conference on Retroviruses and Opportunistic Infections. *Top Antiv Med* **30**, 23 (2022).
16. Dickter, J., *et al.* The 'City of Hope' Patient: prolonged HIV-1 remission without antiretrovirals (ART) after allogeneic hematopoietic stem cell transplantation (aHCT) of CCR5Δ32/Δ32 donor cells for acute myelogenous leukemia (AML). AIDS 2022 Abstract. https://programme.aids2022.org/Abstract/Abstract/?abstractid=12508. (2022).
17. Novembre, J., Galvani, A.P. & Slatkin, M. The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol* **3**, e339 (2005).
18. Hutter, G. More on shift of HIV tropism in stem-cell transplantation with CCR5 delta32/delta32 mutation. *N Engl J Med* **371**, 2437-2438 (2014).
19. Scarborough, R.J., Adams, K.L., Daher, A. & Gatignol, A. Effective inhibition of HIV-1 production by short hairpin RNAs and small interfering RNAs targeting a highly conserved site in HIV-1 Gag RNA is optimized by evaluating alternative length formats. *Antimicrob Agents Chemother* **59**, 5297-5305 (2015).
20. Mandal, D., Feng, Z. & Stoltzfus, C.M. Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J Virol* **84**, 12790-12800 (2010).
21. Knoepfel, S.A., Abad, A., Abad, X., Fortes, P. & Berkhout, B. Design of modified U1i molecules against HIV-1 RNA. *Antiviral Res* **94**, 208-216 (2012).
22. Nimjee, S.M., White, R.R., Becker, R.C. & Sullenger, B.A. Aptamers as Therapeutics. *Annu Rev Pharmacol Toxicol* **57**, 61-79 (2017).
23. Gulick, R.M., *et al.* Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *N Engl J Med* **337**, 734-739 (1997).
24. Hammer, S.M., *et al.* A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* **337**, 725-733 (1997).
25. Das, A.T., *et al.* Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J Virol* **78**, 2601-2605 (2004).
26. Westerhout, E.M., Ooms, M., Vink, M., Das, A.T. & Berkhout, B. HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res* **33**, 796-804 (2005).
27. Boden, D., Pusch, O., Lee, F., Tucker, L. & Ramratnam, B. Human immunodeficiency virus type 1 escape from RNA interference. *J Virol* **77**, 11531-11535 (2003).
28. ter Brake, O., Konstantinova, P., Ceylan, M. & Berkhout, B. Silencing of HIV-1 with RNA interference: a multiple shRNA approach. *Mol Ther* **14**, 883-892 (2006).
29. Chung, J., *et al.* Optimized lentiviral vectors for HIV gene therapy: multiplexed expression of small RNAs and inclusion of MGMT(P140K) drug resistance gene. *Mol Ther* **22**, 952-963 (2014).
30. McIntyre, G.J., *et al.* 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* **6**, 55 (2009).

31. Low, J.T., *et al.* SHAPE-directed discovery of potent shRNA inhibitors of HIV-1. *Mol Ther* **20**, 820-828 (2012).
32. DiGiusto, D.L., *et al.* RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**, 36ra43 (2010).
33. Wolstein, O., *et al.* Preclinical safety and efficacy of an anti-HIV-1 lentiviral vector containing a short hairpin RNA to CCR5 and the C46 fusion inhibitor. *Mol Ther Methods Clin Dev* **1**, 11 (2014).
34. Yam, P.Y., *et al.* Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol Ther* **5**, 479-484 (2002).
35. Scarborough, R.J., Levesque, M.V., Perreault, J.P. & Gatignol, A. Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* **1103**, 31-43 (2014).
36. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. & Gatignol, A. Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp* **115**, e54486 (2016).