Characterization and optimization of U1 interference RNAs for use in gene therapy for HIV

Presented by

Olivier Del Corpo

Division of Experimental Medicine
Department of Medicine

McGill University

Montreal, Quebec, Canada

August 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

© Olivier Del Corpo, 2018

Table of Contents

Acknowledgments:

I would like to thank Dr. Anne Gatignol for providing me the opportunity to work on this project and for her constant guidance, support and assistance. Most of all, I would like to thank her for allowing me to defer my masters to begin my first year of medical school. Working in her lab has taught me more than I can imagine and has helped me develop a critical thought process which I will use throughout my career. I would like to specifically thank Dr. Robert Scarborough for teaching me and constantly helping me throughout my master's project and Aïcha Daher for her expertise, guidance and help with trouble shooting my experiments. Without their support, I doubt I would have been able to complete many of my experiments. I would also like to thank all the other members of the lab for not only being great colleagues but amazing friends. I wish them nothing but the best in their current and future endeavors.

I would like to thank my family for always being there and supporting me throughout my education. Specifically, my mother for constantly pushing me to better myself. Without her, I would not be where I am today.

List of Tables:

Manuscript:

List of Figures:

Introduction:

Manuscript:

Additional Experiments:

List of Abbreviations:

3'ss	3' Splice acceptor site
5'ss	5' Splice donor site
Ago-2	Argonaute 2
AIDS	Acquired immunodeficiency syndrome
APOBEC3G	Apolipoprotein b mRNA editing enzyme catalytic subunit 3G
ARV	Antiretroviral
ASF1/SF2	Alternative splicing factor 1/pre-mRNA-splicing factor
BLAST	basic local alignment search tool
BPS	Branch point sequence
CA	Capsid
cART	Combination antiretroviral therapy
CCR5	C-C chemokine receptor type 5
CD	Cluster of differentiation
CDC	Center for Disease Control and Prevention
CDK9	Cyclin-dependent kinase 9
cDNA	Complementary DNA
CF	Cystic fibrosis
cPPuT	Central polypurine tract
CRF	Circulating recombinant form
CS	Completely-spliced
CXCR4	C-X-C chemokine receptor type 4
Ds	Double stranded
EGFP	Enhanced green fluorescent protein

ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
FL	Full-length
GALT	Gut-associated lymphatic tissue
Gp120	Surface glycoprotein 120
Gp41	Transmembrane glycoprotein 41
HIV	Human immunodeficiency virus
HnRNP	Heterogeneous ribonuclear protein
HSC	Hematopoietic stem cell
HTLV	Human T-lymphotropic virus
IN	Integrase
IS	Incompletely-spliced
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
Kb	Kilobase
LAV	Lymphadenopathy-associated virus
LTR	Long terminal repeats
MA	Matrix
mBBP/SF1	Branchpoint binding protein
MHC	Major histocompatibility complex
NC	Nucleocapsid
Nef	Negative regulatory factor
ORF	Open reading frame
PAP	Poly-A polymerase
PAS	Polyadenylation site

PBS	Primer binding site
PIC	Pre-integration complex
PPuT	Polypurine tract
PPyT	Polypyrimidine tract
PR	Protease
pTEFb	Positive transcription elongation factor b
PTGS	Post transcriptional gene silencing
RISC	RNA-induced silencing complex
RNAi	RNA interference
RRE	Rev response element
RS	Arginine/Serine
RT	Reverse transcription
RTC	Reverse transcription complex
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SIV	Simian immunodeficiency virus
SL	Stem loop
Sm	Smith
SMA	Spinal muscular atrophy
snoTAR	U16 small nucleolar RNA conjugated TAR decoy RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein
SR protein	Serine-arginine rich protein
SS	Single-stranded
TAR	Trans-activation response element

TAT	Trans-activator of transcription
TRBP	TAR RNA binding protein
$tRNA^{Lys3}$	Transfer RNA lysine 3
U1-WT	Wild-type U1 snRNA expressing construct
U1i RNAs	U1 interfering RNAs
U2AF	Heterodimeric U2 auxiliary factor
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U
WT	Wild-Type

Abstract:

Human Immunodeficiency virus (HIV) affects 30.8 to 42.9 million people and remains one of the top 10 causes of death worldwide. Although the use of antiretrovirals has improved the life expectancy of those infected with HIV, it is likely never to be curative due to its inability to eliminate the virus. Furthermore, the high cost of antiretroviral therapy and well-characterized side effects highlights the need for an alternative therapy or even a cure for HIV. One approach being considered to cure HIV infection is the use of gene therapy to make all HIV target cells resistant to HIV replication. Several anti-HIV RNAs have been proposed for use in gene therapy including short hairpin RNAs, decoy RNAs, ribozymes and U1 interfering RNAs (U1i RNAs). Similar to current antiretroviral therapy, antiviral gene therapy will likely require a combinatorial and multiplexing approach to reduce the likelihood of viral resistance and avoid the saturation of any cellular processing pathways. Therefore, there is a constant need to identify and characterize novel antiviral RNAs.

U1i RNA represent attractive candidate for use in HIV-1 gene therapy and several have been designed to target HIV-1 RNA through splicing enhancement or inhibition of mRNA polyadenylation. However, there is a lack of data comparing U1i RNAs that enhance splicing of viral RNA to those that inhibit polyadenylation. To identify optimal U1i RNA antiviral candidates, we constructed and compared the antiviral efficacy of top U1i RNAs identified in previous studies and acting through either mechanism. We also designed and constructed new U1i RNAs targeting the Gag open reading frame of HIV-1 RNA. Here, we show for the first time that U1i RNAs targeting regions upstream of the first splice acceptor site have the potential to inhibit HIV-1 replication through a previously uncharacterized mechanism. We also demonstrate that U1i RNAs that enhance HIV-1 RNA splicing are more effective inhibitors of

viral replication when compared to those that inhibit polyadenylation. Preliminary toxicity studies indicate that U1i RNAs do not induce toxicity at high doses in cultured cells. Lastly, we found that the recognition domain of U1i RNAs can be increased by 6 nucleotides with little effect on their ability to inhibit HIV-1 replication. Overall, our results suggest that U1i RNAs that act through mechanisms other than inhibition of polyadenylation are competitive candidates for use in combination gene therapy with antiviral RNAs targeting HIV-1.

Résumé:

Le virus de l'immunodéficience humaine (VIH) touche de 30,8 à 42,9 millions de personnes et demeure l'une des 10 principales causes de décès dans le monde. Bien que l'utilisation des traitements antirétroviraux ait amélioré l'espérance de vie des personnes infectées par le VIH, elle ne permet pas de guérir complètement de l'infection par le VIH. De plus, le coût élevé ainsi que les effets indésirables bien caractérisés reliés à la thérapie antirétrovirale soulignent la nécessité de trouver une thérapie alternative ou même un traitement de guérison contre le VIH. Une approche envisagée pour guérir l'infection par le VIH est l'utilisation de la thérapie génique pour rendre toutes les cellules cibles du VIH résistantes à la réplication virale. Plusieurs ARN anti-VIH ont été proposés à cet effet, notamment les ARN courts en épingle à cheveux, les ARN leurres, les ribozymes et les ARN interférents U1 (ARN U1i). De la même façon que la thérapie antirétrovirale actuelle, la thérapie génique antivirale nécessitera probablement une approche combinatoire multiple pour réduire la probabilité de résistance virale et éviter la saturation de toute voie cellulaire. Par conséquent, il existe un besoin constant d'identifier et de caractériser de nouveaux ARN antiviraux.

L'ARN U1i est un candidat très prometteur pour une utilisation en thérapie génique contre le VIH-1 et plusieurs ont été conçus pour cibler l'ARN du VIH-1 via l'augmentation de l'épissage ou l'inhibition de la polyadénylation de l'ARNm. Cependant, il y a un manque de données de comparaison des ARN U1i qui augmentent l'épissage de l'ARN viral à ceux qui inhibent la polyadénylation. Dans le but d'identifier les candidats ARN U1i antiviraux optimaux, nous avons construit et comparé l'efficacité antivirale des principaux ARN U1i identifiés dans des études précédentes agissant par l'intermédiaire de l'un ou l'autre de ces mécanismes. En plus, nous avons conçu et construit de nouveaux ARN U1i ciblant le cadre de lecture ouvert de

Gag. Pour la première fois, nous avons montré que les ARNs U1i ciblant des régions en amont du premier site accepteur d'épissage ont le potentiel d'inhiber la réplication du VIH-1 par un mécanisme non caractérisé auparavant. Par ailleurs, les ARN U1i qui augmentent l'épissage de l'ARN du VIH-1 sont des inhibiteurs plus efficaces de la réplication virale que ceux qui inhibent la polyadénylation. Des études de toxicité préliminaires indiquent que les ARN U1i n'induisent pas de toxicité à fortes doses dans les cellules en culture. Enfin, nous avons constaté que le domaine de reconnaissance des ARN U1i pouvait être augmenté de 6 nucléotides, avec la même capacité à inhiber la réplication du VIH-1. Dans l'ensemble, nos résultats suggèrent que les ARN U1i peuvent être ajoutés à l'armée croissante d'ARN antiviraux ciblant le VIH-1.

Chapter 1: Literature Review

1.1 Human Immunodeficiency virus

According to the World Health Organization (WHO)'s global health observatory data (www.who.int), an estimated 30.8 to 42.9 million people are currently infected with human immunodeficiency virus (HIV) worldwide and 1.6 to 2.1 million new infections occurred in 2016. Having claimed over 35 million lives to date, HIV remains a major public health issue. In most cases, infection with HIV is characterized by a week-long acute influenza or mononucleosis-like illness within a month or two after viral infection. Following infection, there is an asymptomatic phase that can last several years and eventually, if untreated, leads to the development of acquired immunodeficiency syndrome (AIDS) due to the scarce amounts of peripheral cluster of differentiation 4^+ ($CD4^+$) T lymphocytes.

1.1.1 *Discovery, history and distribution*

The term AIDS was first coined in September of 1981 by the Center for Disease Control (CDC) to describe a set of mononucleosis-like symptoms seen among a group of homosexual men found in New York City and California (1). Some of these patients went on to develop rare diseases typically found in immunosuppressed individuals, including *Pneumocystis carinni* pneumonia and an atypical cancer called Kaposi's sarcoma (2,3). The first evidence that AIDS was caused by a sexually transmitted infectious agent arose soon after in 1983, when the female partners of previously diagnosed men were found to have AIDS (4). In May of 1983, Barré-Sinoussi and colleagues isolated a retrovirus, which they named Lymphadenopathy associated virus (LAV), from a Caucasian male with symptoms suggestive of AIDS, providing the first hint at a viral etiology (5). It was not until 1984, that a causal link was established by the group of Gallo due to the high frequency of human T lymphocyte leukemia virus (HTLV)-III (termed LAV by Barré-Sinoussi) detected and isolated from patients with AIDS or at risk for AIDS (6).

Soon after, it was determined that HTLV-III and LAV were the same virus and in 1986 the International Committee on the Taxonomy of Viruses declared that the causative agent of AIDS will be called HIV (7). At the same time, Clavel et al. isolated a similar but new human retrovirus in West African patients diagnosed with AIDS which was later termed HIV-2 (8).

HIV-1 and HIV-2 originated through multiple zoonotic transmissions of primate viruses known as simian immunodeficiency viruses (SIV) in West and Central Africa. There are over 40 different species of SIV each of which infect different non-human primates (9). The cross-species transmission is thought to be caused by the hunting and butchering of primates for bushmeat and the keeping of monkeys as pets (10). HIV-1 is believed to be derived from SIVs that infect chimpanzees and gorillas, while HIV-2 appears to be derived from SIVs that infects sooty mangabeys (11).

Independent zoonotic transmission events have generated several HIV lineages. Within the HIV-1 species there are four known lineage groups: M (Major), N (Non-M, non-O), O (Outlier) and P (Figure 1.1). Group M and N originate directly and independently from chimpanzee SIV found in West-Central Africa. HIV-1 group N infection is confined to a small number of individuals living in Cameroon, while HIV-1 group M is responsible for the majority of infections worldwide. HIV-1 group M can be further divided into genetically distinct subtypes A-K. Additionally, coinfection with these group M subtypes has led to the formation of circulating recombinant forms (CRFs) of group M HIV-1. To date, 96 CRFs have been isolated (HIV Los Alamos National Laboratory Database). While the majority of HIV-1 groups cause

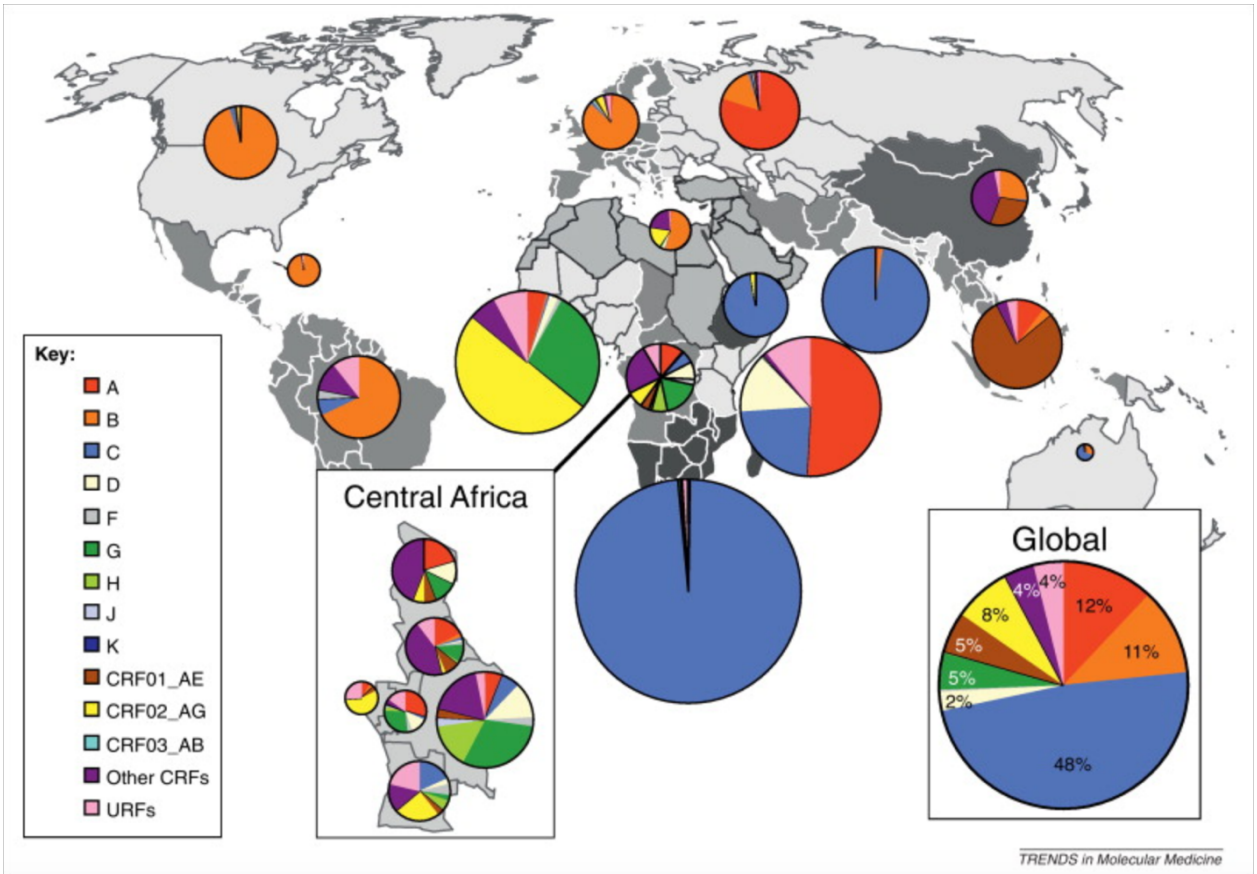

Figure 1.1 Global prevalence and distribution of HIV-1 Group M subtypes. Worldwide distribution of the various HIV-1 group M subtypes (A-K) and circulating recombinant forms (CRF) from 2004-2007 (12).

local infection limited to specific geographic areas, group M spread south via the Congo river to Kinshasa where the global HIV-1 epidemic likely initiated (13). Group O and group P are related to gorilla SIV identified in lowland gorillas living in southern Cameroon. Group O infection represents less than 1% of global HIV-1 infections and has been geographically limited to countries surrounding Cameroon. Lastly, group P was discovered in 2009 and so far, has been identified in only two patients from Yaounde, Cameroon (12).

Since its discovery, HIV-2 has rarely been seen outside of West Africa. Furthermore, overall prevalence rates have been declining, and in most areas HIV-2 is being replaced by HIV-1 (14). The lower viral loads and the lack of mother to infant transmission seen in HIV-2 infection explains the lower transmission rates when compared to HIV-1. Moreover, infection with HIV-2 rarely leads to the development of AIDS. Analogous to HIV-1, there are 8 distinct subtypes of HIV-2 termed groups A-H with group A and B representing the majority of infections (15).

It is important to note that while HIV-1 group M subtype B represents the majority of HIV-1 infections in the Americas and Europe, subtype C is the most prevalent worldwide (Figure 1.1). The majority of research and therapeutic development have focused on subtype B including the research presented in this thesis. However, inter-subtype genetic variations have been characterized to affect the pathogenesis, transmission and clinical management of these different subtypes (16). Therefore, it is imperative that future studies will include research on subtype C and other prominent subtypes or CRFs.

1.1.2 *Acute infection*

Based on the levels of plasma viral RNA and peripheral blood $CD4^+$ T lymphocyte count, HIV-1 infection can be divided into three distinct phases: acute primary infection, chronic infection and AIDS.

HIV-1 infection can occur through several modes of transmission. These include sexual contact, transmission via drug injection, blood transfusions, and transmission to a fetus or infant from an infected mother. However, transmission rates for HIV-1 remain relatively low, ranging from 10% to 0.1% depending on the type of exposure (17). Upon infection, HIV disseminates to gut associated lymphatic tissue (GALT) where it infects $CD4^+$ T lymphocytes. Throughout acute and chronic infection, GALT is considered to be a main site of viral replication. Furthermore, HIV-1 spreads to various different tissues including peripheral lymph nodes, spleen, bone marrow, nervous system and the urogenital tract (18). Following infection, there is a 2-4-week asymptomatic incubation phase which is accompanied by a significant increase in viral load that is reflected in the amount of plasma viral RNA. Increased viremia results in a rapid depletion of $CD4^+$ T lymphocytes and is often accompanied by a short symptomatic phase marked by a flu or mononucleosis-like illness and/or general malaise. Depletion of $CD4^+$ T lymphocytes is mainly due to the action of cytotoxic HIV-1 specific $CD8^+$ T lymphocytes and may be responsible for the initial drop in plasma viral levels seen following peak viral loads (19). During this acute phase, HIV-1 establishes latent reservoirs in memory $CD4^+$ T lymphocytes and possibly other reservoir cells or anatomical reservoirs. Once HIV DNA is integrated into the host chromatin of reservoir cells, it remains transcriptionally repressed and evades current drug therapies. However, it can be activated to reseed the infection once drug therapy is stopped, and represents the major barrier to HIV-1 elimination (20).

1.1.3 *Chronic infection*

Following acute infection, plasma viral load decreases significantly reaching a viral set point and the $CD4^+$ T lymphocytes count returns to near normal levels. The manifestation of a viral setpoint marks the start of the chronic phase of infection. During this phase, HIV-1 causes a slow depletion of circulating and tissue resident $CD4^+$ T lymphocytes, much of which are located in the GALT. T lymphocyte depletion of the GALT allows for systemic translocation of bacterial products and leads to dramatic and sustained immune activation. The chronic immune activation seen in HIV-1 infection mediated by viral antigens and gastrointestinal microbial products results in leukocyte dysregulation and a general increase in the cell turnover rates of $CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, B-cells and natural killer cells. Ultimately, the diverse immune cell populations progressively lose their potential to regenerate, leading to a breakdown of the immune system. Overall, both direct cytopathic and indirect immune mediated effects of viral replication contribute to the depletion of $CD4^+$ T lymphocytes and ultimately results in the development of AIDS (21).

1.1.4 *AIDS*

An HIV infected individual is considered to have AIDS once the peripheral $CD4^+$ T lymphocyte count decreases bellow 200 cells/μL. The lack of adaptive immunity, marked by a loss of $CD4^+$ T lymphocytes, results in a collapse of the immune system making infected individuals more prone to opportunistic infections as well as certain types of AIDS-related cancers (22,23). In certain individuals, the onset of AIDS is preceded by a switch of viral coreceptor use from CCR5 to CXCR4. While CXCR4-using viruses are rarely transmitted, transition of CCR5 to CXCR4 usage is generally correlated with rapidly progressing disease (24).

1.2 HIV-1 Replication Cycle

1.2.1 *Genome and viral structure*

HIV-1 belongs to the lentivirus genus of the retroviridae family. As with all retroviruses, it has the ability to reverse transcribe its single-stranded (SS) positive sense RNA genome into double-stranded (ds) DNA and subsequently integrate its dsDNA genome into the DNA of the infected host cell. Proviral integration allows for HIV-1 to hide from the hosts immune system and be maintained in a dormant/latent state (25).

The HIV genome codes for three major genes *gag, pol* and *env* which are flanked by long terminal repeats (LTRs) containing the viral promoter. Once transcribed, these major polyproteins are subsequently proteolyzed to give rise to individual proteins. The *Gag* open reading frame (ORF) encodes for the major structural proteins capsid (CA), matrix (MA) and nucleocapsid (NC). *Pol* encodes the enzymes reverse transcriptase (RT), protease (PR) and integrase (IN), while *Env* encodes for surface glycoprotein 120 (gp120) and transmembrane glycoprotein 41 (gp41). In addition, HIV-1 has six genes coding for auxiliary and regulatory protein (*tat, rev, vpr, vpu, nef* and *vif*) (Figure 1.2 B).

Each virion contains two copies of full-length (FL) genomic RNA and is enveloped by a host derived lipid membrane during budding. On the surface of the viral envelop there are heterotrimers of gp120 and gp41 used for viral entry along with host cellular proteins. Lining the lipid membrane is a matrix shell composed of the MA protein p17 and surrounding the genomic RNA is an icosahedral core composed of CA protein p24. Within the core capsid are two strands of dimerized FL genomic RNA that are coated by NC protein P7. Infectious virions also contain proteins necessary for reverse transcription, RT, and integration, IN. In addition, auxiliary

proteins Negative regulatory factor (Nef), Viral infectivity factor (Vif) and Viral protein R (Vpr) are packaged into budding viruses (26) (Figure 1.2 A).

1.2.2 *Viral entry, reverse transcription and integration*

Entry: The initial step of the HIV-1 replication cycle is viral binding followed by entry into targeted host cells. Adhesion to the targeted host cell can be mediated by viral Env proteins or host cell membrane proteins incorporated in the viral envelope during budding. Initial attachment can be nonspecifically mediated by Env interacting with negatively charged cell surface heparan sulfate proteoglycans or more specifically via Env interacting with α4β4 integrin or DC-SIGN (27,28). Attachment by any of these mechanisms brings Env into close proximity to the viral receptor CD4 and coreceptor C-C chemokine receptor type 5 (CCR5) or C-X-C chemokine receptor type 4 (CXCR4) (29). The Env glycoprotein is present on the virion surface as heterotrimers of gp41/gp120. The gp120 subunit is responsible for receptor binding and contains five conserved domains and five variable loops. Interaction of viral Env with CD4 results in conformational changes in both CD4 and gp120, exposing the coreceptor binding site on gp120. Tropism for viral coreceptors CCR5 and CXCR4 is determined by the third variable loop in the gp120 protein (30). Coreceptor binding leads to the dissociation of gp120 from gp41 allowing for the insertion of gp41 subunits N-terminal domain into the target membrane. Gp41 undergoes several conformational changes resulting in the formation of a six helix bundle that provides the energy for fusion of the viral and host cell membranes allowing for the release of viral capsid into the cytoplasm (31). A summary of the steps of the HIV-1 replication cycle is shown in Figure 1.3.

Reverse transcription: Immediately after the release of the viral capsid into the host cell cytoplasm, the capsid core undergoes a process of uncoating, leading to the generation of

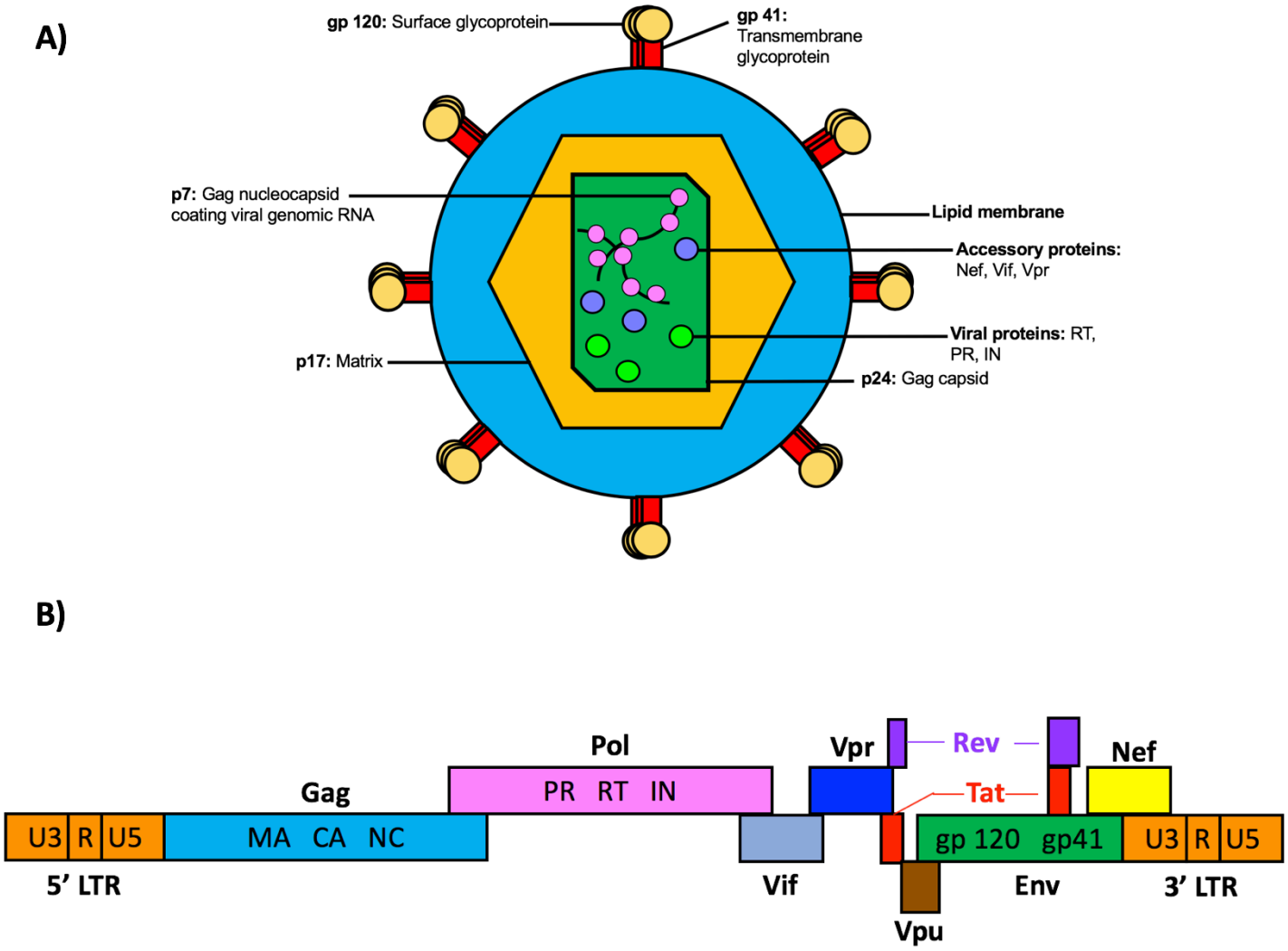

Figure 1.2 Structure of the HIV-1 virus and organization of viral genomic DNA. A) Schematic of the architecture of an infectious HIV-1 virus with viral proteins labeled. B) Outline of full-length genomic viral RNA. Based on information from (26).

reverse-transcription complexes (RTCs) followed by pre-integration complexes (PICs) (32). The distinction between the protein components of RTCs and PICs is unclear; however, reverse transcription is ongoing in RTCs, while PICs are defined as integration competent complexes. Reverse transcription is initiated in the host cell when human transfer RNA lysine 3 ($tRNA^{Lys3}$), which was initially incorporated into infectious viral particles from the previous host cell, binds to the primer binding site (PBS) approximately 180 nucleotides from the 5' end of the viral genomic RNA. $tRNA^{Lys3}$ acts as a primer for the initiation of minus strand DNA synthesis towards the 5' end of viral RNA (33). The RNA of this DNA-RNA duplex created by RT is subsequently degraded by the RNase H domain of RT exposing the newly synthesized minus strand DNA. Since the ends of viral genomic RNA contain direct repeats called R in the LTR, the minus stand DNA is transferred to the 3' end of the viral RNA. Infectious HIV-1 particles contain two copies of viral RNA and transfer can involve the 3' R sequence of either copy of viral RNA (34). After the transfer, minus strand DNA synthesis continues along the FL viral RNA resulting in the synthesis of the FL minus strand DNA. At the same time, RNase H degrades RNA that is part of the DNA-RNA duplex created by RT. The purine-rich sequence called the polypurine tract (PPuT), at the 3' end of the viral genome, and the central PPuT (cPPuT), are relatively resistant to RNase H cleavage and serve as the primers for the synthesis of plus strand DNA by DNA-dependent DNA synthesis using the newly synthesized minus strand DNA as a template (35).

Plus strand synthesis continues copying the minus strand DNA and the first few nucleotides of the $tRNA^{Lys3}$ primer. A second transfer or plus strand transfer occurs after the PBS site is reverse transcribed into minus strand DNA. The first few nucleotides of $tRNA^{Lys3}$ in the 3' end of the plus strand DNA is complementary to the PBS site in the minus strand DNA. These complementary sequences anneal and allow for synthesis of the 5' end of plus strand DNA (36).

A third DNA strand is also formed between the cPPuT and central termination site which is displaced by the elongation of the 5' end of the plus strand DNA, resulting in a discontinuous positive strand. This plus strand overlap or central "DNA flap", has been demonstrated to be an important *cis*-determinant of HIV DNA nuclear import of the PIC (37). Preceding integration, the DNA flap is removed by human flap endonuclease (FEN-1) and the positive strand is made continuous by host cell DNA ligases (38).

Integration: After reverse transcription, the LTRs of the viral dsDNA are processed by IN, producing terminal 3' hydroxyl groups at each LTR. 3' end processing results in the transition of the RCT to the PIC which consists of viral dsDNA, termed complementary DNA (cDNA), as well as viral proteins IN, MA, Vpr and RT. While the viral determinants of nuclear localization have not been fully elucidated, along with the DNA flap structure, these proteins are thought to mediate transport of the PIC from the cytoplasm into the nucleus via nuclear pore complexes (40). The ability of the PIC to enter the nucleus distinguishes lentiviruses from other retroviruses and allows HIV-1 to infect non-dividing cells. Once in the nucleus, IN cuts host cell chromosomal DNA in a staggered fashion and simultaneously joins the viral 3' hydroxyl ends to the 5' phosphates of target cut DNA. Host cellular repair enzymes then catalyze the integration of the unjoined 5' ends of the recombinant DNA intermediate resulting in an integrated provirus (41). Analysis of HIV-1 integration reveals that HIV-1 preferentially integrates within genes that are actively transcribed (42).

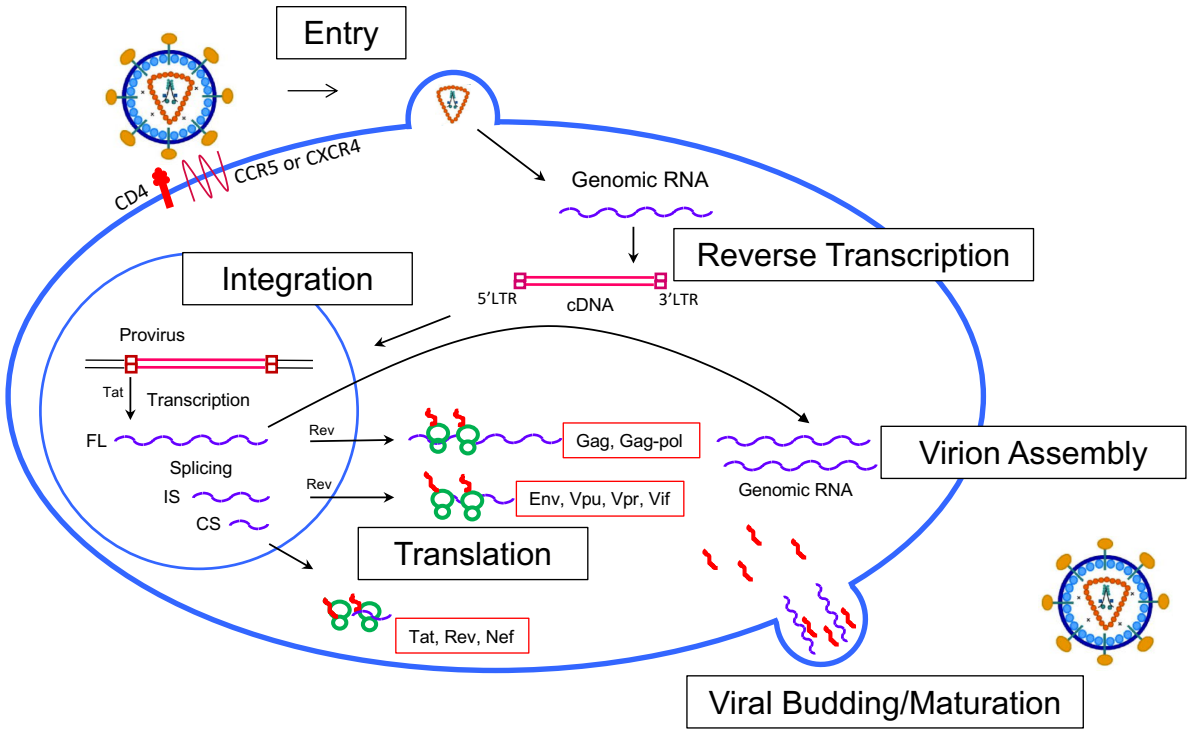

Figure 1.3 Summary of HIV-1 replication cycle. The basic steps of HIV replication are highlighted in the black boxes. FL, IS and CS represent full-length, incompletely spliced and completely spliced mRNA species, respectively. Modified from (39).

1.2.3 *Transcription, translation and viral assembly/budding*

Transcription: Once the viral genome is integrated into the host cell DNA, transcription is mediated by host cellular RNA polymerase II and is initiated at the +1 site located in the R region of the LTR. Early on in transcription, only small amounts of FL 9 kilobase pair (kb) RNA is produced as the provirus is located in nucleosome complexes. Although transcribed, FL RNA cannot be exported from the nucleus due to its size and association with nuclear factors. Instead, it is spliced into smaller 1.8 kb fragments (completely spliced, CS) that are exported to the cytoplasm and translated into Trans-Activator of Transcription (Tat), Rev and Nef proteins. Efficient proviral RNA transcription requires Tat. Once produced, Tat localizes to the nucleus where it complexes with positive transcription elongation factor b (pTEFb) and binds to the trans-activation response element (TAR) stem loop located at the 5' end of each newly transcribed HIV-1 RNA. Cyclin-dependent kinase 9 (CDK9), a subunit of pTEFb, phosphorylates the C-terminal domain of RNA polymerase II allowing for more efficient transcription of the provirus (43).

In the late phase of RNA transcription, the Rev protein localizes to the nucleus and binds to the Rev response element (RRE) present in the Env coding region of the viral incompletely-spliced (IS) and FL RNA. Rev mediates the export of these RNAs into the cytoplasm through the chromosome region maintenance 1/Exportin 1 pathway (44). IS RNAs encode for Vif, Vpr, Vpu, one exon Tat and Vpu/Env proteins, while unspliced FL RNA encodes for Gag and Gag-Pol polyproteins (45).

Translation: Translation of viral proteins occurs in the cytoplasm of the host cells and is mediated by free polyribosomes. FL RNA codes for Gag and Gag-Pol polyproteins; however, the ORFs of Gag and Gag-Pol overlap. Alternative translation of the Gag and Gag-Pol polyprotein is

due to a -1 ribosomal frameshift. Gag and Gag-Pol polyproteins are cleaved by viral PR into their individual protein components while the other polyprotein, Env, is cleaved by a host cell protease, Furin, to produce gp120 and gp41 (46,47).

Assembly and budding: After translation, unprocessed Gag polyprotein is targeted to the plasma membrane by domains in the MA protein. Once located at the plasma membrane, Gag promotes the incorporation of Env glycoproteins and mediates assembly of viral particles. The NC domain of Gag contains two zinc-finger-like domains that bind to HIV-1 FL RNA and mediates its recruitment to forming viral particles. Gag and Gag-Pol assemble into immature viral particles together with two copies of FL RNA. Eventually, these viral particles bud off of infected cells and mature through cleavage of Gag and Gag-Pol by PR forming the individual proteins MA, CA, NC, IN, PR and RT (48).

1.2.4 *Role of HIV-1 Auxiliary Proteins*

Additional proteins that are included in infectious HIV-1 particles are the auxiliary proteins Vif, Vpr and Nef. The other viral auxiliary protein viral protein U (Vpu) is not included in viral particles. While most of these auxiliary proteins are not directly involved in viral replication, they play crucial roles in the evasion of cell-mediated antiviral resistance and modify the cellular environment to ensure viral persistence along with replication.

Interaction with cell surface proteins is crucial for viral entry into targeted host cells; however, the adaptive immune response recognizes infected cells through viral epitope presentation by major histocompatibility complex (MHC) class 1 complexes and targets them for elimination. Both Nef and Vpu proteins modulate the expression of cell surface proteins resulting in evasion of the host immune responses. More specifically, Nef has the ability to decrease the expression of MHC class 1 and class 2 on infected cells (49,50). Furthermore, Nef regulates the

expression of CD4 on the surface of infected cells. Through interactions with the cytoplasmic domain of CD4, Nef targets surface CD4 for lysosomal degradation through the recruitment of clathrin coated pits (51). Vpu binds to CD4 complexed with viral protein Env located in the endoplasmic reticulum and induces proteasomal degradation of CD4 (52). Inhibition of CD4 allows for the release of viral particles from infected cells, which can be inhibited by interactions between cell surface CD4 and viral Env. Along with inhibition of CD4, one of the major roles of Vpu is to counteract tetherin, an interferon induced restriction factor that prevents the release of viral particles from the plasma membrane of infected cells (53).

The auxiliary protein Vif has an essential role during reverse transcription by preventing the cellular RNA-editing enzyme and viral restriction factor apolipoprotein B mRNA editing enzyme catalytic subunit 3G (APOBEC3G) from editing viral RNA. When uninhibited, APOBEC3G deaminates cytosine residues to uracil in minus strand reverse transcripts thereby inhibiting viral spread through loss of genetic integrity (54). Lastly, Vpr assists in transport of the PIC to the nucleus of infected cells and induces cell cycle arrest along with T lymphocyte apoptosis (55).

1.3 Regulation of HIV-1 Alternative RNA Splicing

Alternative RNA splicing of HIV-1 FL RNA is a complex process by which over 40 different spliced mRNAs are generated. Precise modulation of HIV-1 splicing is essential for efficient viral replication as an overproduction or inhibition of host splicing factors drastically affects HIV-1 replication (56). This section will summarize the main components of cellular splicing as well as the mechanism by which the various mRNA species are produced during HIV-1 alternative splicing. A summary of the location of splice sites, exons and splicing elements in the HIV-1 genome is provided in Figure 1.4.

1.3.1 *Eukaryotic splicing regulatory elements*

Eukaryotic genes consist of exons and introns that can be differentially spliced together allowing for a single gene to code for multiple proteins. Exons are sequences of code that will be included in the final mRNA product while introns are sequences that intervene between exons and are excluded from the mRNA products. Present in the majority of RNA introns are the core splicing signals, which include the 5' splice donor site (5'ss), the branchpoint sequence (BPS) and the 3' splice acceptor site (3'ss), containing a stretch of 15 to 40 pyrimidines called the polypyrimidine tract (PPyT). These signals are recognized by a large ribonucleoprotein complex known as the spliceosome, which mediates the successive phosphodiester transfer reactions involved in splicing (57). The spliceosome contains over 100 core proteins and 5 small nuclear ribonucleoproteins (snRNPs): snRNPs U1, U2, U3, U4, U5 and U6. The spliceosomal complex assembles onto an RNA transcript by sequential binding of its various components and each snRNP contains a single small nuclear RNA (snRNA) and over seven associated proteins. Initially, the 5'ss is recognized and bound by the U1 snRNP with the help of alternative splicing factor 1/pre-mRNA-splicing factor (ASF1/SF2). The PPyT is the binding site for heterodimeric U2 auxiliary factors (U2AF) and the BPS is bound by the branchpoint binding protein (mBBP/SF1) which is subsequently replaced by the U2 snRNP as spliceosome formation occurs. Together, binding of U2AF heterodimer and mBBP promote U2 snRNP binding to the 3'ss. U2 snRNP binding promotes the recruitment of U4/U5/U6 trimer leading to mature spliceosome formation and splicing of nascent mRNA transcripts (58). 5'ss are referred to as weak or strong depending on the extent of base pairing with the U1 snRNA. Likewise, 3'ss are referred to as weak or strong depending on the affinity for U2AF and mBBP. Therefore, the extent of homology to the consensus 5'ss and 3'ss dictates the strength of these core splicing signals (59,60).

In mammals, splice sites are predominantly recognized in pairs across exons through a process termed “exon definition”. In this process, spliceosomal factors recognize the 5’ss and the upstream 3’ss. Alternatively, intron definition occurs across the intron with the recognition of 5’ss and downstream 3’ss (61). In addition to core splicing signals, many RNA transcripts contain *cis*-regulatory elements that regulate exon definition. These elements include exonic splicing enhancers (ESEs), exonic splicing silencers (ESS), intronic splicing enhancers (ISE) and intronic splicing silencers (ISS). ESE and ESS promote or inhibit the inclusion of the exon they reside in while ISE and ISS enhance or inhibit the usage of the adjacent splice site or inclusion of the adjacent exon. ESEs function by recruiting members of the serine-arginine rich protein (SR protein) family. SR proteins bind ESEs via an N-terminal RNA recognition motif and interact with proteins that facilitate spliceosomal assembly through a C-terminal arginine/serine (RS) domain (62). On the other hand, ESSs are bound by splicing inhibitory proteins from the heterogeneous ribonuclear protein (hnRNP) family and repress exon definition. HnRNPs represent a group of diverse proteins containing one or more RNA binding domain and splice inhibitory domain and act through a variety of different mechanisms to inhibit splicing. ISE consist of sequences of DNA, the most well characterized is the G triplet (GGG) that enhance recognition of 5’ or 3’ss while ISS have been demonstrated to bind hnRNPs, more specifically hnRNP 1 (58).

1.3.2 *HIV-1 RNA splice sites and regulatory elements*

The production of HIV-1 mRNA requires the host cell splicing machinery to produce IS (4 kb) and CS (1.8 kb) mRNA. As previously mentioned, early in infection, only CS mRNAs are produced and exported to the cytoplasm. CS mRNAs encode for regulatory proteins Tat, Rev and Nef. Tat allows for increased viral mRNA transcription and Rev allows for transport of IS and FL mRNAs to the cytoplasm. HIV-1 RNA contains five 5’ss (D1-D5) and nine 3’ss resulting in

the production of over 40 different viral mRNAs (63). 5'ss D5 and 3'ss A6 are only present in a small number of clade B strains, including the strain HXB2, commonly used to study HIV-1 in the laboratory. The extent of splicing at each of the 3'ss depends on the intrinsic strength of the given splice site as well as the presence of ESE, ESS, ISE and ISS. To make matters more complicated, viral RNA encodes for two non-coding exons between the 3'ss A1 and 5'ss D2 (exon 2) and between the 3'ss A2 and 5'ss D3 (exon 3). IS and CS mRNA species can contain either one or both of these non-coding exons (56).

Several studies have been conducted to determine the intrinsic efficiency of HIV-1 splice sites. 5'ss D1 and D4 were determined to be strong splice sites while D2 and D3 were found to be relatively weak when compared to human consensus 5'ss (64). In addition, viral 3'ss were all shown to be weak 3'ss when compared to consensus 3'ss. ESE sequences were determined to be crucial for splicing at 3'ss A1, A4c, a, b, A5 and A7 while ESSs modulate splicing at 3'ss A2 and A3 (65). Apart from intrinsic efficiency of the various viral splice sites, the relative abundance of the various mRNA transcripts gives insight on the usage of viral 3'ss. The low abundance of mRNA encoding Vif, Vpr and Tat indicates that 3'ss A1, A2 and A3 are infrequently used while the high abundance of mRNA encoding Rev, Env/Vpu and Nef indicates the frequent usage of 3'ss A4c, b, a and A5. 3'ss A7 is also frequently used as over half of total spliced mRNA is CS (66).

1.3.3 *Splicing of viral mRNA:*

The formation of the various HIV-1 mRNAs are depicted in Figure 1.4 and reviewed in (67).

3' splice site A1: Vif mRNA is an IS mRNA formed by splicing 5'ss D1 to 3'ss A1. Along with Vif mRNA, CS and IS Vpr, Tat and Env/Vpu viral transcripts may contain the non-coding exon 2. Several experiments were conducted to determine the role of the 5'ss D2 on splicing

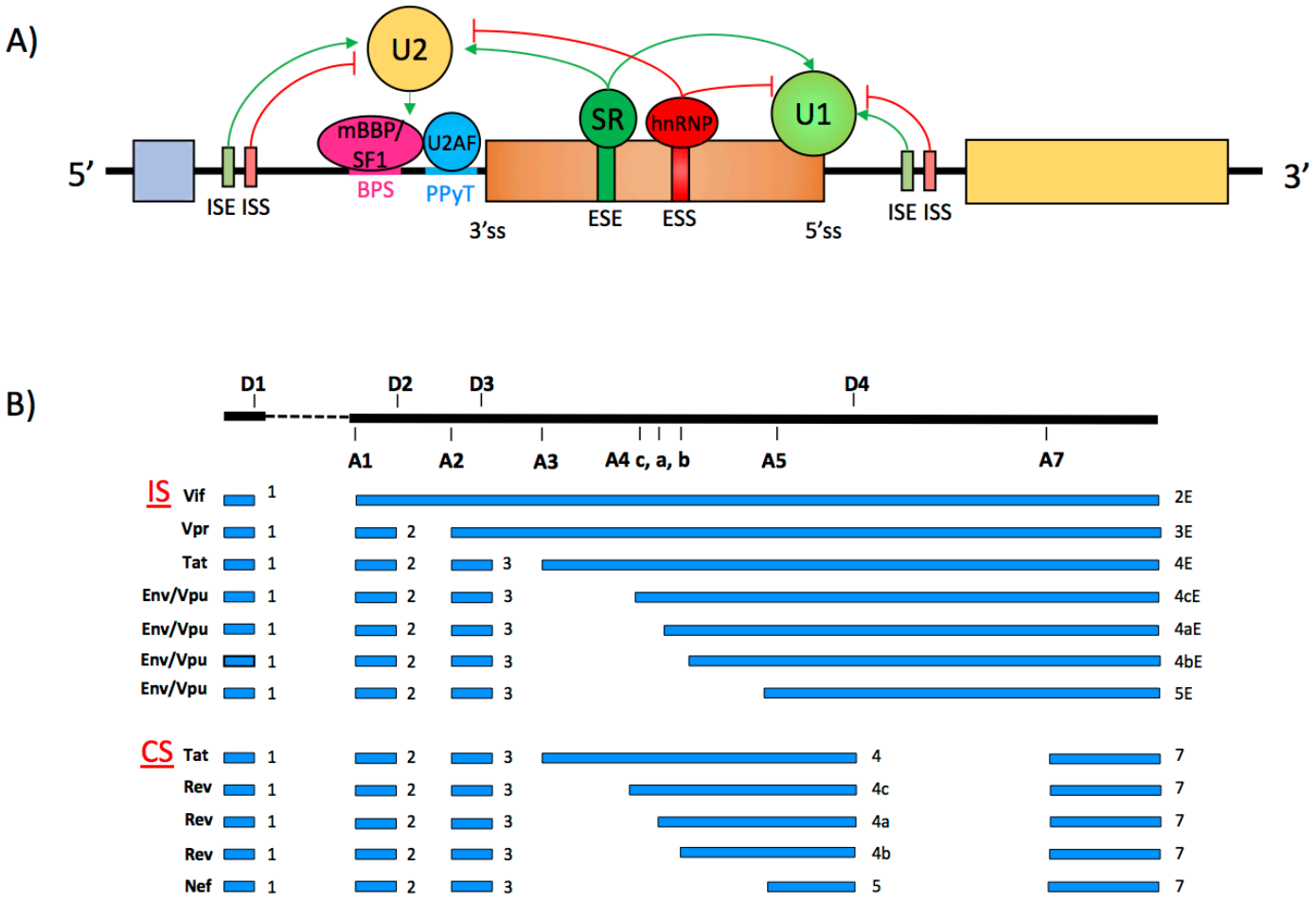

Figure 1.4 Mechanism of exon definition and HIV-1 mRNA splicing. A) The mechanism of exon definition. The core splicing signals 5' splice donor site (5'ss), 3' splice acceptor site (3'ss), branch point sequence (BPS) and the polypyrimidine tract (PPyT) are depicted in the figure along with the various splicing regulatory elements described in section 1.3.1. B) The location of the 5'ss (D) and 3'ss (A) along the HIV-1 genome. The various incompletely spliced (IS) and completely spliced (CS) mRNAs shown below with the contributing exons portrayed. Based on information from (59-61) and (63)

efficiency at the 3'ss A1. Mutant HIV-1 plasmids were constructed in which the homology of the 5'ss D2 to the U1 snRNP was increased or decreased. Mutations which increased affinity to the U1 snRNP resulted in an increase in the inclusion of exon 2 and increase levels of Vif mRNA and protein, while mutations that decreased homology of the 5'ss D2 to consensus 5'ss resulted in a decrease in both Vif mRNA and protein. These results indicate that the 5'ss D2 acts as an enhancer of splicing at 3'ss A1 by the process of exon definition (68). In addition to 5'ss D2, the inclusion of exon 2 is modulated by several ESEs and ESSs (65).

3' splice site A2 and A3: IS Vpr mRNAs are formed by the splicing of 5'ss D1 to 3'ss A2. Non-coding exon 2 and/or non-coding exon 3 can be included in a small amount of Vpr mRNA. Since exon 2 and exon 3 do not contain an AUG sequence, these exons become part of the 5'-leader of these viral transcripts. Splicing at the 3'ss A2 has been shown to be regulated by an ESS in exon 3 termed ESSV (69). Splicing of 5'ss D1, D2 or D3 to 3'ss A3 results in the production of CS two-exon Tat mRNAs or IS one-exon Tat mRNAs. Like Vpr viral transcripts, Tat mRNAs can include exon 2, exon 3 or both. Splicing at A3 is modulated by two ESS elements (ESS2 and ESS2p) and an ESE element (ESE2) (70).

3' splice site A4 c, a, b and A5: These four splice sites are used to produce CS Rev and Nef mRNAs and IS bicistronic Env/Vpu mRNAs. The mRNAs produced by A4c, a, b may include non-coding exon 2, 3, or both. Since the intrinsic strengths of these 3'ss are weak, the activation of these splice sites requires splicing enhancers. Namely, the strong 5'ss D4 and a guanosine-adenosine rich ESE found within exon 5 (defined by 3'ss A5 and 5'ss D4) (71). As for the 5'ss D2, mutational studies were conducted decreasing the affinity of the 5'ss D4 to the U1 snRNA and showed that the process of exon definition is required for the utilization of upstream 3'ss within this cluster (72).

3' splice site A7: Along with the 5'ss D4, usage of 3'ss A7 results in the removal of the 3'-terminal RRE containing intron and the generation of completely spliced 1.8 kb mRNAs encoding Nef, two-exon Tat and Rev. Regulation of the 3'ss A7 is very complex and under the influence of multiple ESEs, ESSs and ISSs (73).

1.4 HIV-1 Therapy

1.4.1 *HIV-1 drug therapy*

Current HIV-1 therapy consists of a cocktail of antiretroviral (ARV) drugs and is referred to as combination ARV therapy (cART). Therapeutic drugs have been designed to inhibit RT, IN, PR and even viral entry by targeting the host cells entry co-receptor CCR5. The nucleotide RT inhibitor, Azidothymidine, was the first drug approved for treatment of HIV by the United States Food and Drug Administration in 1987. Since then, 27 unique ARV drugs have been approved spanning over 5 different therapeutic classes. cART consists of the combination of at least two different classes of ARVs to prevent the emergence of drug resistant viruses seen when using monotherapy. The development of cART in the mid-1990s has drastically changed the outcome of infected individuals. What was once considered to be a death sentence has been converted to a chronic manageable disease (74). However, there are some notable drawbacks. The first of which is that cART will likely never be curative due to the persistence of replication competent cART resistant viral reservoirs and treatment is indefinite. Furthermore, there is residual immune dysregulation in individuals treated with cART resulting in a considerable amount of morbidity, namely higher incidence of cardiovascular, liver, kidney and bone disease (75). More importantly, lack of access to proper medication and medical expertise in developing countries remains a significant concern with antiretroviral therapy reaching only 19.5 million of the potential 42.9 million people infected with HIV in 2016 (WHO). Given these well characterized issues, there is a prominent need for a cure.

1.4.2 *HIV cure research*

HIV cure research can be separated into two broad categories: a sterilizing cure by eliminating all viral reservoirs or a functional cure where there is a lack of viral replication in the absence of cART. Currently, the most studied of the cure strategies is the shock and kill approach whereby latency-reversing agents are used to activate latent infected cells making these cells more susceptible to HIV-1 specific immune clearance and the cytopathic effect of viral replication. However, very little clinical progress has been documented (76). In the absence of complete elimination, a more feasible outcome is a functional cure mediated by better immune control of HIV-1 replication. Various approaches such as transplantation of genetically modified $CD8^+$ T lymphocytes, vaccination with specific HIV-1 antigens and therapeutic cytokine injections are currently under investigation. Furthermore, a lot of research has gone into characterizing a small subset of HIV-1 infected individuals termed elite controllers, who in the absence of cART are able to maintain undetectable plasma viral levels. Lastly, the transplant of HIV-1 resistant cells to control HIV-1 replication has also been an area of intense research (77,78).

1.4.3 *HIV-1 gene therapy*

While no individual has been cured through viral elimination or immune control therapy, one infected individual has attained a functional cure via the allogeneic transplantation of genetically modified hematopoietic stem cells (HSCs) containing a homozygous deletion in one of the co-receptors used for viral entry (CCR5Δ32/Δ32) (79). Timothy Brown, also known as the Berlin patient, has remained aviremic for more than 9 years after discontinuing antiretroviral therapy following his transplant and is considered the only individual effectively cured from HIV. Unfortunately, due to the excessive risk associated with allogeneic HSC transplants and few CCR5Δ32/Δ32 carriers such an approach is not feasible in the majority of HIV-1 infected

individuals and more practical approaches are currently being investigated (80). One approach being considered is the autologous transplantation of *ex-vivo* genetically modified HSCs. In this approach, patients HSCs are purified, expanded and transduced with an integrative viral vector expressing RNA- or protein-based antivirals. These cells are then re-infused, providing a persistent source of HIV-resistant lymphoid and myeloid lineage cells, which include all major HIV target cells (81). Similar to current antiretroviral therapy, antiviral gene therapy requires a combinatorial approach to reduce the likelihood of viral resistance. Furthermore, a multiplexing approach is necessary with different classes of RNA- or protein-based antivirals to avoid the saturation of any cellular processing pathways (82). Several antiviral small RNA classes have been proposed for use in anti-HIV gene therapy including small interfering and short hairpin RNAs (siRNAs and shRNAs) that induce post-transcriptional gene silencing (PTGS) by the RNA interference (RNAi) pathway, decoy RNAs that sequester viral RNA binding proteins in a stoichiometric fashion, ribozymes that catalytically cleave target RNAs and U1 interfering RNAs (U1i RNAs) that either inhibit polyadenylation of target RNAs by interacting with Poly-A Polymerase (PAP) or induce excessive splicing of target RNAs (83). Understanding the mechanism and comparing the effects of these different classes of antiviral RNAs is crucial for the further development of anti-HIV genes to be used in gene therapy.

1.4.3.1 *siRNA & shRNAs:*

Several steps of the HIV replication cycle have been exploited for the development of therapeutic siRNAs and shRNAs. Although many have demonstrated efficacy, only two shRNAs have progressed to clinical trials: an shRNA targeting CCR5 in combination with a gp41-derived entry inhibitor peptide (C46), sponsored by Calimmune (Los Angeles, CA) (84), and an shRNA targeting the overlapping Tat and Rev coding sequence in HIV-1 RNA in combination with an anti-CCR5 ribozyme and TAR decoy RNA, sponsored by the City of Hope (Duarte, CA) (85).

ShRNAs function through the RNAi pathway, which directs homology-dependent PTGS (86). Therapeutic siRNAs or shRNAs, with perfect Watson-Crick base pairing with their target mRNA, induce sequence-specific cleavage by directing the endonuclease Argonaute 2 (Ago-2) to cleave their target mRNA (87). Unlike siRNAs, which transiently silence gene expression after each drug treatment, shRNAs can induce long-term repression of target genes, as they can be constitutively expressed within cells after transduction (82). ShRNAs are transcribed in the nucleus, exported to the cytoplasm and processed by the endonuclease Dicer in complex with the TAR RNA-binding protein (TRBP) (88). Together, Dicer and TRBP process the longer shRNA into a short siRNA (~21-23nt). They then associate with the molecular "scissor" protein Ago-2 to from the RNA-Induced Silencing Complex (RISC). The two strands of the siRNA then separate releasing a "passenger" strand while the "guide" strand is retained by the RISC (86). The RISC then directs homology-dependent cleavage of target mRNAs by Ago-2 upon recognition and base pairing of target mRNAs with the guide RNA (88,89).

Although shRNAs have been shown to have the greatest potency among antisense molecules, there are several safety concerns for the use of shRNAs in anti-HIV gene therapy. The first of which is their ability to saturate the RNAi pathway and disturb endogenous microRNA processing or function (90). Secondly, they may alter the expression of cellular RNAs through microRNA pathways (91) and lastly, they have the potential to elicit innate immune responses in human cells (92). Nonetheless, safe shRNAs can be developed for anti-HIV gene therapy by extensive testing and the selection of potent molecules that can be expressed at low levels.

1.4.3.2 *Decoy RNAs:*

Decoy RNAs are molecules designed to bind and sequester viral RNA binding proteins from their natural targets. There are currently two clinically tested decoy RNAs targeting HIV-1 Rev and Tat proteins. Rev is targeted by an RRE decoy RNA (93), consisting of a 41-nt sequence from the HIV-1 RRE sequence. The RRE decoy RNA sequesters Rev, reducing the nuclear export of FL and IS HIV-1 RNAs. Tat is targeted by a TAR decoy RNA, consisting of the TAR hairpin of HIV-1 RNA. The TAR decoy serves to sequester Tat, resulting in a reduction of transcribed HIV-1 genes. However, both TAR and RRE decoy RNAs have the potential to inhibit RNAi activity, due to their ability to bind TRBP and sequester it from the RISC (94).

Clinical trials involving the TAR decoy expressed in combination with a shRNA and a ribozyme by the City of Hope are ongoing (84). However, clinical trials involving the transduction of HSCs with an RRE decoy have been terminated due to the low level of gene-containing leukocytes present in the peripheral blood one year after engraftment. Nevertheless, these clinical trials have demonstrated the safety and feasibility of gene therapy and decoy RNAs remain competitive candidates for anti-HIV gene therapy (93).

1.4.3.3 *Ribozymes:*

Ribozymes are an attractive candidate for use in anti-HIV gene therapy, as they do not require proteins for catalysis. These catalytic RNAs are found in the genomes of many organisms and thorough research has led to the development of therapeutic ribozymes. With molecular techniques, the substrate recognition domain of ribozymes has been artificially engineered to direct site-specific cleavage of HIV-1 RNAs. Several anti-HIV ribozymes have been clinically tested in a gene therapy setting (95-97) and in 2009, the first phase 2 cell-delivered gene therapy clinical trial involving the transduction of HSCs with a Tat/Rev specific anti-HIV ribozyme was

completed (98). Although the primary efficacy endpoint was not attained, HIV-1 viral loads were constantly lower in patients who underwent treatment (98). Limited efficacy can be attributed to suboptimal kinetics and function of the therapeutic ribozymes *in vivo,* as phase 1 clinical trials demonstrated successful engraftment, differentiation and expansion of autologous HSCs post transduction (96). Moreover, lack of conditioning during this clinical trial prior to HSC infusion could have resulted in poor engraftment (96).

1.4.4 *Recent progress in combinatorial therapy:*

Two combinatorial gene therapies have made it to clinical trials. The first, sponsored by the City of Hope, uses the combination of a shRNA targeting the Tat/Rev coding sequence of HIV-1 RNA (sh5983), a U16 small nucleolar RNA conjugated TAR decoy RNA (snoTAR) and a ribozyme targeting CCR5 mRNA (85). However, antiviral efficacy could not be evaluated due to the scarcity of genetically modified cells preceding HSC transplant. The low frequency of genetically modified HSC was due to the high expression of the three antiviral RNAs from strong RNA Polymerase III U6 promoters, which led to cellular toxicity (85). This group has recently published promising preclinical data after modifying their platform to co-express three small RNAs polycistronically from a single Polymerase II U1 promoter. Furthermore, they have found that the expression of snoTAR resulted in less optimal inhibition of HIV-1 replication and toxicity in transduced HSCs (99). The second combinatorial strategy that has reached clinical trials is sponsored by Calimmune and utilizes a protein-based C46 peptide fusion inhibitor and a shRNA targeting CCR5 mRNA expressed from an RNA Polymerase III H1 promoter (100-102). They have demonstrated efficacy and safety in both cell culture and animal models and clinical trials of HSC transplant have recently been initiated. Although Calimmune's combination has demonstrated promising pre-clinical results, the use of only two inhibitors against viral entry makes such an approach prone to viral resistance.

Collectively, these findings indicate that current antiviral RNA combinations being utilized in clinical trials are suboptimal and there is an ongoing need for the evaluation of alternative anti-HIV RNA candidates alone and in combination to identify an optimal set of antiviral RNAs for use in HIV gene therapy. While much research has gone into the various classes of antiviral RNAs described above, little attention has been given to the newly described U1 interference RNAs also known as U1i RNAs.

1.5 Modified U1 snRNAs:

1.5.1 *Structure of the U1 snRNP*

As discussed previously, the U1 snRNP is an essential molecule in the process of splicing, more specifically in the definition of the 5'ss during spliceosomal assembly. The U1 snRNP consists of 10 proteins which are recruited to a 164 nucleotide U1 RNA molecule. Seven of the proteins are Smith (Sm) proteins which are common to all U snRNPs while three, U1-70K, U1-A and U1-C, are specific to the U1 snRNA (103). The U1 snRNA contains a 10 nucleotide 5'ss recognition domain, 4 stem loop (SL) regions and a single stranded region where the Sm proteins bind. The U1 snRNP naturally binds 5'ss with the sequence CAG|GTAAGTA (vertical bar represents the splice site). Association of the Sm proteins with U1 snRNAs is essential in the biogenesis of these molecules. Once exported to the cytoplasm, Sm proteins complex with the U1 snRNA forming the core RNP domain and are responsible for hypermethylation of the 5' cap structure. Along with the hypermethylated 5' cap structure, nuclear localization signals found in the core RNP domain mediate the nuclear import of U1 snRNP particles (104).

U1-70K and U1-A bind to SL1 and SL2 domains of the U1 RNA molecule, respectively, while U1-C is recruited via protein-protein interactions with U1-70K and Sm proteins. These U1-specific proteins have important roles in aiding the recognition of 5'ss in pre-mRNA

transcripts. The U1-70K protein has an RNA binding domain allowing it to bind SL1 of the U1 snRNA and a C-terminus arginine-serine rich RS domain which mediates interactions with other splicing factors such as ASF/SF2 (105). As previously mentioned, ASF/SF2 promotes recruitment of the U1 snRNP to 5'ss and has a role in bridging 5' and 3'ss (106). The U1-A protein is not essential for splicing mediated by the U1 snRNP molecule as *in vitro* splicing occurs in the absence of U1-A. However, U1-A may have a role in the modulation of 5'ss selection and has been characterized to inhibit polyadenylation when binding directly to specific transcripts (107,108). Lastly, U1-C recruitment to the U1 snRNP molecule is dependent on prior binding of the Sm proteins and U1-70K and facilitates recognition of 5'ss by the U1 snRNP (109). While initially thought to be an essential splicing factor, depletion of U1-C only had an effect on a minor group of weakly defined exons (110). A depiction of the U1 snRNA molecule along with associated proteins can be found in Figure 1.5.

1.5.2 *Therapeutic potential of modified U1 snRNAs*

The first indication that the U1 snRNP had the potential to be used as an RNAi molecule came from a study conducted on the bovine papillomavirus by Furth and colleagues in 1994. In this study, they found that sequences homologous to 5'ss present in the 3' untranslated region, proximal to the polyadenylation site (PAS) of papillomavirus RNA, are necessary for the inhibition of viral late gene expression. The inhibition seen was demonstrated to be dependent upon base pairing of the 5' end of the U1 snRNA to these 5' like splice sequences and was mediated by a decrease in polyadenylation efficiency (111). Soon after, Gunderson et al. showed that U1 snRNP inhibition of pre-mRNA polyadenylation is mediated by a direct interaction between U1-70K and PAP that result in an inhibition of proper polyadenylation (112). Other viruses such as Human papillomavirus 16 and HIV-1 also use the ability of the U1 snRNP to control 3' end processing (113). In the case of HIV-1, binding of the U1 snRNP to the 5'ss D1

results in the suppression of HIV-1 pre-mRNA polyadenylation mediated by the PAS in the 5' LTR and is required for expression of genomic FL RNA (114,115). Moreover, U1 snRNP mediated 3' end processing has also been demonstrated to be crucial in the expression of eukaryotic genes. U1 snRNP suppresses cryptic PASs in eukaryotic introns protecting host cells from aberrant polyadenylation and premature transcript termination (116).

1.5.3 *U1 interfering RNAs*

Based on the ability of the U1 snRNP to inhibit 3' end processing, modified U1 RNA molecules, termed U1 interfering (U1i) RNAs, were constructed and were demonstrated to inhibit the expression of exogenous and endogenous targeted genes in both cell culture and *in vivo* (117-119). To achieve inhibition, the first 10 nucleotides of the 5'ss recognition domain of the U1 RNA molecule are modified to contain complementary sequences to the terminal exon of a targeted gene. U1i RNA molecules are then expressed in plasmids under the control of the U1 promoter or transduced for *in vivo* studies. The inhibitory effect of U1i RNAs is due to the lack of poly(A) tail addition to targeted pre-mRNA transcripts and is mediated by interactions between U1-70K and PAP (118). Non-polyadenylated transcripts are inherently unstable and rapidly degraded by the host cell machinery (120). Importantly, the inhibition seen by U1i RNAs directed to the 3'-terminal exon is not due to the activation of aberrant splicing as binding of U1i RNAs has been demonstrated to activate splicing at non-physiological 5'ss (121). The mechanism of action of U1i RNAs is depicted in Figure 1.5.

Several studies have optimized the construction of U1i RNAs targeting 3'-terminal exons of pre-mRNA. Several aspects to take into consideration are summarized below:

1. U1i RNAs targeting 3'-terminal exons can only inhibit the expression of pre-mRNAs that are polyadenylated by the host cell polyadenylation machinery.

2. For inhibition of polyadenylation, the U1i RNA binding site must be located in 3'-terminal exons and can be up to 1,200 nucleotides upstream from the PAS (118). U1i RNAs can also target sites downstream of the PAS (122). However, these molecules seem to inhibit the cleavage reaction rather than the addition of the poly(A) tail and have not been extensively studied (123).

3. Increasing the length of the U1i RNA recognition domain from 10 to 16 nucleotides towards the 3' end has no effect on the inhibitory potential of these molecules (122). Increasing the recognition domain towards the 5' end of U1i RNA molecule by 1 to 5 nucleotides results in a loss of inhibitory activity. In contrast, molecules with recognition domains increased by 6 nucleotides maintained their inhibitory activity (124).

4. U1i RNA targeting sites in secondary RNA structures or in proximity to certain endogenous U1 binding sites lose their inhibitory activity (124).

5. Single mismatch differences in positions 3 to 8 of the U1i RNA target site abolish inhibitory activity. However, mutations in position 1, 2, 9 and 10 result in a partial loss of inhibitory activity (122,124).

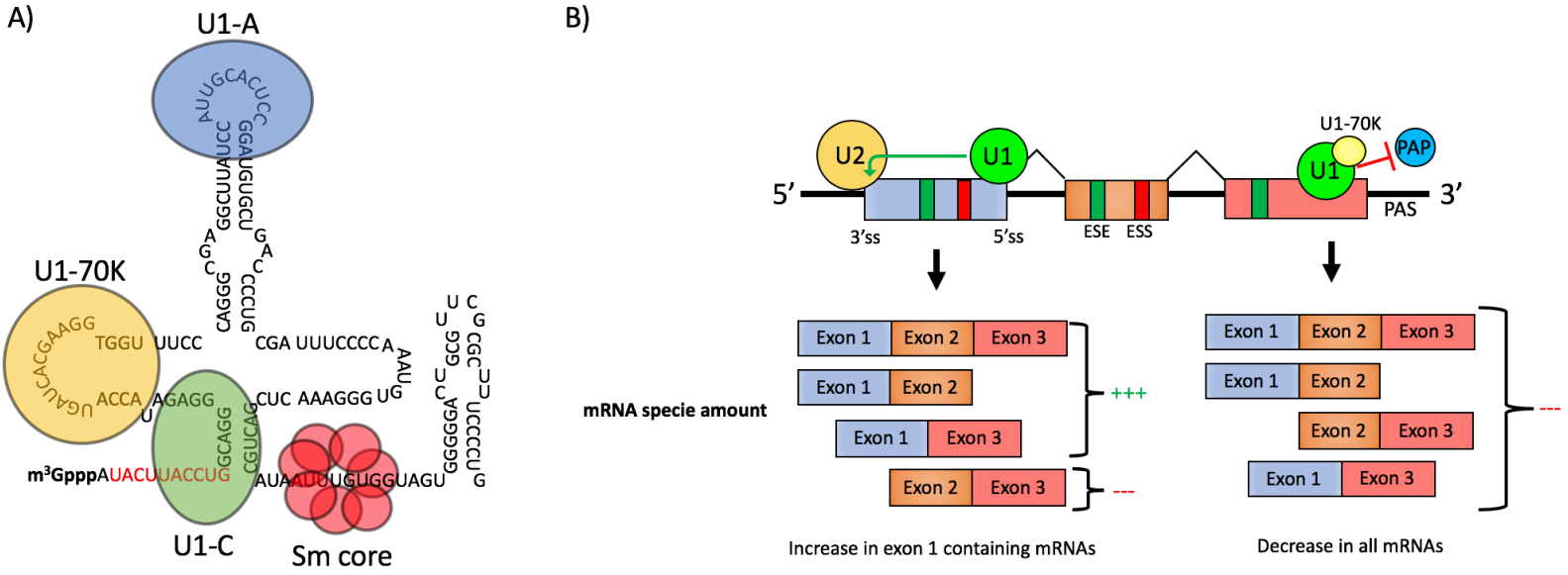

Figure 1.5 U1 snRNP molecule and mechanism of U1 RNA interference. A) The sequence of the U1 snRNA along with the various associated proteins, U1-70K, U1-A, U1-C and 7 Sm proteins. In red is the recognition domain which can be modified to target mRNA transcripts for inhibition. B) The mechanism of action of U1i RNAs targeting 5'ss or downstream of 3'ss and 3'-terminal exons. U1i RNAs directed to 5'ss or downstream of 3'ss increase recognition of the upstream 3'ss and result in an increase in mRNA isoforms that contain the defined exon. Targeting 3'-terminal exons results in a decrease in all mRNA isoforms. Based on information from (103)

6. Inhibitory effect of U1i RNAs targeting the terminal 3' exon is dependent upon proper assembly of the SL1 domain and recruitment of the U1-70K protein (124).

7. Co-expression of two U1i RNAs or a U1i RNA and an shRNA results in a synergistic inhibition of targeted mRNA transcript (118,125,126).

Other U1i RNA have been designed to modulate gene expression by targeting 5'ss and disrupting proper splicing of the targeted gene (127). Since the U1 snRNP has a role in defining the 5'ss of pre-mRNAs, binding of a U1i RNA to a 5'ss will increase the upstream splicing at the 3'ss by the process of exon definition. By this mechanism, U1i RNAs targeting 5'ss induce excessive splicing and reduce the expression of their targeted gene. However, little to no research has gone into determining the specificity, safety and optimization of these molecules as done for U1i RNA targeting terminal exons. Utilizing the same mechanism, therapeutic modified U1 snRNAs were also developed to correct splicing defects seen in genetic diseases.

1.5.4 *Correction of splicing defects*

Over 50% of disease-causing mutations affect splicing and the most frequent mutations cause exon skipping. Therefore, much research is being directed at the utilization of modified U1 snRNAs to correct these exon skipping defects. These molecules are not considered U1i RNAs as they do not inhibit expression of targeted transcripts. When directed to the 5'ss of mutated exons, these modified U1 snRNAs increase the recognition of the upstream 3'ss and reduce the amount of exon skipping.

One of the best characterized disease involving exon skipping is spinal muscular atrophy (SMA). However, exon skipping is also implicated in various cancers, cystic fibrosis (CF) and Heamophilia B disorder. In 2012, Fernandez et al. demonstrated that by modifying the 5' end of U1 snRNA molecules, exon-specific U1 snRNAs can be designed that are able to correct the

exon skipping defects in the genes involved in SMA, CF and Heamophilia B (128). This was taken one step further by Rogalska et al., who developed transgenic SMA mice that had a germline-integrated exon-specific U1 snRNA molecule designed by Fernandez et al. to rescue the exon skipping defect in SMA. They demonstrated that expression of the SMA specific U1 snRNA significantly improved the disease symptoms as well as the survival of these SMA mice (129). More recently, studies published by Danadon et al., delivered exon-specific U1 RNAs utilizing an adeno-associated virus and effectively reversed familial dysautonomia, a rare genetic disease cause by an intronic point mutation (130). Importantly, *in vivo* expression of these modified U1 snRNA molecules did not cause any detectable toxicity. However, experiments are still required to analyze the specificity of these modified U1 RNA molecules and more detailed safety studies will be needed before they can advance to clinical trials.

1.5.5 *U1i RNAs targeting HIV-1*

To date, there have been three studies conducted utilizing U1i RNAs as a means to inhibit HIV-1 replication. The first was conducted by Sajic et al, in 2007 (131). To identify target sites in the 3'-terminal exon, consensus sequences of the Env and Nef regions of all circulating M and O clades of HIV-1 were analyzed for highly conserved 10 nucleotide sites. These 10 nucleotide sites were compared to the human genome using Basic Local Alignment Search Tool (BLAST) to exclude molecules that would target human genes. This led to the development of 15 different U1i RNAs targeting conserved sites in the terminal exon of HIV-1. Inhibition of HIV-1 was measured by Western blot, measuring p24 protein expression. Of the 15 U1i RNAs developed, two (U1 αHIV-1 #1 and #3) demonstrated complete inhibition of p24 expression. The inhibition seen was not due to gross toxicity and expression of these U1i RNAs did not reduce the expression of non-target RNA secreted alkaline phosphatase when co-expressed from a different plasmid. Furthermore, mutational studies demonstrated that inhibition of p24 expression was

dependent on the recruitment of U1-70K to the SL1 domain, the recruitment of the Sm proteins and to a certain extent on the recruitment of U1-A to the SL2 domain. Importantly, cells stably expressing U1i RNAs targeting HIV-1 showed reduced capacity to support HIV replication.

In an independent study conducted by Knoepfel et al., 14 new U1i RNA molecules targeting the 3'-terminal exon of HIV-1 were designed (132). This screen included molecules targeting the 3' LTR, specifically in the U3 and R regions. Inhibition of HIV-1 production was measured using p24 ELISA. Of the 14 U1i RNA molecules, U1i RNA J, C and BD were found to be potent inhibitors of p24 production. However, the expression of U1i RNAs in this study resulted in significant non-specific suppression of Firefly luciferase reporter plasmid indicating the lack of specificity of these molecules. Nonetheless, lentiviral vectors with various U1i RNA constructs were then used to transduce SupT1 cells. However, stably transduced cells did not have the ability to inhibit viral replication (132).

Lastly, in a study conducted by Mandal et al., U1i RNAs were constructed to inhibit HIV-1 replication by promoting excessive splicing at HIV-1 splice acceptor sites A1, A2 and A3 (133). To promote excessive splicing at 3'ss A1 and A2, U1i RNAs were modified to bind 5'ss D2 and D3. Since the 5'ss D4 is considered to be strong, to induce excessive splicing at the 3'ss A3, U1i RNAs were designed to target regions downstream of A3 and upstream of 3'ss A4c, a and b. Expression of U1i RNAs targeting the 5'ss D2 resulted in a drastic decrease in FL mRNA and a concomitant increase in Vif mRNA. Furthermore, RT-PCR demonstrated that expression of D2 specific U1i RNAs resulted in an increase in mRNA species containing non-coding exon 2 as expected by exon definition. Expression of U1i RNAs targeting 5'ss D3 resulted in a decrease in FL mRNA and an increase in IS and CS mRNA containing non-coding exon 3. Finally, expression of U1i RNAs targeting regions downstream of 3' splice acceptor A3 resulted in a

decrease of FL mRNA and an increase in IS and CS mRNA species containing exon 4. Expression of these U1i RNAs drastically reduced viral production when co-transfected with an HIV-1 proviral plasmid (pNL4-3) and inhibited viral replication in transduced T lymphocytes.

References:

1. (1981) Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR. Morbidity and mortality weekly report*, **30**, 305-308.

2. Centers for Disease, C. (1981) Pneumocystis pneumonia--Los Angeles. *MMWR. Morbidity and mortality weekly report*, **30**, 250-252.

3. Centers for Disease, C. (1982) Update on Kaposi's sarcoma and opportunistic infections in previously healthy persons--United States. *MMWR. Morbidity and mortality weekly report*, **31**, 294, 300-291.

4. (1983) Immunodeficiency among female sexual partners of males with acquired immune deficiency syndrome (AIDS) - New York. *MMWR. Morbidity and mortality weekly report*, **31**, 697-698.

5. Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C. *et al.* (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **220**, 868-871.

6. Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B. *et al.* (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, **224**, 500-503.

7. Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P. *et al.* (1986) What to call the AIDS virus? *Nature*, **321**, 10.

8. Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M.A., Santos-Ferreira, M.O., Laurent, A.G., Dauguet, C., Katlama, C., Rouzioux, C. *et al.* (1986) Isolation of a new human retrovirus from West African patients with AIDS. *Science*, **233**, 343-346.

9. Aghokeng, A.F., Ayouba, A., Mpoudi-Ngole, E., Loul, S., Liegeois, F., Delaporte, E. and Peeters, M. (2010) Extensive survey on the prevalence and genetic diversity of SIVs in primate bushmeat provides insights into risks for potential new cross-species transmissions. *Infect Genet Evol*, **10**, 386-396.

10. Hahn, B.H., Shaw, G.M., De Cock, K.M. and Sharp, P.M. (2000) AIDS as a zoonosis: scientific and public health implications. *Science*, **287**, 607-614.

11. Sharp, P.M. and Hahn, B.H. (2011) Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*, **1**, a006841.

12. Hemelaar, J. (2012) The origin and diversity of the HIV-1 pandemic. *Trends Mol Med*, **18**, 182-192.

13. Keele, B.F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M.L., Bibollet-Ruche, F., Chen, Y., Wain, L.V., Liegeois, F. *et al.* (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science*, **313**, 523-526.

14. Hamel, D.J., Sankale, J.L., Eisen, G., Meloni, S.T., Mullins, C., Gueye-Ndiaye, A., Mboup, S. and Kanki, P.J. (2007) Twenty years of prospective molecular epidemiology in Senegal: changes in HIV diversity. *AIDS Res Hum Retroviruses*, **23**, 1189-1196.

15. Popper, S.J., Sarr, A.D., Gueye-Ndiaye, A., Mboup, S., Essex, M.E. and Kanki, P.J. (2000) Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. *J Virol*, **74**, 1554-1557.

16. Hemelaar, J. (2013) Implications of HIV diversity for the HIV-1 pandemic. *J Infect*, **66**, 391-400.

17. Patel, P., Borkowf, C.B., Brooks, J.T., Lasry, A., Lansky, A. and Mermin, J. (2014) Estimating per-act HIV transmission risk: a systematic review. *AIDS*, **28**, 1509-1519.

18. Wong, J.K. and Yukl, S.A. (2016) Tissue reservoirs of HIV. *Curr Opin HIV AIDS*, **11**, 362-370.

19. Swanstrom, R. and Coffin, J. (2012) HIV-1 pathogenesis: the virus. *Cold Spring Harb Perspect Med*, **2**, a007443.

20. Deeks, S.G., Overbaugh, J., Phillips, A. and Buchbinder, S. (2015) HIV infection. *Nat Rev Dis Primers*, **1**, 15035.

21. Moir, S., Chun, T.W. and Fauci, A.S. (2011) Pathogenic mechanisms of HIV disease. *Annu Rev Pathol*, **6**, 223-248.

22. Biberfeld, P., Ensoli, B., Sturzl, M. and Schulz, T.F. (1998) Kaposi sarcoma-associated herpesvirus/human herpesvirus 8, cytokines, growth factors and HIV in pathogenesis of Kaposi's sarcoma. *Curr Opin Infect Dis*, **11**, 97-105.

23. Masur, H., Brooks, J.T., Benson, C.A., Holmes, K.K., Pau, A.K. and Kaplan, J.E. (2014) Prevention and treatment of opportunistic infections in HIV-infected adults and adolescents: Updated Guidelines from the Centers for Disease Control and Prevention, National Institutes of Health, and HIV Medicine Association of the Infectious Diseases Society of America. *Clin Infect Dis*, **58**, 1308-1311.

24. Berger, E.A., Murphy, P.M. and Farber, J.M. (1999) Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol*, **17**, 657-700.

25. Dahabieh, M.S., Battivelli, E. and Verdin, E. (2015) Understanding HIV latency: the road to an HIV cure. *Annu Rev Med*, **66**, 407-421.

26. Frankel, A.D. and Young, J.A. (1998) HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem*, **67**, 1-25.

27. Arthos, J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D., Xiao, Z., Veenstra, T.D., Conrad, T.P., Lempicki, R.A. *et al.* (2008) HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. *Nat Immunol*, **9**, 301-309.

28. Geijtenbeek, T.B., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Middel, J., Cornelissen, I.L., Nottet, H.S., KewalRamani, V.N., Littman, D.R. *et al.* (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, **100**, 587-597.

29. Nisole, S. and Saib, A. (2004) Early steps of retrovirus replicative cycle. *Retrovirology*, **1**, 9.

30. Clapham, P.R. and McKnight, A. (2001) HIV-1 receptors and cell tropism. *Br Med Bull*, **58**, 43-59.

31. Wilen, C.B., Tilton, J.C. and Doms, R.W. (2012) HIV: cell binding and entry. *Cold Spring Harb Perspect Med*, **2**.

32. Zhang, H., Dornadula, G., Orenstein, J. and Pomerantz, R.J. (2000) Morphologic changes in human immunodeficiency virus type 1 virions secondary to intravirion reverse transcription: evidence indicating that reverse transcription may not take place within the intact viral core. *J Hum Virol*, **3**, 165-172.

33. Kleiman, L. (2002) tRNA(Lys3): the primer tRNA for reverse transcription in HIV-1. *IUBMB Life*, **53**, 107-114.

34. Panganiban, A.T. and Fiore, D. (1988) Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science*, **241**, 1064-1069.

35. Telesnitsky, A. and Goff, S.P. (1997) In Coffin, J. M., Hughes, S. H. and Varmus, H. E. (eds.), *Retroviruses*, Cold Spring Harbor (NY).

36. Hu, W.S. and Hughes, S.H. (2012) HIV-1 reverse transcription. *Cold Spring Harb Perspect Med*, **2**.

37. Zennou, V., Petit, C., Guetard, D., Nerhbass, U., Montagnier, L. and Charneau, P. (2000) HIV-1 genome nuclear import is mediated by a central DNA flap. *Cell*, **101**, 173-185.

38. Hameau, L., Jeusset, J., Lafosse, S., Coulaud, D., Delain, E., Unge, T., Restle, T., Le Cam, E. and Mirambeau, G. (2001) Human immunodeficiency virus type 1 central DNA flap: dynamic terminal product of plus-strand displacement dna synthesis catalyzed by reverse transcriptase assisted by nucleocapsid protein. *J Virol*, **75**, 3301-3313.

39. Gorry, P.R., Ong, C., Thorpe, J., Bannwarth, S., Thompson, K.A., Gatignol, A., Vesselingh, S.L. and Purcell, D.F. (2003) Astrocyte infection by HIV-1: mechanisms of restricted virus replication, and role in the pathogenesis of HIV-1-associated dementia. *Curr HIV Res*, **1**, 463-473.

40. Jayappa, K.D., Ao, Z. and Yao, X. (2012) 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.

41. Engelman, A., Mizuuchi, K. and Craigie, R. (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, **67**, 1211-1221.

42. Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R. and Bushman, F. (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, **110**, 521-529.

43. Ott, M., Geyer, M. and Zhou, Q. (2011) The control of HIV transcription: keeping RNA polymerase II on track. *Cell Host Microbe*, **10**, 426-435.

44. Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W. and Luhrmann, R. (1995) The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell*, **82**, 475-483.

45. Swanstrom, R. and Wills, J.W. (1997) In Coffin, J. M., Hughes, S. H. and Varmus, H. E. (eds.), *Retroviruses*, Cold Spring Harbor (NY).

46. Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988) Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*, **331**, 280-283.

47. Decroly, E., Vandenbranden, M., Ruysschaert, J.M., Cogniaux, J., Jacob, G.S., Howard, S.C., Marshall, G., Kompelli, A., Basak, A., Jean, F. *et al.* (1994) The convertases furin and PC1 can both cleave the human immunodeficiency virus (HIV)-1 envelope glycoprotein gp160 into gp120 (HIV-1 SU) and gp41 (HIV-I TM). *J Biol Chem*, **269**, 12240-12247.

48. Freed, E.O. (2015) HIV-1 assembly, release and maturation. *Nat Rev Microbiol*, **13**, 484-496.

49. Stumptner-Cuvelette, P., Morchoisne, S., Dugast, M., Le Gall, S., Raposo, G., Schwartz, O. and Benaroch, P. (2001) HIV-1 Nef impairs MHC class II antigen presentation and surface expression. *Proc Natl Acad Sci U S A*, **98**, 12144-12149.

50. Lubben, N.B., Sahlender, D.A., Motley, A.M., Lehner, P.J., Benaroch, P. and Robinson, M.S. (2007) HIV-1 Nef-induced down-regulation of MHC class I requires AP-1 and clathrin but not PACS-1 and is impeded by AP-2. *Mol Biol Cell*, **18**, 3351-3365.

51. Chaudhuri, R., Lindwasser, O.W., Smith, W.J., Hurley, J.H. and Bonifacino, J.S. (2007) Downregulation of CD4 by human immunodeficiency virus type 1 Nef is dependent on clathrin and involves direct interaction of Nef with the AP2 clathrin adaptor. *J Virol*, **81**, 3877-3890.

52. Margottin, F., Bour, S.P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K. and Benarous, R. (1998) A novel human WD protein, h-beta TrCp, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif. *Mol Cell*, **1**, 565-574.

53. Neil, S.J., Zang, T. and Bieniasz, P.D. (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, **451**, 425-430.

54. Harris, R.S. and Liddament, M.T. (2004) Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*, **4**, 868-877.

55. Kogan, M. and Rappaport, J. (2011) HIV-1 accessory protein Vpr: relevance in the pathogenesis of HIV and potential for therapeutic intervention. *Retrovirology*, **8**, 25.

56. Stoltzfus, C.M. and Madsen, J.M. (2006) Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. *Curr HIV Res*, **4**, 43-55.

57. Wongpalee, S.P. and Sharma, S. (2014) The pre-mRNA splicing reaction. *Methods Mol Biol*, **1126**, 3-12.

58. Wang, Z. and Burge, C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*, **14**, 802-813.

59. Roca, X., Sachidanandam, R. and Krainer, A.R. (2005) Determinants of the inherent strength of human 5' splice sites. *RNA*, **11**, 683-698.

60. Wu, S., Romfo, C.M., Nilsen, T.W. and Green, M.R. (1999) Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, **402**, 832-835.

61. Zheng, Z.M. (2004) Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J Biomed Sci*, **11**, 278-294.

62. Graveley, B.R. (2000) Sorting out the complexity of SR protein functions. *RNA*, **6**, 1197-1211.

63. Purcell, D.F. and Martin, M.A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol*, **67**, 6365-6378.

64. O'Reilly, M.M., McNally, M.T. and Beemon, K.L. (1995) Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology*, **213**, 373-385.

65. Kammler, S., Otte, M., Hauber, I., Kjems, J., Hauber, J. and Schaal, H. (2006) The strength of the HIV-1 3' splice sites affects Rev function. *Retrovirology*, **3**, 89.

66. Karn, J. and Stoltzfus, C.M. (2012) Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med*, **2**, a006916.

67. Stoltzfus, C.M. (2009) Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. *Adv Virus Res*, **74**, 1-40.

68. Exline, C.M., Feng, Z. and Stoltzfus, C.M. (2008) Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. *J Virol*, **82**, 3921-3931.

69. Bilodeau, P.S., Domsic, J.K., Mayeda, A., Krainer, A.R. and Stoltzfus, C.M. (2001) RNA splicing at human immunodeficiency virus type 1 3' splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element. *J Virol*, **75**, 8487-8497.

70. Zahler, A.M., Damgaard, C.K., Kjems, J. and Caputi, M. (2004) SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing. *J Biol Chem*, **279**, 10077-10084.

71. Caputi, M., Freund, M., Kammler, S., Asang, C. and Schaal, H. (2004) A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *J Virol*, **78**, 6517-6526.

72. Asang, C., Hauber, I. and Schaal, H. (2008) Insights into the selective activation of alternatively used splice acceptors by the human immunodeficiency virus type-1 bidirectional splicing enhancer. *Nucleic Acids Res*, **36**, 1450-1463.

73. Staffa, A. and Cochrane, A. (1995) Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol Cell Biol*, **15**, 4597-4605.

74. Palmisano, L. and Vella, S. (2011) A brief history of antiretroviral therapy of HIV infection: success and challenges. *Ann Ist Super Sanita*, **47**, 44-48.

75. Lederman, M.M., Funderburg, N.T., Sekaly, R.P., Klatt, N.R. and Hunt, P.W. (2013) Residual immune dysregulation syndrome in treated HIV infection. *Adv Immunol*, **119**, 51-83.

76. Deeks, S.G., Lewin, S.R., Ross, A.L., Ananworanich, J., Benkirane, M., Cannon, P., Chomont, N., Douek, D., Lifson, J.D., Lo, Y.R. *et al.* (2016) International AIDS Society global scientific strategy: towards an HIV cure 2016. *Nat Med*, **22**, 839-850.

77. Deeks, S. (2014) Towards an HIV cure. *J Int AIDS Soc*, **17**, 19479.

78. Henrich, T.J. and Lelievre, J.D. (2018) Progress towards obtaining an HIV cure: slow but sure. *Curr Opin HIV AIDS*.

79. Hutter, G., Nowak, D., Mossner, M., Ganepola, S., Mussig, A., Allers, K., Schneider, T., Hofmann, J., Kucherer, C., Blau, O. *et al.* (2009) Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med*, **360**, 692-698.

80. Petz, L.D., Burnett, J.C., Li, H., Li, S., Tonai, R., Bakalinskaya, M., Shpall, E.J., Armitage, S., Kurtzberg, J., Regan, D.M. *et al.* (2015) Progress toward curing HIV infection with hematopoietic cell transplantation. *Stem Cells Cloning*, **8**, 109-116.

81. Kiem, H.P., Jerome, K.R., Deeks, S.G. and McCune, J.M. (2012) Hematopoietic-stem-cell-based gene therapy for HIV disease. *Cell Stem Cell*, **10**, 137-147.

82. Burnett, J.C. and Rossi, J.J. (2012) RNA-based therapeutics: current progress and future prospects. *Chem Biol*, **19**, 60-71.

83. Blazquez, L. and Fortes, P. (2015) U1 interference (U1i) for antiviral approaches. *Adv Exp Med Biol*, **848**, 51-69.

84. Wolstein, O., Boyd, M., Millington, M., Impey, H., Boyer, J., Howe, A., Delebecque, F., Cornetta, K., Rothe, M., Baum, C. *et al.* (2014) 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.

85. DiGiusto, D.L., Krishnan, A., Li, L., Li, H., Li, S., Rao, A., Mi, S., Yam, P., Stinson, S., Kalos, M. *et al.* (2010) RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med*, **2**, 36ra43.

86. Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199-208.

87. Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. and Conklin, D.S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev*, **16**, 948-958.

88. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell*, **15**, 185-197.

89. Daniels, S.M. and Gatignol, A. (2012) The multiple functions of TRBP, at the hub of cell responses to viruses, stress, and cancer. *Microbiol Mol Biol Rev*, **76**, 652-666.

90. Castanotto, D., Sakurai, K., Lingeman, R., Li, H., Shively, L., Aagaard, L., Soifer, H., Gatignol, A., Riggs, A. and Rossi, J.J. (2007) Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res*, **35**, 5154-5164.

91. Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L. and Linsley, P.S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*, **12**, 1179-1187.

92. Whitehead, K.A., Dahlman, J.E., Langer, R.S. and Anderson, D.G. (2011) Silencing or stimulation? siRNA delivery and the immune system. *Annu Rev Chem Biomol Eng*, **2**, 77-96.

93. Kohn, D.B., Bauer, G., Rice, C.R., Rothschild, J.C., Carbonaro, D.A., Valdez, P., Hao, Q., Zhou, C., Bahner, I., Kearns, K. *et al.* (1999) A clinical trial of retroviral-mediated transfer of a rev-responsive element decoy gene into CD34(+) cells from the bone marrow of human immunodeficiency virus-1-infected children. *Blood*, **94**, 368-371.

94. Daniels, S.M., Sinck, L., Ward, N.J., Melendez-Pena, C.E., Scarborough, R.J., Azar, I., Rance, E., Daher, A., Pang, K.M., Rossi, J.J. *et al.* (2015) HIV-1 RRE RNA acts as an RNA silencing suppressor by competing with TRBP-bound siRNAs. *RNA Biol*, **12**, 123-135.

95. Michienzi, A., Castanotto, D., Lee, N., Li, S., Zaia, J.A. and Rossi, J.J. (2003) RNA-mediated inhibition of HIV in a gene therapy setting. *Ann N Y Acad Sci*, **1002**, 63-71.

96. Amado, R.G., Mitsuyasu, R.T., Rosenblatt, J.D., Ngok, F.K., Bakker, A., Cole, S., Chorn, N., Lin, L.S., Bristol, G., Boyd, M.P. *et al.* (2004) Anti-human immunodeficiency virus hematopoietic progenitor cell-delivered ribozyme in a phase I study: myeloid and lymphoid reconstitution in human immunodeficiency virus type-1-infected patients. *Hum Gene Ther*, **15**, 251-262.

97. Wong-Staal, F., Poeschla, E.M. and Looney, D.J. (1998) 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.

98. Mitsuyasu, R.T., Merigan, T.C., Carr, A., Zack, J.A., Winters, M.A., Workman, C., Bloch, M., Lalezari, J., Becker, S., Thornton, L. *et al.* (2009) Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nat Med*, **15**, 285-292.

99. Chung, J., Scherer, L.J., Gu, A., Gardner, A.M., Torres-Coronado, M., Epps, E.W., Digiusto, D.L. and Rossi, J.J. (2014) 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.

100. Shimizu, S., Hong, P., Arumugam, B., Pokomo, L., Boyer, J., Koizumi, N., Kittipongdaja, P., Chen, A., Bristol, G., Galic, Z. *et al.* (2010) A highly efficient short hairpin RNA potently down-regulates CCR5 expression in systemic lymphoid organs in the hu-BLT mouse model. *Blood*, **115**, 1534-1544.

101. Shimizu, S., Kamata, M., Kittipongdaja, P., Chen, K.N., Kim, S., Pang, S., Boyer, J., Qin, F.X., An, D.S. and Chen, I.S. (2009) Characterization of a potent non-cytotoxic shRNA directed to the HIV-1 co-receptor CCR5. *Genet Vaccines Ther*, **7**, 8.

102. An, D.S., Donahue, R.E., Kamata, M., Poon, B., Metzger, M., Mao, S.H., Bonifacino, A., Krouse, A.E., Darlix, J.L., Baltimore, D. *et al.* (2007) 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.

103. Stark, H., Dube, P., Luhrmann, R. and Kastner, B. (2001) Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*, **409**, 539-542.

104. Fischer, U. and Luhrmann, R. (1990) An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science*, **249**, 786-790.

105. Guiro, J. and O'Reilly, D. (2015) Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley Interdiscip Rev RNA*, **6**, 79-92.

106. Jamison, S.F., Pasman, Z., Wang, J., Will, C., Luhrmann, R., Manley, J.L. and Garcia-Blanco, M.A. (1995) U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic Acids Res*, **23**, 3260-3267.

107. Tarn, W.Y. and Steitz, J.A. (1995) Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *Proc Natl Acad Sci U S A*, **92**, 2504-2508.

108. Workman, E., Veith, A. and Battle, D.J. (2014) U1A regulates 3' processing of the survival motor neuron mRNA. *J Biol Chem*, **289**, 3703-3712.

109. Will, C.L., Rumpler, S., Klein Gunnewiek, J., van Venrooij, W.J. and Luhrmann, R. (1996) In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Res*, **24**, 4614-4623.

110. Rosel-Hillgartner, T.D., Hung, L.H., Khrameeva, E., Le Querrec, P., Gelfand, M.S. and Bindereif, A. (2013) A novel intra-U1 snRNP cross-regulation mechanism: alternative splicing switch links U1C and U1-70K expression. *PLoS Genet*, **9**, e1003856.

111. Furth, P.A., Choe, W.T., Rex, J.H., Byrne, J.C. and Baker, C.C. (1994) Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. *Mol Cell Biol*, **14**, 5278-5289.

112. Gunderson, S.I., Polycarpou-Schwarz, M. and Mattaj, I.W. (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell*, **1**, 255-264.

113. Goraczniak, R. and Gunderson, S.I. (2008) The regulatory element in the 3'-untranslated region of human papillomavirus 16 inhibits expression by binding CUG-binding protein 1. *J Biol Chem*, **283**, 2286-2296.

114. Ashe, M.P., Pearson, L.H. and Proudfoot, N.J. (1997) The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J*, **16**, 5752-5763.

115. Ashe, M.P., Griffin, P., James, W. and Proudfoot, N.J. (1995) Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev*, **9**, 3008-3025.

116. Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L. and Dreyfuss, G. (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*, **468**, 664-668.

117. Beckley, S.A., Liu, P., Stover, M.L., Gunderson, S.I., Lichtler, A.C. and Rowe, D.W. (2001) Reduction of target gene expression by a modified U1 snRNA. *Mol Cell Biol*, **21**, 2815-2825.

118. Fortes, P., Cuevas, Y., Guan, F., Liu, P., Pentlicky, S., Jung, S.P., Martinez-Chantar, M.L., Prieto, J., Rowe, D. and Gunderson, S.I. (2003) Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc Natl Acad Sci U S A*, **100**, 8264-8269.

119. Koornneef, A., van Logtenstein, R., Timmermans, E., Pisas, L., Blits, B., Abad, X., Fortes, P., Petry, H., Konstantinova, P. and Ritsema, T. (2011) AAV-mediated in vivo knockdown of luciferase using combinatorial RNAi and U1i. *Gene Ther*, **18**, 929-935.

120. Carpousis, A.J., Vanzo, N.F. and Raynal, L.C. (1999) mRNA degradation. A tale of poly(A) and multiprotein machines. *Trends Genet*, **15**, 24-28.

121. Cohen, J.B., Snow, J.E., Spencer, S.D. and Levinson, A.D. (1994) Suppression of mammalian 5' splice-site defects by U1 small nuclear RNAs from a distance. *Proc Natl Acad Sci U S A*, **91**, 10470-10474.

122. Liu, P., Gucwa, A., Stover, M.L., Buck, E., Lichtler, A. and Rowe, D. (2002) Analysis of inhibitory action of modified U1 snRNAs on target gene expression: discrimination of two RNA targets differing by a 1 bp mismatch. *Nucleic Acids Res*, **30**, 2329-2339.

123. Vagner, S., Ruegsegger, U., Gunderson, S.I., Keller, W. and Mattaj, I.W. (2000) Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA*, **6**, 178-188.

124. Abad, X., Vera, M., Jung, S.P., Oswald, E., Romero, I., Amin, V., Fortes, P. and Gunderson, S.I. (2008) Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Res*, **36**, 2338-2352.

125. Blazquez, L., Gonzalez-Rojas, S.J., Abad, A., Razquin, N., Abad, X. and Fortes, P. (2012) Increased in vivo inhibition of gene expression by combining RNA interference and U1 inhibition. *Nucleic Acids Res*, **40**, e8.

126. Abad, X., Razquin, N., Abad, A. and Fortes, P. (2010) Combination of RNA interference and U1 inhibition leads to increased inhibition of gene expression. *Nucleic Acids Res*, **38**, e136.

127. Kato, K., Hitomi, Y., Imamura, K. and Esumi, H. (2002) Hyperstable U1snRNA complementary to the K-ras transcripts induces cell death in pancreatic cancer cells. *Br J Cancer*, **87**, 898-904.

128. Fernandez Alanis, E., Pinotti, M., Dal Mas, A., Balestra, D., Cavallari, N., Rogalska, M.E., Bernardi, F. and Pagani, F. (2012) An exon-specific U1 small nuclear RNA (snRNA) strategy to correct splicing defects. *Hum Mol Genet*, **21**, 2389-2398.

129. Rogalska, M.E., Tajnik, M., Licastro, D., Bussani, E., Camparini, L., Mattioli, C. and Pagani, F. (2016) Therapeutic activity of modified U1 core spliceosomal particles. *Nat Commun*, **7**, 11168.

130. Donadon, I., Pinotti, M., Rajkowska, K., Pianigiani, G., Barbon, E., Morini, E., Motaln, H., Rogelj, B., Mingozzi, F., Slaugenhaupt, S.A. *et al.* (2018) 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.

131. Sajic, R., Lee, K., Asai, K., Sakac, D., Branch, D.R., Upton, C. and Cochrane, A. (2007) Use of modified U1 snRNAs to inhibit HIV-1 replication. *Nucleic Acids Res*, **35**, 247-255.

132. Knoepfel, S.A., Abad, A., Abad, X., Fortes, P. and Berkhout, B. (2012) Design of modified U1i molecules against HIV-1 RNA. *Antiviral Res*, **94**, 208-216.

133. Mandal, D., Feng, Z. and Stoltzfus, C.M. (2010) Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J Virol*, **84**, 12790-12800.

Chapter 2: Manuscript

Title: Characterization of novel and potent U1 interference RNAs targeting HIV-1

In accordance with the "guidelines for thesis preparation", this thesis will be presented in a manuscript-based format.

Olivier Del Corpo[1, 2], Ryan Goguen[1, 3], Camille Malard[1, 3], Aïcha Daher[1], Susan Colby-Germinario[1], Robert J. Scarborough[1,3*] and Anne Gatignol[1, 2, 3*]
[1] Lady Davis Institute for Medical Research, Montreal, Quebec, H3T 1E2, Canada
[2] Department of Medicine, Division of Experimental Medicine, McGill University, Montreal, Quebec, H3A 0G4, Canada
[3] Department of Microbiology and Immunology, McGill University, Montreal, Quebec, H3A 0G4, Canada

This manuscript is currently in preparation for submission to Nucleic Acids Research

Contribution of the authors:

ODC, RJS and AG conceived and designed the experiments

ODC performed experiments represented in Figures 1 to 7.

RG helped to set-up conditions and performed a Northern blot similar to Figure 4A

CM performed cloning of the lentiviral vectors used in Figure 7

AD and SCG provided technical guidance throughout the project

RJS and AG supervised the project

ODC, RJS and AG wrote the manuscript

*Co-corresponding authors

* To whom correspondence should be addressed. Email: robert.scarborough@mail.mcgill.ca or anne.gatignol@mcgill.ca

Rationale for manuscript:

Several antiviral small RNA classes have been proposed for use in anti-HIV gene therapy with some even advancing to clinical trials. However, research has indicated that the current combinations of small RNA molecules in clinical trials are suboptimal and there is still a need to identify novel antiviral RNA molecules. U1i RNAs are attractive candidates to be used in gene therapy against HIV-1 as they have been demonstrated to be synergistic when expressed with other U1i RNAs or shRNAs and are potent inhibitors of HIV-1 replication. To date, there has been three different studies utilizing U1i RNAs to target and inhibit HIV-1 replication. However, there is a lack of data comparing the efficacy of anti-HIV U1i RNAs targeting the 3'-terminal exon to those that target upstream of 3'ss. Our goal was to compare top U1i RNA candidates from these studies in order to identify molecules to be further researched for gene therapy. Furthermore, the lab has previously characterized a site in the Gag open reading frame that when targeted by an anti-HIV-1 ribozyme and shRNA resulted in a drastic decrease in HIV-1 replication. We hypothesized that U1i RNAs targeting this site in Gag can be effective suppressors of viral replication. We also hypothesized, that by comparing different anti-HIV-1 U1i RNAs, we can identify a set of highly potent and non-toxic molecules to be used in future studies. The aims of this study were:

1. Construct and test U1i RNAs targeting the Gag open reading frame for their ability to inhibit HIV-1 replication.
2. Compare top U1i RNAs that induce excessive splicing to those that inhibit polyadenylation in terms of inhibition of viral replication, effects on viral RNA, toxicity and protein requirements.
3. Elongate the recognition domain of U1i RNAs and determine the effect on the efficacy of these molecules as well as their specificity.

Characterization of novel and potent U1 interference RNAs targeting HIV-1

ABSTRACT:

U1 interference (U1i) RNAs can be designed to correct splicing defects and target pathogenic RNA, such as HIV-1 RNA. They represent promising candidates for HIV-1 gene therapy and several have been designed to target HIV-1 RNA through splicing enhancement or inhibition of mRNA polyadenylation. Here we show for the first time that effective U1i RNAs can be designed to inhibit HIV-1 production by targeting sites upstream from the first splice acceptor site in the Gag coding region of HIV-1 RNA. One Gag U1i RNA was as effective at inhibiting HIV-1 production as some of the top U1i RNAs that enhance HIV-1 RNA splicing and more effective than top U1i RNAs that inhibit polyadenylation of HIV-1 RNA. Our results suggest that the U1i RNA targeting Gag acts through a mechanism distinct from splicing enhancement or inhibition of polyadenylation. By changing the length of the U1i RNA recognition domains, we show that an increase by 6 nucleotides has little effect on the antiviral efficacy of these molecules. Overall our results indicate that, U1i RNAs acting through mechanisms other than polyadenylation inhibition, represent the best candidates for HIV-1 gene therapy and several optimized candidates from the different classes are described.

INTRODUCTION:

Despite the availability of numerous effective antiretroviral therapies (ARTs), HIV-1 remains a global epidemic and is responsible for a considerable amount of morbidity and mortality. Current HIV-1 treatment consists of a combination of anti-retroviral therapy (cART) which can successfully control HIV-1 replication and prevent the onset of acquired immunodeficiency syndrome (AIDS) (1). While cART has drastically improved the lifespan of individuals infected with HIV, it does have limitations, including drug-related toxicity, patient compliance issues, high costs and emergence of viral resistance (2). Moreover, the persistence of viral latent reservoirs during cART leads to rapid rebound of viremia following treatment cessation requiring cART to be a lifelong therapy (3). Due to these well-known limitations associated with current HIV-1 therapy, a cure for HIV-1 infection remains a highly desirable goal to improve the lives of HIV-1 infected individuals.

In the past 10 years, a substantial amount of research has focused on a gene-based cure for HIV. One approach involves the use of antiviral RNAs such as short hairpin RNAs (shRNAs), ribozymes, decoy and aptamer RNAs, designed to target and reduce HIV-1 replication (4). However, viral escape remains a significant issue (5). As with cART, gene therapy will require a combination of antiviral genes to prevent the development of resistant virus. Although several clinical trials have begun, there still remains a need for the identification and characterization of novel and potent antiviral RNAs.

The U1 small nuclear RNA (U1 snRNA), in complex with seven Smith (Sm) proteins and three U1-specific proteins (U1-70K, U1-A and U1-C), is a fundamental component of the spliceosome, a ribonucleoprotein (RNP) complex that catalyzes precursor mRNA splicing (6). During the early steps of spliceosome assembly, 5’ splice donor sites (5’ss) (CAGGTAAGTA)

of pre-mRNAs are recognized by the U1 snRNA through RNA-RNA interactions with the 5' recognition domain of the U1 snRNA (Figure 1A). U1 snRNP binding along with the recognition of the upstream 3' splice acceptor sites (3'ss) polypyrimidine tract (PPyT) by the U2AF heterodimeric cellular splicing factor and the branch point sequence by branch point binding protein (SF1/mBBP) allows for recruitment of the U2 snRNP and proper formation of the spliceosomes catalytic core. Spliceosomal assembly across exons leads to splicing by a process termed "exon definition" (7,8). The U1 snRNP has also been implicated in repressing 3' end polyadenylation of pre-mRNAs via interactions with *cis*-elements located upstream or downstream of polyadenylation sites (PAS) (9). Inhibition of 3' end processing is mediated by interactions between U1-specific U1-70K protein and the Poly-A-polymerase (PAP) (10). Transcripts that lack a poly(A) tails are inherently unstable and are rapidly degraded by host cell machinery (11).

U1 interference (U1i) is a technique to inhibit the expression of a targeted gene by utilizing the properties of the U1 snRNP to inhibit 3' end polyadenylation when targeting 3'-terminal exons or by promoting splicing when bound to a 5'ss or downstream of a 3'ss by the process of exon definition. Inhibition is achieved by modifying the 5' recognition domain of U1 snRNAs to contain sequences complementary to regions in the terminal exon or downstream of a 3'ss of a targeted transcript (Figure 1B). These modified U1 snRNAs are often referred to as U1i RNAs (12) and some studies have shown that they have a synergistic inhibitory effect on mRNA expression when combined with other U1i RNAs or shRNAs that use the RNAi pathway to target mRNAs (13,14).

To date, there has been three independent studies utilizing U1i RNAs to inhibit HIV-1 replication. Two of these have designed U1i RNAs targeting highly conserved sequences in the

3’ end of HIV-1 RNA and demonstrated that U1i RNAs that inhibit polyadenylation are potent inhibitors of HIV-1 production in cell culture (15,16). The last study demonstrated that U1i RNAs designed to interfere with HIV-1 splicing by targeting 5’ss or downstream of 3’ss were also potent inhibitors of HIV-1 production (17). However, these U1i RNAs have shown mixed efficacy against HIV-1 replication in transduced T lymphocyte cell lines with one study demonstrating zero efficacy against HIV-1 replication (16,17). While all three of these studies identified potent anti-HIV-1 U1i RNAs, there has yet to be a study comparing the efficacy of these various molecules. Moreover, a comparison between U1i RNAs that induce excessive splicing to those that inhibit polyadenylation has not been performed.

In this study, we identify and characterize a novel and potent U1i RNA targeting the Gag open reading frame of HIV-1. After comparing its efficacy to previous U1i molecules targeting splicing or polyadenylation of viral transcripts, we found that U1i RNAs that induce excessive splicing and the U1i targeting Gag are more potent than U1i RNAs that inhibit polyadenylation, furthermore, that they depend on the recruitment of the U1-70K protein. We were able to increase the recognition domains of U1i RNAs with little effect on their antiviral efficacy and found that a U1i RNA that induces excessive splicing inhibits HIV-1 replication in transduced T lymphocytes at levels similar to an shRNA currently in clinical trials for anti-HIV-1 gene therapy.

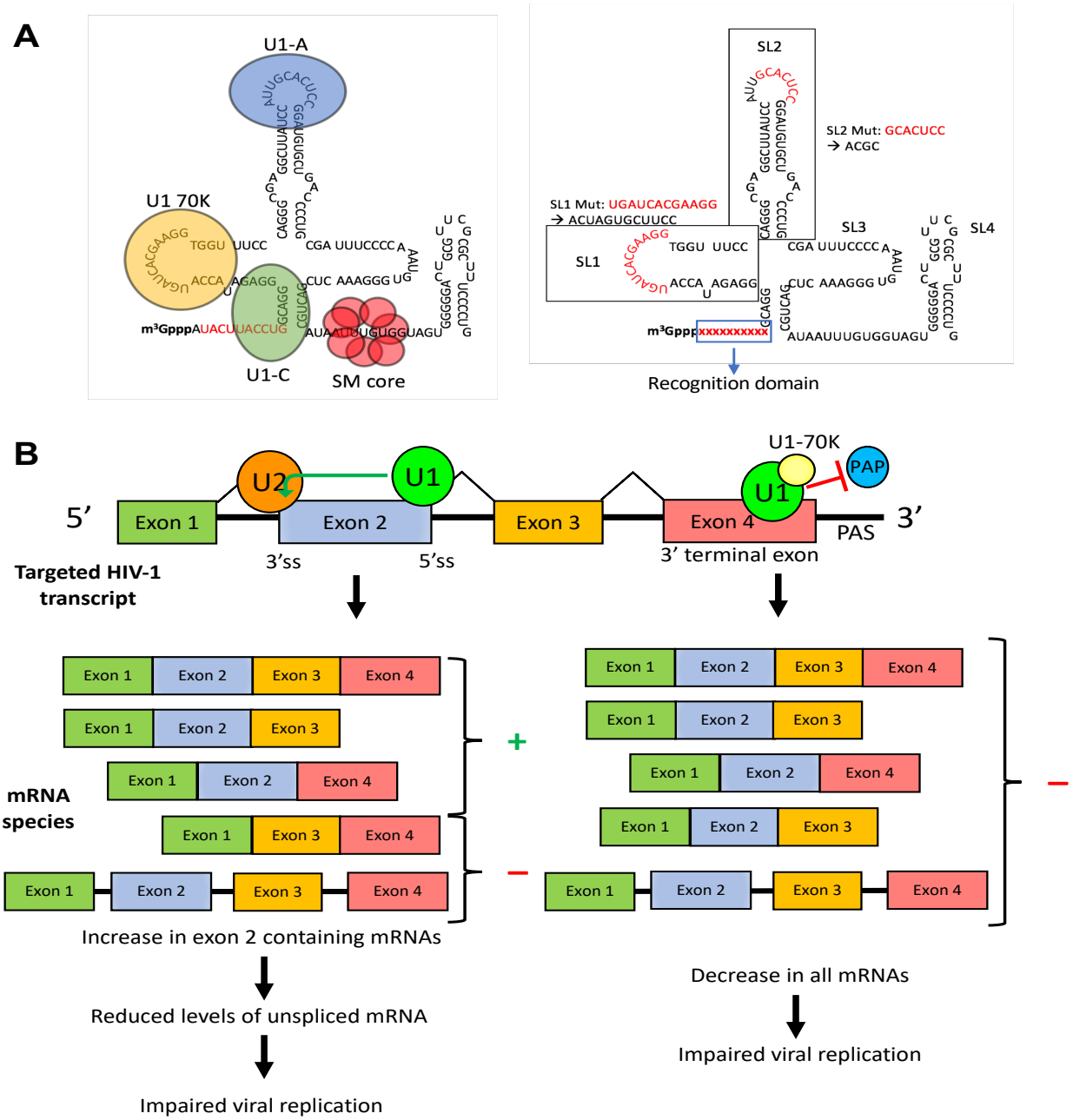

Figure 1. Structure of the U1 snRNP and mechanism of action of U1i RNAs. (A) Left, The U1 snRNP with associated proteins U1-70K, U1-A, U1-C and Sm. Right, a U1i RNA in which the U1 snRNA recognition domain is changed to be complementary to a target RNA sequence. Stem loops (SL) 1 and 2 mutated sequences used for domain mutation experiment are illustrated. (B) Depiction of the mechanism of action of U1i RNAs targeting 5'slice donor sites (5'ss) or 3'- terminal exons of targeted HIV-1 mRNA. Left, U1i RNAs targeted to a 5'ss or downstream of a 3' splice acceptor site (3'ss) enhance splicing at the upstream 3'ss resulting in an increase in mRNA species containing a particular exon and a decrease in unspliced RNA and mRNA species that do not include that particular exon. Right, binding of U1i RNAs to the 3'-terminal exon of mRNAs results in an inhibition of polyadenylate polymerase (PAP) at the polyadenylation site (PAS) and a decrease in all mRNA species.

MATERIAL AND METHODS:

Plasmid constructs

UBC plasmids expressing wild-type (WT) and HIV-1 specific U1 snRNAs (U1-Rev and U1-Env) were provided by Dr. Cochrane (15). All other HIV specific U1 snRNA plasmids were created using polymerase chain reaction (PCR)-mutagenesis on the WT UBC plasmid (U1-WT). The 5' (mutagenic) primers used to create the various U1i RNAs are shown in Table 1. The 3' primer was, 5'-AGTGCCAAGCTTGCATGCCAGCAGGTC-3'. A base change (underlined) was introduced into the 3' primer to remove a Pst I site found in the U1-WT plasmid and allow for Pst I digestion discrimination against WT plasmids. PCR products and the U1-WT plasmid were digested with restriction enzymes Bgl II and Hind III (Fermentas) overnight at 37°C and ligated using T4 DNA ligase (InVitrogen). Ligated products were then used to transform DH5α bacteria and mini-preps were performed using Pure Link HiPure™ Plasmid Miniprep Kit (Invitrogen). Primers for the elongated U1i RNAs are shown in Table S1.

U1 snRNA stem loop 1 and 2 mutants (SL1Mut, SL2Mut) described in (15) were generated with the Q5 site directed mutagenesis kit (New England Biolabs) using the primer pairs provided in supplementary Table S2. For SL1Muts, site directed mutagenesis was performed on the individual U1 plasmids while for SL2Muts site directed mutagenesis was performed on the U1-WT plasmid and the various U1 plasmids were cloned using PCR-mutagenesis, as described above. For all plasmids, correct construction was confirmed by sequencing, using forward (5'-CCCAGTCAGGAGGTTGTAAAACG-3') and reverse (5'-AGGGGATAACAATTTCACACAGG- 3') primers.

Cells and transfections

Transfections were performed as described in (18). Briefly, Human embryonic kidney (HEK) 293T cells (ATCC) were maintained in Dulbecco's Modified Eagle's medium with high glucose (HyClone) containing 10% fetal bovine serum (HyClone) and 1% penicillin/streptomycin (Life Technologies). Cells were seeded at 2.25 x 10^5 cells/ml 24 hrs prior to transfection in 96, 24, 12 or 6 well plates in a volume of 150, 500, 1000 or 2000 µl, respectively. Plasmid transfections were performed using TransIT®-LT1 (Mirus) according to the manufacturer's protocol. For all transfection experiments, assays were performed 48 hrs after transfection.

HIV-1 production assay

Viral production was determined by measuring the activity of HIV-1 reverse transcriptase (RT) in the supernatant of co-transfected cells, as previously described (19). Briefly, HEK293T cells were co-transfected with 100 ng of HIV-1 molecular clone pNL4-3 (Genebank M19921) and different amounts of U1i RNA plasmids in 24 well plates. 48 hrs after co-transfection, 5 µl of supernatant was incubated for 2h in 50 µl of RT cocktail (60 mM Tris-HCl, 75 mM KCl, 5 mM MgCl2, 1.04 mM Ethylenediaminetetraacetic acid, 1% NP-40, 10 µg/ml Poly(A), 0.33 µg/ml oligo dT, 8 mM dithiothreitol (DTT, C4H10O2S2) and [^{32}P] dTTP (3000 Ci/mmol)). 5 µl of the reaction mixture was then spotted onto a glass fiber Di Ethyl Amino Ethyl (DEAE) filter mat (PerkinElmer) and left to dry for 10 min. Filter mats were washed 5 x 5 min with 2X saline sodium citrate (SSC) buffer (20X SSC buffer: 3M NaCl and 0.3M sodium citrate), followed by 2 x 1 min washes in 95% ethanol. Counts per minute (cpm) were measured using a microplate scintillation counter (MicroBeta TriLux). RT data were normalized to cells co-transfected with negative control plasmids (U1-WT, U1-SL1-WT, U1-SL2-WT).

Cell viability assay

Cell viability was estimated by measuring the metabolism of WST-1 (4-[3-(4-Iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolio]-1,3-benzene disulfonate) (Sigma-Aldrich) to formazan (dark red) in HEK293T cells transfected with U1 plasmids in 96 well plates. As a positive control for cell toxicity, 10 mM of 3% w/v H_2O_2 (Sigma-Aldrich) was added to control wells at the time of transfection. Forty-eight hours after transfection, 100 µl of cell culture media was removed from each well and 20 µl of WST-1 reagent was added. Cells were then incubated for 4 hrs at 37°C and 5% CO_2. Plates were read at 450 nm using a Benchmark Plus Microplate spectrophotometer (BioRad). All results were normalized to cells transfected with U1-WT.

Sequence conservation estimates

The QuickAlign tool on www.hiv.lanl.gov/content/index was used to estimate the conservation at the nucleotide level of sites targeted by U1i RNAs among circulating HIV-1 strains as previously described (19,20). Briefly, sequence alignments containing all complete HIV-1 genome sequences (3666 at the time of analysis) were generated using the QuickAlign tool. Jalview alignment editor (Version 2.9) was used to calculate the % conservation for each nucleotide in the selected target site expressed as the % of sequences containing the consensus nucleotide at each position.

HIV intracellular RNA expression

HEK293T cells were transfected with 400 ng of HIV-1 pNL4-3 and 1000 ng of U1 plasmids in 6 well plates. 48 hrs after co-transfection, cell lysates were harvested in TRIzol (Invitrogen). Total RNA was isolated using phenol chloroform extraction followed by clean-up using RNeasy Mini Kits (QIAGEN). 20 µg of RNA was resolved on a 1% agarose gel (1 g agarose in 84.6 ml water, 10 ml of 10 x MOPS buffer (200 mM MOPS, 50mM sodium adetate,

10 mM EDTA, 0.1% DEPC, pH 7.0), 5.4 ml of 37% formaldehyde; Emd Millipore and 10 µl RedSafe™) and transferred to a Hybond-N membrane (GE Healthcare Amersham). Equal loading was determined by UV visualization of the agarose gel prior to the transfer. Blots were hybridized with DIG-Labelled DNA probes targeting the HIV-1 LTR (generated by PCR using the primers 5'-CTAATTCACTCCAAAGAAGA-3' and 5'-TGCTAGAGATTTTCCACACTG-3'). DIG-labelling, hybridization, washes and visualization of RNA were performed as in the DIG DNA Labelling and Detection Kit protocol (Roche).

RT-PCR analysis of viral RNA species

Total RNA was isolated as indicated above. One microgram of DNase treated RNA was used for cDNA synthesis using ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs). Viral cDNA was subsequently amplified using forward primer 5'-CTGAGCCTGGGAGCTCTCTGGC-3' and reverse primer 5'-TCATTGCCACTGTCTTCTGCTCT-3' for 4 kb viral RNA and forward primer 5'-CTGAGCCTGGGAGCTCTCTGGC-3' and reverse primer 5'-CCGCAGATCGTCCCAGATAAG-3' for 2 kb viral RNAs. The PCR products were resolved on a 3% agarose gel (25 µl of RedSafe™) and visualized with UV light.

HIV-1 infection assay

U1i RNA expression cassettes from this study and shRNA expression cassettes generated in our previous study (20) were sub-cloned into HIV-7-eGFP plasmid (donated by Dr. Rossi) (21). Lentiviral vectors were produced by co-transfecting the HIV-7-eGFP plasmids with a VSV-G expression plasmid (from Dr. Rossi) and a 2nd generation packaging plasmid (pax2, addgene #12260) into HEK293T cells. The supernatants were collected 48 hrs later and the lentiviral particles were concentrated using Lenti-X (Clontech) according to the manufacturers protocol.

Lentiviruses were tittered on SupT1 cells using the % GFP-positive cells as a readout and a multiplicity of infection (MOI) of 4 was used to infect SupT1 cells for the HIV-1 infection assay. 48 hrs after infection with lentiviruses, cells we resorted for GFP expression. 24 hrs after, sorted cells were plated at 1 X 10^5 cells/well in a 96 well plate and infected with HIV-1 in triplicate. HIV-1 replication was monitored by measuring HIV-1 RT activity in cultured supernatants of infected cells.

RESULTS:

U1i RNAs targeting the Gag coding sequence of HIV-1 RNA are effective inhibitors of HIV-1 production.

For antisense-based molecules targeting HIV-1 RNA, an important consideration for their development as therapies is that they target highly conserved sequences. In a previous study, we identified a conserved target site in the Gag coding sequence of HIV-1 RNA that was accessible to inhibition by a ribozyme and an shRNA (20). To determine whether the target sequence is also accessible to inhibition by U1i RNAs, four U1i RNAs (U1-Gag1, 2, 3, and 4) were designed, targeting overlapping regions within the Gag target site (Table 1, Figure 2A). The effects of U1i RNAs on HIV-1 production were evaluated by co-transfection of HEK293T cells with HIV-1 molecular clone pNL4-3 and increasing concentrations of U1i RNA constructs. HIV-1 production was compared by measuring the activity of the HIV-1 RT enzyme in culture supernatants, and data were expressed as a percentage of RT activity in cells co-transfected with the WT U1 snRNA expressing construct. Out of the four U1i RNAs targeting Gag, U1-Gag1 and U1-Gag2 dose dependently inhibited HIV-1 production, with U1-Gag2 providing the most potent inhibition (Figure 2B). The Gag target site is not downstream of a splice acceptor site, nor is it in the 3'-terminal exon, suggesting that U1-Gag1 and U1-Gag2 inhibit HIV-1 production through a mechanism distinct from previously designed U1i RNAs targeting HIV-1 replication (15-17).

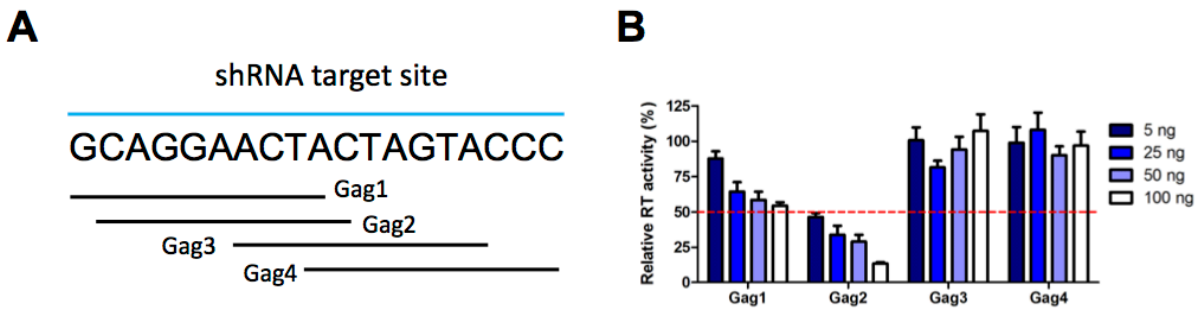

Figure 2. Noncanonical U1i RNAs targeting the Gag open reading frame inhibit HIV-1 production. (A) Outline of the site targeted by constructed U1i RNAs as well as an shRNA described in our previous study. (B) Inhibition of HIV-1 production by U1-Gag1, U1-Gag2, U1-Gag3 and U1-Gag4 when transfected at 5, 25, 50, 100 ng. The dotted line in the graph represents the level at which 50% of HIV-1 production is inhibited. Results are expressed as a relative percentage of HIV-1 inhibition to the U1-WT transfected cells. All data are represented as the mean +/- standard error means (SEM) of at least three independent transfections with 3 replicates (n = 9).

U1-Gag2 is more effective compared to top U1i RNAs targeting polyadenylation and has similar effects compared to top U1i RNAs that enhance HIV-1 RNA splicing.

To directly compare the effects of U1-Gag2 to other U1i RNAs targeting HIV-1 RNA we constructed the top three U1i RNAs designed to inhibit polyadenylation, identified in the initial U1i RNA screen (15), and the top five candidates designed to induce excessive splicing of HIV-1 RNA (17). U1i RNAs that inhibit polyadenylation all target the 3' proximal region of HIV-1, while those that induce excessive splicing target 5'ss D2 or D3 and sequences downstream of the first, second and third 3'ss, A1, A2 and A3. The target sequences for the U1i RNAs are provided in Table 1 and their positions within the HIV-1 RNA open reading frames are illustrated in Figure 3A. All U1i RNAs constructed inhibited HIV-1 production in a dose dependent manner (Figure 3B). However, U1-Gag2 and U1i RNAs designed to elicit excessive splicing (U1-D2, D3, T3, T4, T6) were found to be the most potent candidates, providing a 50 % inhibition of viral production at around 5 ng or less of input DNA, compared to around 25 ng for U1i RNAs that inhibit polyadenylation (U1-Env, Ref, Nef). Of the U1i RNAs that inhibited polyadenylation, only U1-Env was able to provide a near complete inhibition of viral production at higher doses, similar to U1-Gag2 and U1i RNA molecules that induce excessive viral RNA splicing. These results suggest that U1i RNAs acting through mechanisms other than polyadenylation inhibition are more potent at inhibiting HIV-1 production. Moreover, that U1-Gag2 is a competitive candidate for further development.

Using the HIV Los Alamos database and methods previously developed in our lab (19,20), we estimated the conservation at the nucleotide level of sites targeted by U1i RNAs among all complete HIV-1 sequences available in the database (3666 at the time of analysis). Sites targeted by U1-D2, U1-D3 and U1-Gag2 had the highest conservation among the HIV-1

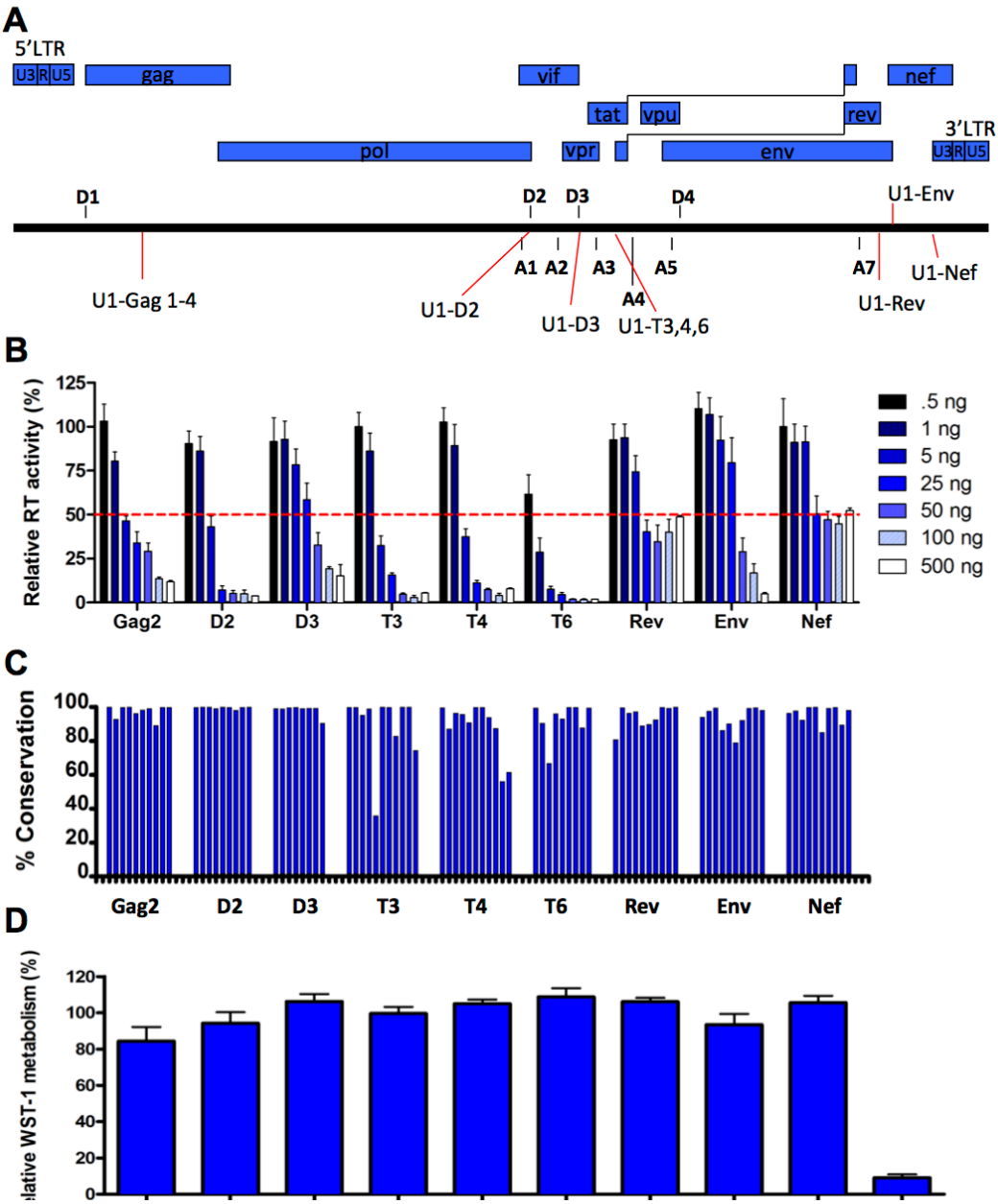

Figure 3. U1i RNAs that induce excessive splicing and U1-Gag2 are more potent inhibitors of HIV-1 when compared to U1i RNAs that inhibit polyadenylation. (A) Outline of the HIV-1 genome and organization of the major splice donors and acceptors. (B) Inhibition of HIV-1 production by U1i RNAs at 0.5, 1, 5 ,25, 50 and 100 ng. The dotted line in the graph represents the level at which 50% of HIV-1 production is inhibited. Results are expressed as a relative percentage of HIV-1 inhibition to U1-WT co-transfected cells. (C) Conservation estimates of sequences targeted by the various U1i RNAs. (D) Toxicity of U1i RNA molecules. HEK293T cells were mock transfected with U1-WT or transfected with 1000 ng/ml of U1i RNAs targeting Gag, U1i RNAs that induce excessive splicing and U1i RNAs that inhibit polyadenylation. Cells treated with H_2O_2 were used as a positive control. Cell viability was estimated by measuring the metabolism of WST-1 48 hrs post transfection. Data are expressed relative to U1-WT treated cells (%). All data are represented as the mean +/- SEM from three experiments with three replicates (n = 9).

strains in the database with over 90% conservation at each nucleotide (Figure 3C). Sites targeted by U1-Env, U1-Rev, U1-Nef and U1-T6 had an intermediate level of conservation with most nucleotides conserved in greater than 80% of the strains and sites targeted by U1-T3 and U1-T4 had the lowest conservation, with several nucleotides conserved in less than 60% of the strains (Figure 3C).

A significant concern for the development of antiviral RNAs for use in gene therapy is their potential to cause cellular toxicity, as this can alter the interpretation of experimental results and limit their potential for use in clinical settings. To evaluate whether the reduction of HIV-1 production seen when expressing U1i RNAs is due to toxicity in HEK293T cells, we assessed the ability of U1i RNAs to affect cell viability. HEK293T cells were transfected with U1i RNAs at 1000 ng/ml of cultured supernatant, a dose several folds higher than their maximal inhibitory concentrations. Cell viability was estimated using the WST-1 colorimetric assay and data were expressed as relative WST-1 metabolism compared to cells transfected with the U1-WT construct. While treatment of cells with 10 mM of H_2O_2 resulted in significant cell death, there were no significant differences between U1i RNA-transfected and U1-WT transfected cells (Figure 3D). Based on the lack of cellular toxicity, conservation of their target sites and their ability to inhibit HIV-1 production, U1-Gag2, D2, T6, Rev, Env and Nef were selected for further evaluation.

U1i RNA expression alters HIV-1 RNA accumulation

To determine whether the U1-Gag2 inhibited HIV-1 production by enhancing HIV-1 RNA splicing, we next compared the effects of the selected U1i RNAs on HIV-1 RNA accumulation by Northern blot following co-transfection with pNL4-3 in HEK293T cells (Figure 4A). Consistent

with previous results (17), co-transfection with U1-D2 and U1-T6 resulted in a near complete reduction of full-length (FL) HIV-1 RNA and an increase in incompletely spliced (IS) or both IS and completely spliced (CS) RNA species. In contrast, U1-Gag2, U1-Rev and U1-Nef expression resulted in a decrease of all RNA species, suggesting that these molecules do not enhance HIV-1 RNA splicing. Unexpectedly, U1-Env expression resulted in an increase in CS RNA, suggesting that at least part of its effect on HIV-1 production can be attributed to its ability to enhance HIV-1 RNA splicing.

To further explore the effects of U1i RNAs on HIV-1 RNA expression, we next constructed cDNA libraries from the RNA extracted from co-transfected HEK-293T cells and used previously described primer sets (17,22) to amplify CS and IS cDNAs. The position of these primers and major exons used for the generation of HIV-1 RNA splice variants are illustrated in Figure 4B. The products of the amplification were run on an agarose gel and visualized with UV light (Figure 4C). Consistent with previous results (17), U1-D2, targeting the 5'ss D2, increased inclusion of exon 2 or exon 2E. This is shown by the increase in CS mRNA species 1.2.5.7 (*nef*3*)* and IS 1.2E (*vif*E*)* and 1.2.5E (*env*5) and a decrease in other viral mRNA species. Also consistent with the previous study, expression of U1-T6 resulted in increased inclusion of exon 4 or exon 4E. This is demonstrated by the increase in CS species 1.4.7 (*Tat1*), 1.2.4.7 (*Tat*2)*,* and 1.3.4.7 (*Tat*3), as well as IS spliced species 1.4E (*Tat*5). As seen with the Northern blot results, U1-Env had major effects on the splicing pattern of HIV-1 RNA. These results suggest that like U1-T6, U1-Env acts through increasing inclusion of exon 4 and to a lesser extent, exon 4E. This is shown by the increase expression of CS species 1.4.7 (*Tat1)* and 1.3.4.7 (*Tat*3*)* and IS species 1.4E (*Tat*5) (Figure 4C). Although there were no apparent effects of U1-Rev on splicing in the Northern blot (Figure 4A), a small effect was detected when we looked at the IS species, where a modest increase in 1.4E was observed, suggesting that, like U1-

Env and U1-T6, U1-Rev can lead to an increase in the inclusion of exon 4E. These results demonstrate that some U1i RNAs targeting the terminal exon of HIV-1 RNA can enhance HIV-1 RNA splicing and that some of their effects on HIV-1 production can be attributed to this mechanism. The results also demonstrate that enhanced HIV-1 RNA splicing does not contribute to the inhibition of HIV-1 production provided by U1-Gag2 and U1-Nef.

HIV-1 suppression by all U1i RNAs is dependent on the proper function of the SL1 domain.

The endogenous U1 snRNA molecule exerts its function in the form of a ribonucleoprotein (RNP) complex that consists of three specific U1 proteins, U1-A, U1-70K and U1-C, and 7 Sm proteins (23). To determine whether the inhibition of HIV-1 production by the U1i RNA molecules was dependent upon proper assembly of the U1 snRNP, we mutated the stem loop 1 (SL1) domain of all the U1i RNAs in a manner that results in a loss of U1-70K binding and the stem loop 2 (SL2) domain in a fashion that results in the loss of U1-A binding (as previously demonstrated (15)). The mutated sequences are illustrated in Figure 1A. While, U1-70K binding and to a lesser extent U1-A binding is required for inhibition by U1i RNAs that inhibit polyadenylation, the U1 snRNP domain requirements of U1i RNAs that induce excessive splicing have yet to be determined (10,15).

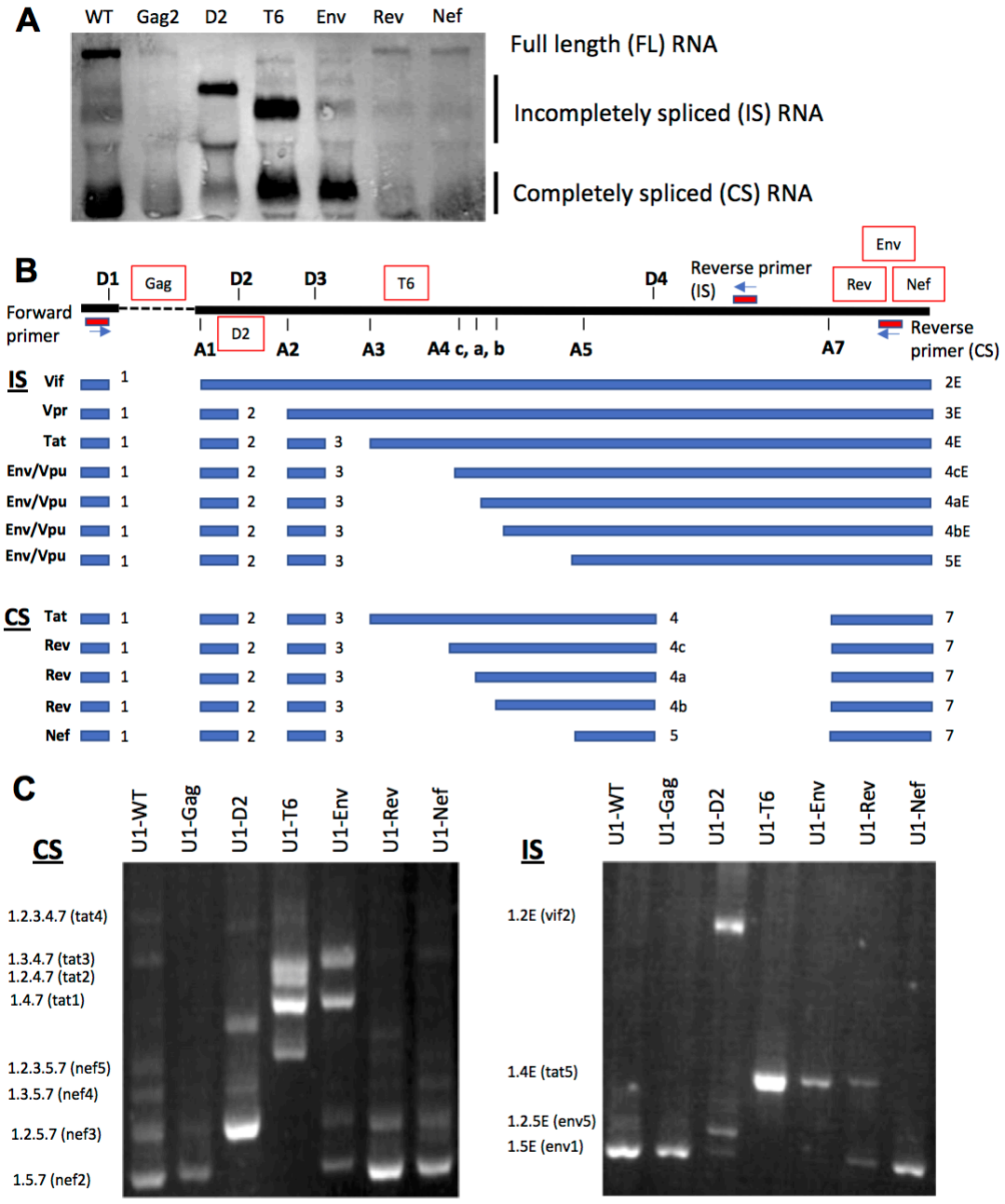

Figure 4. The effects of U1i RNA expression on HIV-1 mRNA. (A) Total RNA isolated from HEK293T cells transfected with 400 ng of pNL4-3 in the presence or absence of U1i RNAs (1000 ng) were analyzed by Northern blotting. Full-length (FL), incompletely spliced (IS) and completely spliced (CS) RNA species are indicated. (B) The numerous HIV-1 mRNA species produced by alternative splicing at the various splice donor (D1-5) and splice acceptor (A1-7) sites are indicated. Primers used for RT-PCR analysis are depicted in red for IS and CS. (C) RT-PCR analysis of completely spliced (1.8 kb) and singly spliced (4.0 kb) RNA species in the presence of U1i RNAs or U1-WT expressing plasmid.

The effects of the SL1 and SL2 mutants on HIV-1 production were compared with the nonmutated versions of the selected U1i RNAs at different doses (Figure 5). Inhibition of HIV-1 production by all U1i RNAs was found to be dependent on the binding of U1-70K to the SL1 domain as little to no inhibition was observed for the SL1 mutants at all doses evaluated (Figure 5A-F). These results suggest that the inhibition of HIV-1 production by all U1i RNAs evaluated

was dependent on the formation of a U1 snRNP complex and not the result of antisense RNA effects alone. At higher doses, SL1 mutants U1-T6 and U1-D2 had some effects on HIV-1 production (Figure 5 A, B). This could be the result of antisense effects or may be related to competitive inhibition of 3'ss or 5'ss by mutated U1i RNAs, impeding the normal binding and subsequent splicing by the spliceosome complex. Mutation of the SL2 domain had little to no effect on U1-D2 and U1-T6 that were designed to induce excessive splicing (Figure 5A, B), suggesting that binding of the U1 SL2 loop to the U1-A protein is dispensable for their effects. In contrast, SL2 mutations abolished the inhibitory effects of U1i RNAs designed to inhibit polyadenylation (U1-Rev and U1-Nef) (Figure 5C, D). Although U1-Env was designed to inhibit polyadenylation, our results in Figure 4 demonstrated that it had major effects on HIV-1 RNA splicing. Consistent with a predominant effect on HIV-1 production through splicing enhancement, mutation of the SL2 domain did not have major effects on inhibition provided by U1-Env (Figure 5E). Although it was demonstrated that the U1-Gag2 does not enhance HIV-1 RNA splicing (Figure 4), the SL2 mutant of this U1i RNA also had similar effects compared to the nonmutated version (Figure 5F), suggesting that binding to the U1-A protein is also dispensable for its mechanism of inhibition.

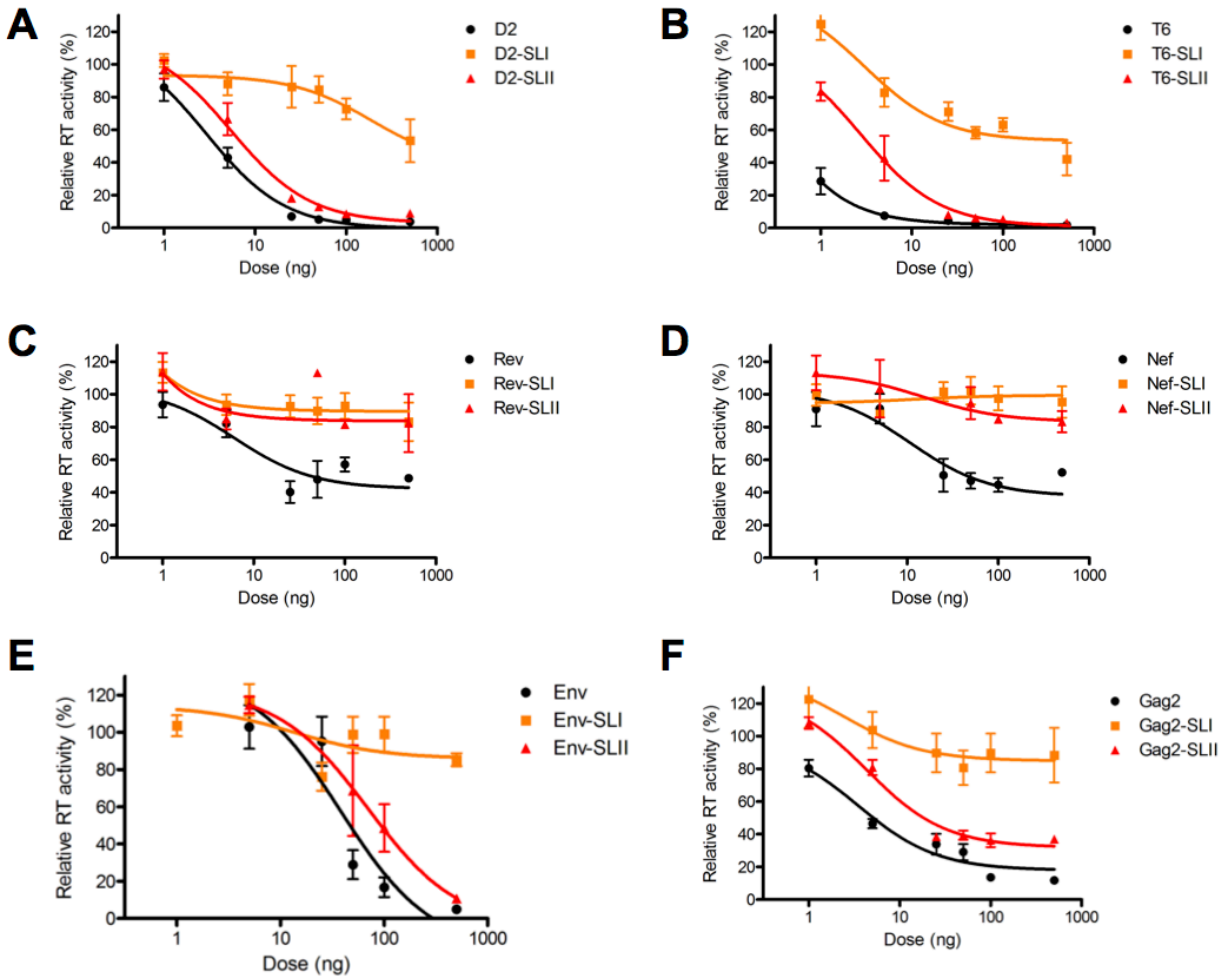

Figure 5. HIV-1 suppression by all U1i RNAs is dependent on the proper assembly of the SL1 domain and recruitment of U1-70K while only U1i RNAs that inhibit polyadenylation depend on the SL2 domain and recruitment of U1-A. (A-F) The effects of U1 snRNA SL1 and SL2 domain mutations on the suppression of HIV-1 production. Dose response curves were created of HEK293T cells co-transfected with 1- 500 ng of mutated U1i RNAs and 100 ng of pNL4-3. RT activity measurements were log transformed and a nonlinear regression log(inhibitor) versus response equation with least squares (ordinary) fit was determined using GraphPad Prism. All data-points are represented as the mean +/- SEM of at least 3 independent transfections with 3 replicates (n = 9).

U1i RNAs with increased recognition domains maintain their antiviral efficacy

U1i RNAs typically have recognition domains of 10-11 nucleotides and increasing its length would improve the specificity by decreasing off-target effects. Increasing the recognition domain towards the 3' end of the molecule by 6 nucleotides did not severely compromise the inhibitory effect of U1i RNAs in one study and could improve their specificity for their intended target sites (24). In contrast, in another study, extending the recognition domain towards the 5' end of the U1i RNA molecule abolished or severely limited its inhibitory activity of U1i RNAs except for molecules that had a recognition domain increased to 16 nucleotides long towards the 5' end (25). To evaluate the impact of increasing the recognition domain on the selected U1i RNAs targeting HIV-1 production, we increased the recognition domains by 3 or 6 nucleotides towards the 3' or 5' end of the viral genome. These length variants were named according to the direction of the increased recognition domain (towards the 5' or 3' end of the viral genome) and by the number of nucleotides added (3'-3, 3'-6, 5'-3 and 5'-6) and were compared to the original U1i RNAs for effects on HIV-1 production (Figure 6). U1i RNAs that induce excessive splicing (U1-D2, U1-T6) maintained their ability to inhibit viral production even when the recognition domain was increased by six nucleotides towards the 3' or 5' end of the viral genome (Figure 6A, B). For U1-Rev, increases of three nucleotides had a negative impact on inhibitory activity, whereas increases of six nucleotides did not have any major effects (Figure 6C). While an increase of three nucleotides on the 5' end of the U1-Nef target site had a negative impact on its inhibitory activity, an increase of six nucleotides resulted in a more effective inhibition (Figure 6D). Increases in the recognition domain of U1-Env towards the 5' end of the viral genome abolished its ability to inhibit HIV-1 production, while increasing it at the 3' end of the viral genome had no effect (Figure 6E). Lastly, increasing the recognition domain in either

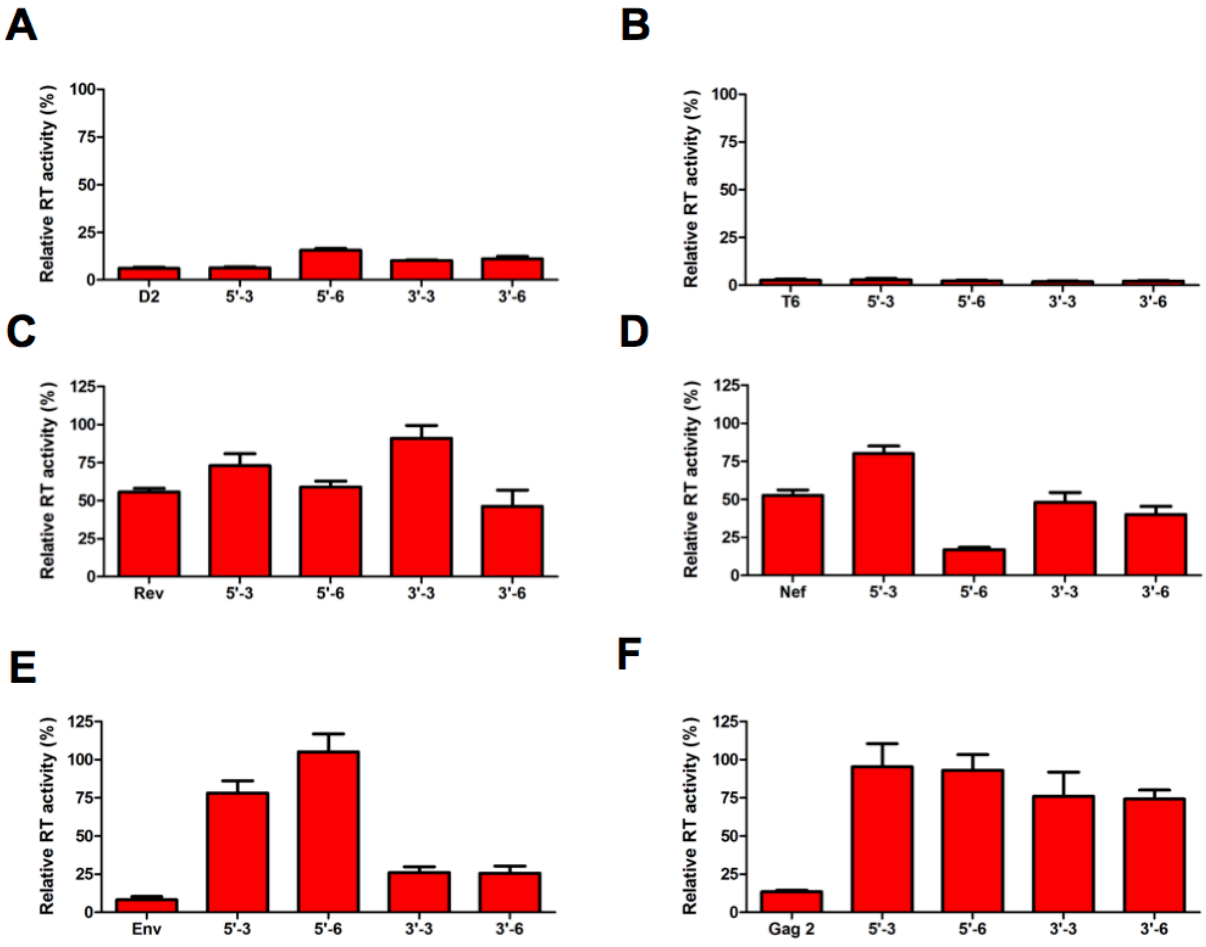

Figure 6. Increasing the recognition domain of U1i RNA molecules is feasible with minimal loss to their inhibitory activity. (A-F) The recognition domains of U1i RNAs were increased by 3 or 6 nucleotides going towards the 3' or 5' end of the viral genome. These length variants are named according to the direction of the increased recognition domain (towards the 5' or 3' end of the viral genome) and by the number of nucleotides added (3'-3, 3'-6, 5'-3 and 5'-6). HEK293T cells were co-transfected with the various constructs and pNL4-3 at 100ng. As a comparison cells were transfected with non-elongated constructs. Results are expressed as a relative RT percentage of HIV-1 inhibition to the WT-pUCB co-transfected cells. All data are represented as the mean +/- SEM from three experiments with three replicates (n = 9).

direction for U1-Gag2 eliminated its ability to inhibit viral replication (Figure 6F). These results suggest that, with the exception of U1-Gag2, increasing the recognition domain of U1i RNAs targeting HIV-1 production can be done with negligible impacts on antiviral effects and in one case (U1-Nef 5'-6) it can even improve the antiviral effects.

Cells transduced with U1-T6 support lower levels of HIV-1 replication than cells transduced with a short hairpin RNA in clinical trials.

To evaluate the potential of U1i RNAs that inhibit polyadenylation and those that enhance HIV-1 RNA splicing to inhibit HIV-1 replication, we used an HIV-based lentiviral vector expressing an enhanced green fluorescent protein (eGFP) reporter gene (HIV-7) (21), to transduce SupT1 cells with cassettes expressing U1-WT, U1-Nef, U1-T6 and an shRNA targeting the tat/rev region of HIV-1 RNA (shRNA-tat/rev). Previous U1i RNA studies have shown either no efficacy or various efficacies to inhibit viral replication over an 8 to 10 day period (16,17). Following transduction, cells were sorted for GFP expression to obtain similar populations of transduced cells and eliminate high and low expressing cells. The sorted cells were infected with HIV-1 NL4-3 and viral replication was monitored by measuring HIV-1 RT activity in culture supernatants (Figure 7). In agreement with the previous study (16), cells transduced with U1-Nef, which our results suggest acts exclusively through inhibition of polyadenylation (Figure 4), supported viral replication with similar kinetics as the empty vector (HIV-7) and U1-WT transduced cells. In contrast, U1-T6 transduced cells had very low levels of viral replication up to 10 days post infection, with even lower levels compared to shRNA-tat/rev transduced cells (Figure 7). These results suggest that U1i RNAs that induce excessive splicing, such as U1-T6, are highly effective at restricting HIV-1 replication in T cells and may be even more effective compared to an shRNA found in clinical trials (26).

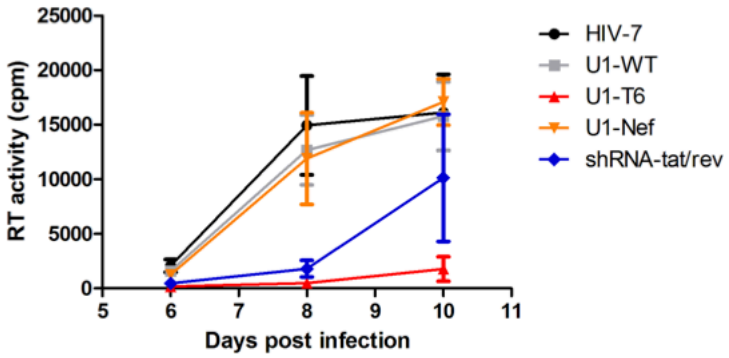

Figure 7. Inhibition of HIV-1 replication in T lymphocytes expressing U1i RNA molecules. SupT1 cells were transduced with HIV based lentiviral vectors (HIV-7-eGFP) expressing U1-WT, U1-T6, U1-Nef and shRNA-tat/rev at an MOI of 4. As a negative control cells were transfected with empty HIV-7-eGFP. The cell lines were then infected with NL4-3 and viral replication was monitored for 10 days by measuring HIV-1 RT activity in cultured supernatant. The mean RT activity (cpm) in cultured supernatants across two independent infections performed in triplicates (n = 6) is shown for the 10 days following infection.

DISCUSSION:

Similar to current antiretroviral therapy, antiviral gene therapy requires a combinatorial and multiplexing approach to reduce the likelihood of viral resistance. While several RNA based antiviral genes are being tested in pre-clinical and clinical settings (26-30), saturation of cellular pathways, suboptimal antiviral inhibition and cellular toxicity remain significant concerns (31-33). Modified U1 snRNAs, termed U1i RNAs, have been demonstrated to be an attractive candidate for gene therapy targeting HIV-1. Multiple U1i RNA molecules targeting the 3'-terminal exon of HIV-1 transcript were demonstrated to be potent inhibitors of HIV-1 production, although no resistance to HIV-1 replication was observed upon transduction into T lymphocytes. U1i RNAs targeting HIV-1 5'ss and sites downstream from 3'ss, induced excessive splicing of HIV-1 RNA inhibited HIV-1 production and one reduced viral replication in T lymphocytes. Our aim was to construct novel U1i RNAs targeting the Gag 1498 site characterized in our previous studies (20,34), to compare its efficacy to U1i RNAs targeting the HIV-1 3'-terminal exon and those enhancing HIV-1 RNA splicing to determine the best molecule to use in gene therapy setting.

Our first goal in this study was to determine if U1i RNAs targeting a highly conserved site in the Gag open reading frame could inhibit HIV-1 replication. We constructed four U1i RNA molecules. While most demonstrated some ability to inhibit HIV-1 replication, U1-Gag2 was found to be the most potent (Figure 2). Expression of U1-Gag2 reduced all HIV-1 RNA isoforms as seen by Northern blotting and RT-PCR analysis (Figure 4A, C) suggesting that it does not act through enhancing HIV-1 RNA splicing. The viral inhibition seen in HEK293T cells transfected with U1-Gag2 was dependant on SL1 sequence integrity, strongly suggesting that proper assembly of an RNP through binding of the 70K protein to the SL1 domain was required for function rather than an antisense mechanism alone (Figure 5F). While it is possible that U1-

Gag2 could be acting through inhibition of polyadenylation, the fact that it was not dependent on binding of the U1-A protein to the SL2 domain, whereas both polyadenylation inhibitors U1-Rev and U1-Nef were, suggests that it acts through a distinct mechanism. In addition, of all the U1i RNAs evaluated, only U1-Gag2 completely lost its activity when its recognition domain was increased by three or six nucleotides in either direction along its target site (Figure 5). Altogether our results suggest that U1-Gag2 is acting through a mechanism distinct from polyadenylation inhibition or enhancement of HIV-1 RNA splicing.

A possible explanation is that U1-Gag2 sequesters U1 RNA proteins resulting in suboptimal splicing at the D1 splice donor site and a subsequent decrease in all viral RNA isoforms. Another mechanism can be that U1-Gag2 is interfering with D1 splice site utilization by inhibiting the action of cellular splice enhancer factors. This hypothesis is compatible with intronic splicing enhancer sequences located in the Gag open reading frame that regulate D1 splice site usage (35,36). Future studies will elucidate the mechanism of U1-Gag2 and determine whether other U1i RNAs can inhibit HIV-1 production by targeting sites upstream from acceptor site 1.

Our main goal is to find the most potent molecules in each antiviral RNA class to be used in an anti-HIV-1 gene therapy setting. Therefore, we compared new and previously characterized U1i RNAs for their ability to inhibit HIV-1 production when transfected at various doses. We found that U1i RNAs that induce excessive splicing are more potent than those that inhibit HIV-1 RNA polyadenylation (Figure 3B). The inhibition by both U1i RNAs that induced excessive splicing and those that inhibited polyadenylation was dependant on the SL1 domain (Figure 5). This is in concordance with previous studies that demonstrated that the action of the endogenous U1 snRNP depends on proper association of the U1-70K with the SL1 domain of the U1

molecule (10,37). Surprisingly, mutations in the SL2 domain that inhibit recruitment of U1-A abolished the inhibition of viral replication by U1-Rev and U1-Nef. Previously, Sajic et al. showed that mutations in the SL2 domain of U1i RNA molecules targeting the 3'-terminal exon resulted in a reduction in antiviral effects but did not completely abolish their effects. In contrast, our results indicate otherwise when looking at viral replication using an RT assay and are coherent with a recent publication indicating that U1-A binding is dispensable for the improvement of exon and intron definition by therapeutic U1 snRNAs designed to correct exon skipping (37). One possible explanation for this observed discrepancy is a blockage after the production of Gag protein that inhibits viral replication. Furthermore, their experiments were conducted with the HxBruR$^-$/RI$^-$ HIV-1 plasmid which expresses a HIV-1 provirus with deletions in the RT and integrase genes.

The most potent U1i RNAs altered the expression of viral RNA isoforms. The expression of U1-D2 resulted in an increase in IS HIV-1 RNA (Figure 4A) and inclusion of exon 2 (Figure 4C) while expression of U1-T6 resulted in an increase of both IS RNA and CS viral RNA and an increase in inclusion of exon 4. In addition, U1-Env, which was previously characterized to inhibit HIV-1 polyadenylation induced excessive splicing in our study as demonstrated by the increase in CS RNA (Figure 4A) and in mRNA species containing exon 4 (Figure 4C). During HIV-1 replication, exclusion of the intron between 5' splice donor D4 and 3' splice acceptor A7 is required to generate CS mRNA coding for two-exon Tat, Rev and Nef. This process is dependent on the complex interplay between exon splice silencer (ESS) and exon splice enhancer elements (ESE) located downstream of the A7 splice acceptor site. Both, ESS and ESE modulate the binding of U2AF, a subunit of the U2 snRNP, to the PPyT which stabilizes the U2 snRNP binding to the 3'ss (38). Furthermore, U1 snRNP binding proximal to a 3'ss has been demonstrated to increase the recruitment of U2AF to the PPyT (39). Therefore, we

hypothesize that U1-Env increases recognition of splice acceptor site A7 by increasing recruitment of U2AF. This is further exemplified by U1-Rev which binds in proximity to U1-Env and is seen to increase inclusion of exon 4 in IS mRNA (Figure 4C). However, U1-Rev induces a general reduction in viral RNA species on Northern blot (Figure 4A), suggesting that the inhibition seen is mainly mediated through the inhibition of polyadenylation.

An important consideration when designing molecules to be used in HIV-1 gene therapy is that they target highly conserved sites and can cover the majority of circulating HIV-1 strains. While the majority of U1i RNAs targeted conserved sites in the HIV-1 genome, U1-Gag2, D2, T6 and Env were found to be the most potent and to target relatively highly conserved sites (Figure 3C). When compared to WT transfected cells, none of the U1i RNAs induced cellular toxicity indicating that the inhibition seen was not due to cell death. Although *in vivo* experiments in mice have demonstrated that long term expression of U1i RNAs does not result in any detectable toxicity, with a recognition domain of 10-11 nucleotides the specificity of these molecules remains a significant concern (13,40). U1i RNAs that inhibit polyadenylation were previously shown to distinguish between target sequences that differ by one or two nucleotides (24). Furthermore, a single mismatch in positions 3 to 8 of the U1i RNA recognition domain that differ from GU completely abrogated its silencing activity (25). In contrast, expression of U1i RNAs, in one study, resulted in non-specific inhibition of a Renilla expressing plasmid emphasizing the need to better explore the specificity of U1i RNAs with therapeutic potential (16).

With the aim of increasing the specificity of U1i RNA molecules, we increased the length of the recognition domains. Although several studies have demonstrated that maximal inhibition of targeted RNA is achieved with a 10 to 11 nucleotide long target site, molecules that inhibit

polyadenylation can have their recognition domain increased by 6 nucleotides towards the 3' end of the molecule with little effect on their inhibitory activity (24). In other studies, increasing the recognition domain of U1i RNAs by 1 to 5 nucleotides towards the 5'end of the molecule completely abolished their activity, whereas molecules whose recognition domains were increased by 6 nucleotides maintained their inhibitory activity and expression levels (25,41). With the exception of U1-Gag2 and U1-Env, we found that the recognition domain of all U1i RNAs evaluated could be increased without affecting their activity and for U1-Nef, one permutation even increased its activity (Figure 6D). Taken together, these results suggest that the specificity of U1i RNAs could be increased without affecting their potency and could result in a more targeted inhibition.

Based on the efficacy of the molecules and the conservation of their respective target site, we decided to use U1-T6 and U1-Nef to stably transduce SupT1 cells and determine their ability to restrict HIV-1 infection/replication. SupT1 cells were transduced with lentiviral vectors expressing the various U1i RNA molecules. Cells were challenged with NL4-3. Expression of U1-T6 resulted in a greater than 90% inhibition of viral production. Furthermore, when compared to a shRNA targeting HIV Tat/Rev, U1-T6 demonstrated superior inhibition of HIV-1 replication highlighting its higher potency and clinical potential.

In conclusion, we have demonstrated that U1i RNAs that induce excessive splicing are potent inhibitors of HIV-1 replication. Their distinct mechanism of action and ability to act synergistically with shRNAs make them very attractive candidates to be used in a gene therapy approach for HIV-1(14). While specificity remains an ongoing concern for these molecules, our results suggest that it is possible to increase the recognition domain of some very active U1i RNAs to 16 nucleotides and maintain their efficacy. Future studies will focus on determining the

long-term safety and toxicity of these molecules as well as their synergistic or antagonistic potential when combined with other antiviral RNA technologies such as ribozymes, decoy RNAs, aptamers and shRNAs.

TABLES:

Table 1. Summary of U1i RNA target sites in HIV-1

U1i RNA Name	Primer 5'-3'	Target Sequence 5'-3'	Location	Reading Frame
U1-Gag1	GGCCCAAGATCTCA**TAGTTCCTGC**GCAGG	GCAGGAACTA	1498	Gag
U1-Gag2	GGCCCAAGATCTCA**GTAGTTCCTG**GCAGG	CAGGAACTAC	1499	Gag
U1-Gag3	GGCCCAAGATCTCA**TACTAGTAGT** GCAGG	ACTACTAGTA	1504	Gag
U1-Gag4	GGCCCAAGATCTCA**GGGTACTAGT**GCAGG	ACTAGTACCC	1507	Gag
U1-D2	GGCCCAAGATCTCA**ACCTTCACCTG**GCAGG	AGGTGAAGG	4961	Gag-Pol (D2)
U1-D3	GGCCCAAGATCTCA**ATTCCTACCTG**GCAGG	AGGTAGGA	5262	Vif (D3)
U1-T3	GGCCCAAGATCTCA**TGGATGCTTCC**GCAGG	GGAAGCATCCA	5861	Tat/Vpr
U1-T4	GGCCCAAGATCTCA**TTTAGGCTGAC**GCAGG	GTCAGCCTAAA	5876	Tat
U1-T6	GGCCCAAGATCTCA**TTACAAGCAG**TGCAGG	CTGCTTGTA	5887	Tat
U1-Rev	GGCCCAAGATCTCA**GTAGCTGAAG**GCAGG	CTTCAGCTAC	8509	Rev exon 2
U1-Env	GGCCCAAGATCTCA**TTATAGCAAA**GCAGG	TTTGCTATAA	8776	Env
U1-Nef	GGCCCAAGATCTCA**ACTCCGGATG**GCAGG	CATCCGGAGT	9381	LTR/Nef

Position of targeted sequence is relative to HIV-1 provirus NL4-3

SUPPLEMENTAL MATERIAL:

Table S1. Primers used for elongating recognition domain		
Name	Target DNA sequence 5'-3'	Primer sequence 5'-3'
U1-Gag-2 5'-3	*TAG*CAGGAACTAC	GGCCCAAGATCTCA**GTAGTTCCTGCTA**GCAGG*GGAGA*
U1-Gag-2 5'-6	*ACATAG*CAGGAACTAC	GGCCCAAGATCTCA**GTAGTTCCTGCTATGT**GCAGG*GGAGA*
U1-Gag-2 3'-3	CAGGAACTAC*TAG*	GGCCCAAGATCTCA**CTAGTAGTTCCTG**GCAGG*GGAGA*
U1-Gag-2 3'-6	CAGGAACTAC*TAGTAC*	GGCCCAAGATCTCA**GTACTAGTAGTTCCTG**GCAGG*GGAGA*
U1-D2-5'-3	GAAAGGTGAAGGG	GGCCCAAGATCTCA**CCCTTCACCTGTTC**GCAGG*GGAGA*
U1-D2-5'-6	CTGGAAAGGTGAAGGG	GGCCCAAGATCTCA**CCCTTCACCTGTTCCAG**GCAGG*GGAGA*
U1-D2-3'-3	AGGTGAAGGG*GCA*	GGCCCAAGATCTCA**TGCCCCTTCACCT**GCAGG*GGAGA*
U1-D2-3'-6	AGGTGAAGGG*GCAGTA*	GGCCCAAGATCTCA**TACTGCCCCTTCACCT**GCAGG*GGAGA*
U1-T6-5'-3	AAACTGCTTGTAA	GGCCCAAGATCTCA**TTACAAGCAGTTT**GCAGG*GGAGA*
U1-T6-5'-6	CTAAAACTGCTTGTAA	GGCCCAAGATCTCA**TTACAAGCAGTTTTAG**GCAGG*GGAGA*
U1-T6-3'-3	CTGCTTGTAACCA	GGCCCAAGATCTCA**TGGTTACAAGCAG**GCAGG*GGAGA*
U1-T6-3'-6	CTGCTTGTAACCAATT	GGCCCAAGATCTCA**AATTGGTTACAAGCAG**GCAGG*GGAGA*
U1-Rev-5'-3	*CCT*CTTCAGCTAC	GGCCCAAGATCTCA**GTAGCTGAAGAGG**GCAGG*GGAGA*
U1-Rev-5'-6	*GTGCCT*CTTCAGCTAC	GGCCCAAGATCTCA**GTAGCTGAAGAGGCAC**GCAGG*GGAGA*
U1-Rev-3'-3	CTTCAGCTAC*CAC*	GGCCCAAGATCTCA**GTGGTAGCTGAAG**GCAGG*GGAGA*
U1-Rev-3'-6	CTTCAGCTAC*CACCGC*	GGCCCAAGATCTCA**GCGGTGGTAGCTGAAG**GCAGG*GGAGA*
U1-Env-5'-3	GATTTTGCTATAA	GGCCCAAGATCTCA**TTATAGCAAAATC**GCAGG*GGAGA*
U1-Env-5'-6	AAGGATTTTGCTATAA	GGCCCAAGATCTCA**TTATAGCAAAATCCTT**GCAGG*GGAGA*
U1-Env-3'-3	TTTGCTATAAGAT	GGCCCAAGATCTCA**ATCTTATAGCAAA**GCAGG*GGAGA*
U1-Env-3'-6	TTTGCTATAAGATGGG	GGCCCAAGATCTCA**CCCATCTTATAGCAAA**GCAGG*GGAGA*
U1-Nef-5'-3	CTGCATCCGGAGT	GGCCCAAGATCTCA**ACTCCGGATGCAG**GCAGG*GGAGA*
U1-Nef-5'-6	GAGCTGCATCCGGAGT	GGCCCAAGATCTCA**ACTCCGGATGCAGCTC**GCAGG*GGAGA*
U1-Nef-3'-3	CATCCGGAGTACT	GGCCCAAGATCTCA**AGTACTCCGGATG**GCAGG*GGAGA*
U1-Nef-3'-6	CATCCGGAGTACTACA	GGCCCAAGATCTCA**TGTAGTACTCCGGATG**GCAGG*GGAGA*

Table S2. Primers used for site directed mutagenesis

SL1Mut	Forward Primer	Reverse Primer
U1-Gag2	GCTTCCTGGTTTTCCCAGGGCGAG	ACTAGTTGGTATCTCCCCTGCCAG
U1-D2	GCTTCCTGGTTTTCCCAGGGCGAG	ACTAGTTGGTATCTCCCCTGCCAG
U1-T6	GCTTCCTGGTTTTCCCAGGGCGAG	ACTAGTTGGTATCTCCCCTGCACTG
U1-Rev	GCTTCCTGGTTTTCCCAGGGCGAG	ACTAGTTGGTATCTCCCCTGCCTTC
U1-Nef	GCTTCCTGGTTTTCCCAGGGCGAG	ACTAGTTGGTATCTCCCCTGCCATC
SL2Mut	CATTGCACTCACGCGGATGTGCTGACC	GATAAGCCTCGCCCTGGG

Primers were designed using the NEBaseChanger V1.2.6
*U1-Env was cloned using initial cloning primers on SL1Mut U1-WT plasmid

REFERENCES:

1. Jones, J., Taylor, B.S., Tieu, H.V. and Wilkin, T.J. (2017) CROI 2017: Advances in Antiretroviral Therapy. *Top Antivir Med*, **25**, 51-67.

2. Carr, A. (2003) Toxicity of antiretroviral therapy and implications for drug development. *Nat Rev Drug Discov*, **2**, 624-634.

3. Chun, T.W., Justement, J.S., Murray, D., Hallahan, C.W., Maenza, J., Collier, A.C., Sheth, P.M., Kaul, R., Ostrowski, M., Moir, S. *et al.* (2010) Rebound of plasma viremia following cessation of antiretroviral therapy despite profoundly low levels of HIV reservoir: implications for eradication. *AIDS*, **24**, 2803-2808.

4. Burnett, J.C. and Rossi, J.J. (2012) RNA-based therapeutics: current progress and future prospects. *Chem Biol*, **19**, 60-71.

5. Das, A.T., Brummelkamp, T.R., Westerhout, E.M., Vink, M., Madiredjo, M., Bernards, R. and Berkhout, B. (2004) Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition. *J Virol*, **78**, 2601-2605.

6. Will, C.L. and Luhrmann, R. (2011) Spliceosome structure and function. *Cold Spring Harb Perspect Biol*, **3**.

7. Berget, S.M. (1995) Exon recognition in vertebrate splicing. *J Biol Chem*, **270**, 2411-2414.

8. Robberson, B.L., Cote, G.J. and Berget, S.M. (1990) Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol*, **10**, 84-94.

9. Vagner, S., Ruegsegger, U., Gunderson, S.I., Keller, W. and Mattaj, I.W. (2000) Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA*, **6**, 178-188.

10. Gunderson, S.I., Polycarpou-Schwarz, M. and Mattaj, I.W. (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell*, **1**, 255-264.

11. Carpousis, A.J., Vanzo, N.F. and Raynal, L.C. (1999) mRNA degradation. A tale of poly(A) and multiprotein machines. *Trends Genet*, **15**, 24-28.

12. Blazquez, L. and Fortes, P. (2015) U1 interference (U1i) for antiviral approaches. *Adv Exp Med Biol*, **848**, 51-69.

13. Blazquez, L., Gonzalez-Rojas, S.J., Abad, A., Razquin, N., Abad, X. and Fortes, P. (2012) Increased in vivo inhibition of gene expression by combining RNA interference and U1 inhibition. *Nucleic Acids Res*, **40**, e8.

14. Abad, X., Razquin, N., Abad, A. and Fortes, P. (2010) Combination of RNA interference and U1 inhibition leads to increased inhibition of gene expression. *Nucleic Acids Res*, **38**, e136.

15. Sajic, R., Lee, K., Asai, K., Sakac, D., Branch, D.R., Upton, C. and Cochrane, A. (2007) Use of modified U1 snRNAs to inhibit HIV-1 replication. *Nucleic Acids Res*, **35**, 247-255.

16. Knoepfel, S.A., Abad, A., Abad, X., Fortes, P. and Berkhout, B. (2012) Design of modified U1i molecules against HIV-1 RNA. *Antiviral Res*, **94**, 208-216.

17. Mandal, D., Feng, Z. and Stoltzfus, C.M. (2010) Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J Virol*, **84**, 12790-12800.

18. Scarborough, R.J., Adams, K.L., Del Corpo, O., Daher, A. and Gatignol, A. (2016) Evaluation of the Efficacy And Toxicity of RNAs Targeting HIV-1 Production for Use in Gene or Drug Therapy. *J Vis Exp*.

19. Scarborough, R.J., Levesque, M.V., Perreault, J.P. and Gatignol, A. (2014) Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol*, **1103**, 31-43.

20. Scarborough, R.J., Levesque, M.V., Boudrias-Dalle, E., Chute, I.C., Daniels, S.M., Ouellette, R.J., Perreault, J.P. and Gatignol, A. (2014) 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.

21. Yam, P.Y., Li, S., Wu, J., Hu, J., Zaia, J.A. and Yee, J.K. (2002) Design of HIV vectors for efficient gene delivery into human hematopoietic cells. *Mol Ther*, **5**, 479-484.

22. Purcell, D.F. and Martin, M.A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol*, **67**, 6365-6378.

23. Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K., Li, J. and Nagai, K. (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature*, **458**, 475-480.

24. Liu, P., Gucwa, A., Stover, M.L., Buck, E., Lichtler, A. and Rowe, D. (2002) Analysis of inhibitory action of modified U1 snRNAs on target gene expression: discrimination of two RNA targets differing by a 1 bp mismatch. *Nucleic Acids Res*, **30**, 2329-2339.

25. Abad, X., Vera, M., Jung, S.P., Oswald, E., Romero, I., Amin, V., Fortes, P. and Gunderson, S.I. (2008) Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Res*, **36**, 2338-2352.

26. DiGiusto, D.L., Krishnan, A., Li, L., Li, H., Li, S., Rao, A., Mi, S., Yam, P., Stinson, S., Kalos, M. *et al.* (2010) RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med*, **2**, 36ra43.

27. An, D.S., Donahue, R.E., Kamata, M., Poon, B., Metzger, M., Mao, S.H., Bonifacino, A., Krouse, A.E., Darlix, J.L., Baltimore, D. *et al.* (2007) 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.

28. Wolstein, O., Boyd, M., Millington, M., Impey, H., Boyer, J., Howe, A., Delebecque, F., Cornetta, K., Rothe, M., Baum, C. *et al.* (2014) 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.

29. ter Brake, O., Legrand, N., von Eije, K.J., Centlivre, M., Spits, H., Weijer, K., Blom, B. and Berkhout, B. (2009) Evaluation of safety and efficacy of RNAi against HIV-1 in the human immune system (Rag-2(-/-)gammac(-/-)) mouse model. *Gene Ther*, **16**, 148-153.

30. Tebas, P., Stein, D., Binder-Scholl, G., Mukherjee, R., Brady, T., Rebello, T., Humeau, L., Kalos, M., Papasavvas, E., Montaner, L.J. *et al.* (2013) Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV. *Blood*, **121**, 1524-1533.

31. Chung, J., Scherer, L.J., Gu, A., Gardner, A.M., Torres-Coronado, M., Epps, E.W., Digiusto, D.L. and Rossi, J.J. (2014) 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.

32. van Gestel, M.A., van Erp, S., Sanders, L.E., Brans, M.A., Luijendijk, M.C., Merkestein, M., Pasterkamp, R.J. and Adan, R.A. (2014) shRNA-induced saturation of the microRNA pathway in the rat brain. *Gene Ther*, **21**, 205-211.

33. Amado, R.G., Mitsuyasu, R.T., Rosenblatt, J.D., Ngok, F.K., Bakker, A., Cole, S., Chorn, N., Lin, L.S., Bristol, G., Boyd, M.P. *et al.* (2004) Anti-human immunodeficiency virus hematopoietic progenitor cell-delivered ribozyme in a phase I study: myeloid and lymphoid reconstitution in human immunodeficiency virus type-1-infected patients. *Hum Gene Ther*, **15**, 251-262.

34. Scarborough, R.J., Adams, K.L., Daher, A. and Gatignol, A. (2015) 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.

35. Asang, C., Erkelenz, S. and Schaal, H. (2012) The HIV-1 major splice donor D1 is activated by splicing enhancer elements within the leader region and the p17-inhibitory sequence. *Virology*, **432**, 133-145.

36. Schaub, M.C., Lopez, S.R. and Caputi, M. (2007) Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. *J Biol Chem*, **282**, 13617-13626.

37. Rogalska, M.E., Tajnik, M., Licastro, D., Bussani, E., Camparini, L., Mattioli, C. and Pagani, F. (2016) Therapeutic activity of modified U1 core spliceosomal particles. *Nat Commun*, **7**, 11168.

38. Staffa, A. and Cochrane, A. (1995) Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol Cell Biol*, **15**, 4597-4605.

39. Hoffman, B.E. and Grabowski, P.J. (1992) U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev*, **6**, 2554-2568.

40. Koornneef, A., van Logtenstein, R., Timmermans, E., Pisas, L., Blits, B., Abad, X., Fortes, P., Petry, H., Konstantinova, P. and Ritsema, T. (2011) AAV-mediated in vivo knockdown of luciferase using combinatorial RNAi and U1i. *Gene Ther*, **18**, 929-935.

41. Fortes, P., Cuevas, Y., Guan, F., Liu, P., Pentlicky, S., Jung, S.P., Martinez-Chantar, M.L., Prieto, J., Rowe, D. and Gunderson, S.I. (2003) Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc Natl Acad Sci U S A*, **100**, 8264-8269.

Chapter 3: Additional Experiments

The experiments presented in this chapter were performed to complement the manuscript presented in Chapter 2, but were not included in the version that will be submitted. I performed the experiments shown in figure 8. The objectives were to analyze the activity of the U1i RNAs at the protein level.

Methods:

HIV-1 intracellular protein expression and immune activation

HEK293T cells were co-transfected with 400 ng of HIV pNL4-3 and 1000 ng of different U1i RNA plasmids in 6 well plates. 48 hrs after co-transfection, cell lysates were harvested in 50 µl of lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM EDTA, 10 % v/v glycerol, 1 % v/v NP-40) containing antiproteases and antiphosphatases (Roche). 100 µg of total protein was resolved on a 10% denaturing polyacrylamide gel and transferred to a Hybond ECL nitrocellulose membrane (GE Healthcare) as previously described (1). Membranes were first incubated with HIV-1 anti-p24 capsid (183-H12-5c) at a 1/1000 dilution. Membranes were stripped by incubation with Re-Blot Plus Strong (Millipore) for 10 mins and then incubated with anti-GAPDH (sc-322333, Santa Cruz) at a dilution of 1/500. Following primary antibody incubation, membranes were incubated with anti-mouse IgG-horseradish peroxidase (GE healthcare) at a 1/5000 dilution for 1 hr. Bands were visualized with chemiluminescence reagent plus (ECL) (GE Healthcare) on films. The relative amount of Gag (p55/GAG) precursor and capsid (p24/CA) products were determined by quantification of the chemiluminescence of each protein band using Image J densitometry software (Version 10.2). All ratios were expressed relative to the Gag and capsid bands of cells co-transfected with U1-WT and pNL4-3.

Results:

U1i RNAs that induce excessive splicing alter HIV-1 Gag processing

It has been previously shown that mutations in HIV-1 second 5'ss, D2, which increase its homology to the endogenous U1 snRNA results in Gag protein processing defects (4). Therefore, we next asked whether U1i RNAs, which inhibit polyadenylation or induce excessive splicing, also affected Gag processing. To examine this question, total cellular protein extracts of HEK293T cells co-transfected with pNL4-3 and different U1i RNAs were analyzed for expression of the Gag polyprotein (p55), the matrix-capsid intermediate (p41) and fully processed CA (p24) protein of HIV-1 using an antibody directed against the capsid protein (Figure 8). To evaluate effects on processing of the Gag polyprotein, the ratios of processed Gag (p24) and Gag polyprotein (p55) to WT transfected cells was determined using Image J densitometry software (Figure 8C). U1i RNAs designed to target excessive splicing (U1-D2 and U1-T6) strongly reduced the ratio of processed Gag (p24), suggesting that they can inhibit Gag processing as previously reported (5). In contrast, U1-Gag2 partially reduced Gag processing and U1i RNAs designed to inhibit polyadenylation did not significantly affect Gag processing (Figure 8C). Preliminary experiments demonstrate that U1-Env also has the ability to alter Gag processing (Figure 8D). These results suggest that the more effective inhibition of HIV-1 production observed for U1i RNAs targeting HIV-1 splicing compared to those that inhibit polyadenylation, may be due to an additional effect of the U1i RNAs on Gag processing.

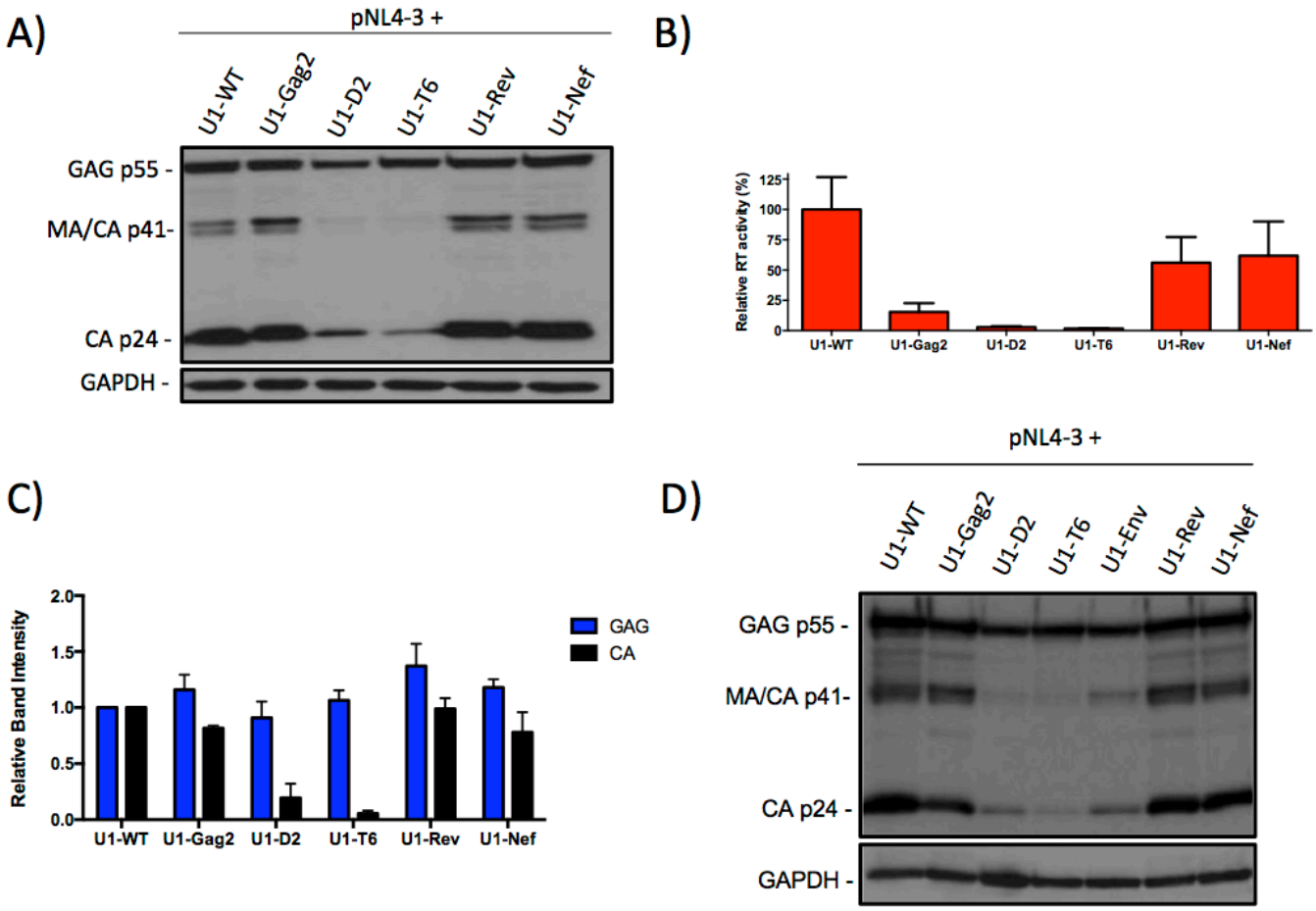

Figure 8. U1i RNAs that induce excessive splicing alter Gag processing. (A) Analysis of Gag processing in protein isolated from HEK293T cells transfected with 400 ng of pNL4-3 alone or in combination with 1000 ng of U1-WT, U1i RNAs that induce excessive splicing (U1-D2, T6) or U1i RNAs that inhibit polyadenylation (U1-Rev, Nef). 48 hrs post transfection cell were harvested and the processing of p24 was assessed by Western blot. Gag precursor (p55), intermediate (p41) and capsid (p24) are indicated. (B) Relative RT activity in cultured supernatant of transfected HEK293T cells. Results are expressed as a relative RT percentage of HIV-1 inhibition to the U1-WT co-transfected cells. (C) p55 and p24 band intensities were calculated relative to U1-WT co-transfected cells using ImageJ software. All data are represented as the mean +/- SEM from three independent experiments (n = 3). (D) Preliminary analysis of Gag processing by U1-Env (n = 1).

Chapter 4:Reflection and Future Directions

Specificity and Toxicity of U1i RNAs

A big concern for the use of these U1i RNA molecules in gene therapy to inhibit HIV-1 replication is their specificity. With a 10 nucleotide recognition domain, binding sites are expected to occur by chance every 10^6 base pairs representing roughly 3000 occurrences in the human genome (8). Since exons represent 2% of the human genome and U1i RNAs are believed to only inhibit mRNAs when targeting regions within 3'-terminal exons, it was thought that rationale design of specific U1i RNAs was possible. However, it has been shown that U1i RNAs targeting not only 5'ss but also downstream of 3'ss have the ability to induce splicing by the process of exon definition (5). Moreover, we demonstrate that U1i RNAs targeting a site that is upstream of a 3'ss has the ability to inhibit HIV-1 replication. Therefore, it is necessary to perform experiments to determine the overall specificity of these molecules especially in the case of U1i RNAs that target 5'ss.

There are around 300 000 5'ss found in the human genome. The majority of these have modest complementarity with the U1 snRNA with a mean of 6 out of 10 nucleotides. Binding of U1i RNAs to sequences with little complementarity is regulated by various ESE and ISE (9). Since little homology is required, designing U1i RNAs targeting 5'ss may result in significant off-target effects. Looking at the recognition domain of U1-D2 in our study (5'-CCTTCACCCTG-3') and comparing it to that of the endogenous U1 snRNA (5'-TACTTACCTG-3'), there are 7 nucleotides in common (underlined) indicating that there is a high potential for off-target effects. The same is true for U1-D3 which also has 7 nucleotides in common with WT-U1. In order to avoid potential off-target effects, U1i RNAs designed to enhance splicing should target regions upstream of 3'ss as they likely have lower homology to

the endogenous U1 snRNA. This is exemplified by U1-T3 and T4 which only have 1 nucleotide in common with the endogenous U1 snRNA and U1-T6 which has 3 nucleotides in common. To further avoid off-target effects when designing U1i RNAs, sequences should be checked using BLAST for potential binding sites in human genes and should not contain nucleotides A, G and T at positions 2, 8 and 9, respectively, to avoid targeting cryptic 5'ss (10). Future experiments will consist of screening regions downstream of the 3'ss A1, A2, A3, A4, A5 and A7 in the HIV-1 genome to identify conserved sites that can be safely targeted by U1i RNAs. Newly constructed U1i RNAs will be tested for their anti-HIV-1 capacity to identify the most potent compounds. Once we have identified several candidate molecules, we will then attempt to increase their specificity and perform toxicity studies.

To potentially increase the specificity of these molecules, we have elongated the recognition domain by 6 nucleotides towards the 5' or 3' end of the molecule. Surprisingly, molecules with 16 nucleotide-long recognition domains had the same antiviral efficacy as unmodified molecules. However, whether or not this increases the specificity of these molecules has yet to be determined. Abad et al. found that mutating the U1-binding site positions 3 to 8 of the targeted transcripts made them resistant to U1i RNA molecules, revealing the critical importance of these positions for inhibition (11). However, single point mutations in positions 1, 2, 9 and 10 were tolerated. Therefore, for molecules with elongated recognition domains, there is a potential that the added nucleotides do not contribute to the overall specificity of the molecules and that U1i specificity depends on the 8-nucleotide central sequence.

Future studies will consist of introducing single and subsequently multiple point mutations in the recognition domain of the elongated U1i RNA molecule and determining their ability to inhibit HIV-1 replication using the same methods as our study. Furthermore, we will

construct elongated U1i RNA molecules with mutations in the SL1 U1-70K recognition domain to determine if the inhibition seen when expressing these modified U1i RNAs is mediated by antisense mechanisms or is dependent upon proper assembly of these molecules. It would also be interesting to create U1i RNAs that have elongated recognition domains towards the 3' and 5' end of the molecule. After identifying potent and specific anti-HIV-1 U1i RNAs, preliminary toxicity studies will consist of testing their ability to induce cellular toxicity by the WST-1 assay and to activate innate immune responses. Specifically, the expression of interferon-stimulated genes and activation of dsRNA immune sensors will be assessed by Western blotting. While costly, microarray studies should be performed to determine if expression of these modified molecules alter the levels of human mRNA expression through off-target effects.

U1i RNAs and Gag Processing

The relationship between U1i RNAs that induce excessive splicing of HIV-1 RNA and Gag processing is not entirely clear. One possible explanation for the defect in Gag protein processing in cells transfected with U1-D2 is that the overproduction of Vif protein inhibits proper proteolytic processing of Gag (12). However, this is likely not the case as U1-T6 and U1-Env which do not increase the production of Vif mRNA also have an impact on Gag processing. There is a possibility that particle assembly is affected by the reduced amount of FL mRNA seen when cells are transfected with U1-D2, U1-T6 and U1-Env. This seems unlikely as U1-Gag2, U1-Rev and U1-Nef also reduce the amount of FL HIV-1 RNA and have no effect on Gag processing. Therefore, we think that the defect in Gag processing is due to a failure of Gag to assemble and associate with the plasma membrane. This mechanism is likely specific to U1i RNAs that induce excessive splicing. We have begun to do immunofluorescent studies staining for U1i RNA in transfected cells using antibodies against p24 and co-staining for cellular membrane protein in order to demonstrate lack of Gag recruitment, but these studies were

inconclusive. Therefore, future experiments will consist of using a cell fractionation kit to isolate the cellular membrane of HIV-1 infected cells transfected with U1i RNAs and perform Western blots to determine the recruitment of p24.

Synergistic Potential of U1i RNAs

The increased inhibition seen when combining two antiviral RNA molecules is very desirable in antiviral gene therapy as synergism allows for potent viral inhibition with decreased doses of the antiviral molecules. This is therapeutically relevant as shRNAs have been characterized to have dose-dependent off-target effects (13). Moreover, high expression of antiviral RNAs from strong RNA pol III U6 promoters has resulted in cellular toxicity in combinatorial anti-HIV gene therapy clinical trials ran by the City of Hope (14). Therapeutically, achieving high inhibition of viral RNA production is essential, as an inhibition of 97.5% of HIV-1 RNA production would leave twice as many viral RNA copies when compared to an inhibition of 98.75%. Because the replication of a few viral particles may be sufficient to sustain viral infection in individuals, achieving the strongest inhibition remains a priority (15). One way of achieving high viral RNA inhibition is through the expression of antiviral molecules that act synergistically. Previously, U1i RNAs have been characterized to have a synergistic antiviral effect when co-expressed with another U1i RNA or shRNA targeting the same mRNA transcript. However, not all combinations of U1i RNAs and shRNAs result in synergistic inhibition (16). Therefore, future studies will consist of co-expressing potent anti-HIV U1i RNAs and shRNAs to identify synergistic combinations. Seemingly, these combinations will have to be identified empirically as the mechanism of synergistic inhibition obtained when combining shRNAs and U1i RNAs has yet to be determined. Since U1i RNAs inhibit viral RNA production at the level of the nucleus, it would be interesting to see if there is synergism when co-expressed with other

antiviral RNA molecules that target cytoplasmic viral RNA such as ribozymes and aptamer RNAs.

Concluding Remarks

In summary, there is an emergent international effort aimed at finding a cure for HIV infection. A huge concern in almost all cure strategies is the risk to benefit ratio. Delivery of cART may not be accessible for many individuals. However, those who do have access and adhere to the drug regimen generally have a good prognosis (17). Nonetheless, life-long cART has been estimated to cost around 420,000 to 755,000$ while the cost of an autologous bone marrow transplant with additional care and treatment has been estimated to costs around 250,000 to 300,000$ per patient. Furthermore, HSC based gene therapy has been shown to be safe and feasible (18,19). Although there are many hurdles that must be addressed, the identification of synergistic combinations of antiviral genes is essential for the development of HSC based gene therapy to cure HIV. Here we demonstrate that U1i RNA could make a great addition to current combinatorial anti-HIV-1 gene therapy and warrant future studies to characterize their specificity and potential synergism with other antiviral RNA molecules.

References:

1. Scarborough, R.J., Levesque, M.V., Boudrias-Dalle, E., Chute, I.C., Daniels, S.M., Ouellette, R.J., Perreault, J.P. and Gatignol, A. (2014) 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.

2. Scarborough, R.J., Levesque, M.V., Perreault, J.P. and Gatignol, A. (2014) Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol*, **1103**, 31-43.

3. Michienzi, A., Li, S., Zaia, J.A. and Rossi, J.J. (2002) A nucleolar TAR decoy inhibitor of HIV-1 replication. *Proc Natl Acad Sci U S A*, **99**, 14047-14052.

4. Mandal, D., Feng, Z. and Stoltzfus, C.M. (2008) Gag-processing defect of human immunodeficiency virus type 1 integrase E246 and G247 mutants is caused by activation of an overlapping 5' splice site. *J Virol*, **82**, 1600-1604.

5. Mandal, D., Feng, Z. and Stoltzfus, C.M. (2010) Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *J Virol*, **84**, 12790-12800.

6. Foucquier, J. and Guedj, M. (2015) Analysis of drug combinations: current methodological landscape. *Pharmacol Res Perspect*, **3**, e00149.

7. DiGiusto, D.L., Krishnan, A., Li, L., Li, H., Li, S., Rao, A., Mi, S., Yam, P., Stinson, S., Kalos, M. *et al.* (2010) RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med*, **2**, 36ra43.

8. Makalowski, W. (2001) The human genome structure and organization. *Acta Biochim Pol*, **48**, 587-598.

9. Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*, **72**, 291-336.

10. Blazquez, L. and Fortes, P. (2013) U1 snRNP control of 3'-end processing and the therapeutic application of U1 inhibition combined with RNA interference. *Curr Mol Med*, **13**, 1203-1216.

11. Abad, X., Vera, M., Jung, S.P., Oswald, E., Romero, I., Amin, V., Fortes, P. and Gunderson, S.I. (2008) Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Res*, **36**, 2338-2352.

12. Akari, H., Fujita, M., Kao, S., Khan, M.A., Shehu-Xhilaga, M., Adachi, A. and Strebel, K. (2004) High level expression of human immunodeficiency virus type-1 Vif inhibits viral infectivity by modulating proteolytic processing of the Gag precursor at the p2/nucleocapsid processing site. *J Biol Chem*, **279**, 12355-12362.

13. Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F. and Kay, M.A. (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*, **441**, 537-541.

14. Chung, J., Scherer, L.J., Gu, A., Gardner, A.M., Torres-Coronado, M., Epps, E.W., Digiusto, D.L. and Rossi, J.J. (2014) 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.

15. Blazquez, L., Gonzalez-Rojas, S.J., Abad, A., Razquin, N., Abad, X. and Fortes, P. (2012) Increased in vivo inhibition of gene expression by combining RNA interference and U1 inhibition. *Nucleic Acids Res*, **40**, e8.

16. Abad, X., Razquin, N., Abad, A. and Fortes, P. (2010) Combination of RNA interference and U1 inhibition leads to increased inhibition of gene expression. *Nucleic Acids Res*, **38**, e136.

17. Kiem, H.P., Jerome, K.R., Deeks, S.G. and McCune, J.M. (2012) Hematopoietic-stem-cell-based gene therapy for HIV disease. *Cell Stem Cell*, **10**, 137-147.

18. Schackman, B.R., Gebo, K.A., Walensky, R.P., Losina, E., Muccio, T., Sax, P.E., Weinstein, M.C., Seage, G.R., 3rd, Moore, R.D. and Freedberg, K.A. (2006) The lifetime cost of current human immunodeficiency virus care in the United States. *Medical care*, **44**, 990-997.

19. DiGiusto, D.L., Stan, R., Krishnan, A., Li, H., Rossi, J.J. and Zaia, J.A. (2013) Development of hematopoietic stem cell based gene therapy for HIV-1 infection: considerations for proof of concept studies and translation to standard medical practice. *Viruses*, **5**, 2898-2919.