

HIV, ADAR Editing, The Rev-RRE Pathway and Antisense RNA-mediated Inhibition

Siripong Tongjai
Lampang, Thailand

Bachelor of Sciences in Biology,
University of Virginia, 2009

A Dissertation presented to the Graduate Faculty of the University of Virginia in
Candidacy for the Degree of Doctor of Philosophy

Department of Biochemistry and Molecular Genetics

University of Virginia
May 2017

HIV, ADAR Editing, The Rev-RRE Pathway and Antisense RNA-mediated Inhibition

Siripong Tongjai
Lampang, Thailand

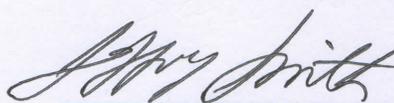
Bachelor of Sciences in Biology,
University of Virginia, 2009

A Dissertation presented to the Graduate Faculty of the University of Virginia in
Candidacy for the Degree of Doctor of Philosophy

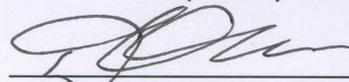
Department of Biochemistry and Molecular Genetics

University of Virginia

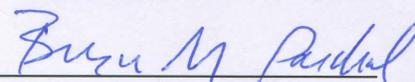
May 2017



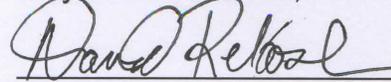
Jeffrey S. Smith, Ph.D. (Chairperson)



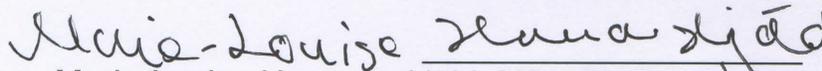
David T. Auble, Ph.D.



Bryce M. Paschal, Ph.D.



David Rekosh, Ph.D.



Marie-Louise Hammarskjöld, MD., Ph.D. (Mentor)

Abstract

In this study we demonstrate that a 287-bp long dsRNA panhandle structure can form in a 937-nucleotide HIV1 Antisense (AS) RNA that is expressed in a lentivirus vector which was previously shown to efficiently inhibit HIV expression. The presence of the panhandle showed a ten-fold better inhibition than an AS RNA that lacked the panhandle. Significant inhibition was also achieved with a vector that maintained only the 287-bp HIV panhandle, but substituted a heterologous EGFP sequence in the loop between the panhandle. Furthermore, expression of an AS RNA with an antisense/sense EGFP panhandle, rather than the one derived from HIV, led to efficient inhibition of both viral production and EGFP expression from the NL4-3-derived proviral clone pHIG, which contains an EGFP sequence inserted into the HIV genome. In all cases, efficient inhibition was only observed when the AS RNA trafficked on the Rev/RRE pathway. Thus our experiments demonstrate that the Rev/RRE pathway can be utilized to achieve efficient inhibition using long AS RNAs, especially when the generated AS RNA can form a double-stranded panhandle between sense and antisense sequences.

Sanger and Next Generation Sequence (NGS) analysis of the AS RNA expressed from the Lentivirus vector expressing the 937-nt AS env RNA with the panhandle showed that multiple A-to-G changes could be detected in the panhandle region in cDNA obtained from AS RNA in cells transfected with the vector alone. This was likely to be the result of editing of the dsRNA by Adenosine Deaminases Acting on RNA (ADAR) enzymes. Removal of the RRE

from the vector, expression without Rev or redirection of the AS RNA to the Nxf1 export pathway, resulted in a significant reduction in the observed A-to-G changes in the panhandle. This indicates that RNA that traffics on the Rev/RRE pathway intersects with the ADAR machinery in a way that is different from mRNA trafficking on other export pathways. This is further supported by the fact that over expression of huADAR1 and huADAR2s interferes with Rev/RRE-driven AS inhibition. In the presence of HIV target mRNA, A-to-G changes were also observed in the regions of AS and target mRNA outside of the panhandle, which would be expected to form intermolecular double-stranded hybrids in transfected cells. However, compared to the changes in the AS RNA, only very low levels of editing were observed in the target mRNA. Thus these findings speak against a previously proposed mechanism for AS inhibition that involves retention of antisense/sense hybrids in the nuclear matrix after A-to-I editing.

Processing of edited AS dsRNA into small RNAs that function as miRNAs/siRNAs represents a novel alternative mechanism that could explain the efficient inhibition. This is consistent with the fact that the partial depletion of the human Tudor Staphylococcal Nuclease (huTudor-SN or SND1), implied in the processing of edited dsRNAs, led to a less effective AS RNA-mediated inhibition. Furthermore, NGS analysis showed that AS RNA-derived miRNAs/siRNAs could be specifically detected in cells during Rev/RRE-mediated AS inhibition. Taken together, our results demonstrate a potential novel mechanism for AS RNA-mediated inhibition dependent on RNA trafficking on the Rev/RRE pathway.

Acknowledgements

I would like to thank my mentors Drs. Lou Hammarskjöld and David Rekosh for their guidance, support, and care during my training. Also, I would like to thank my committee members, Drs. Jeff Smith, David Auble, and Bryce Paschal for their support and advice over the years. Many thanks to all HamRek lab members past and present for their generosity and friendship. Particularly, I would like to thank “The Master”, Dr. Yeou-Cherng Bor, for his invaluable guidance. I would also like to thank Susan Prasad for her support over the years, and Dr. Chringma Sherpa and her family for their friendship. Many thanks to Eileen Trainum, Debbie Sites, and Sandy Weirich for their administrative assistance and support.

I would like to thank Drs. Eric Houpt and Carla Green for giving me the opportunity to work in their labs during my undergraduate years at UVA. I also would like to thank Dr. Misty Rae Gilbert, Dr. Mami Taniuchi, and Suzanne Stroup for their friendship and advice. Especially, I would like to thank Drs. Jirakan Nunkaew, Salinporn Kittiwatanakul, Nattawut Anuniwat, Kiadtisak Saenboonruang, Parinya Anantachaisilp, and Rujira Chaysiri, for being there when I needed them the most. Thanks to all of my friends back in Thailand and in the States since there are so many, I cannot mention them all here. Thanks to Will Coleman and his family for showing me their generosity and true Southern hospitality.

I would like to thank my beloved teachers, Mrs. Saowanee Anuphantanan, Ms. Banyen Jantararittikul, Ms. Choorat Aksornsakbun, Mrs. Kinnorn Thakum,

Mr. Opart Chaisongkhram, and Mrs. Nongyao Kanhasiri, for their invaluable guidance and great care generously given to me as a student at Lampang Kalayani School and as one of their own children. I would like to thank Mrs. Pamela McDonald, my beloved advisor at Westminster School, as well as, all of my teachers back in Thailand and at Westminster School, whom I do not mention here, for their generous supports. I have come this far because of all of you.

Finally, I would like to thank my own family for supporting me, encouraging me, and simply being there by my side although we are separated by great physical distance. Thanks Mom and Dad, for everything. You are the inspiration that gets me through all of the difficulties in my life. Both of you mean the world to me, and I hope I make you proud. Thanks Tew (or “Bai Tong”) and my lovely cousin, Sister Noi, for your kind words that have been cheering me up. Thanks to all of my relatives back home. I hope I can make all of you proud.

Dedication

In memory of His Majesty King Bhumibol Adulyadej.

In memory of Mrs. Khamnoi Chatkumpaeng and Mrs. Phong Tongjai, my beloved grandmothers.

Table of Contents

Abstract	1
Acknowledgements	3
Dedication	5
Table of Contents	6
List of Figures	10
List of Tables	14
Chapter 1 – Introduction	15
Human Immunodeficiency Virus (HIV)	15
<i>Viral entry</i>	17
<i>Uncoating</i>	17
<i>Reverse Transcription</i>	18
<i>Integration</i>	20
<i>HIV gene expression and viral mRNA export</i>	21
<i>Viral assembly and release</i>	23
<i>Viral maturation</i>	24
RNA Export Pathways	26
<i>The HIV-1 Rev/RRE Export Pathway</i>	26
<i>The MPMV CTE-mediated nuclear export pathway</i>	28
<i>An Overview of host cell mRNA Export</i>	30
Naturally occurring Antisense RNAs (AS RNAs)	34
<i>Naturally occurring retroviral antisense RNAs</i>	35
<i>The transcription of naturally occurring antisense HIV-1 RNA</i>	35
<i>HIV-1 AS RNA and its potential role in viral latency</i>	36

Table of Contents (continued)

Lentiviral vectors	38
<i>Three generations of Lentiviral vectors</i>	39
<i>Improvement of the transfer vectors</i>	41
<i>Pseudotyping of Lentiviral vectors</i>	42
<i>The use of Lentiviral vectors in immunotherapy</i>	42
Antisense RNA and gene therapy	43
RNA interference (RNAi)	50
ADAR editing	57
Project Rationale	64
Chapter 2 – A double-stranded RNA panhandle promotes antisense RNA-mediated inhibition of HIV viral production in a Rev/RRE dependent manner.	65
Results	
1) <i>An AS RNA that contains a 287 bp panhandle is able to inhibit HIV replication more efficiently than an AS RNA that lacks it.</i>	67
2) <i>An AS RNA containing a double-stranded RNA panhandle with only a 287-nt HIV specific AS sequence is sufficient for efficient inhibition.</i>	71
3) <i>An antisense vector that delivers an AS EGFP RNA with a panhandle is able to efficiently inhibit expression of a GFP containing Lentiviral vector, in an HIV Rev/RRE pathway-dependent manner.</i>	73
Discussion	
1) <i>A 287-bp panhandle enhances the AS RNA-mediated inhibition of HIV-1 replication.</i>	76
2) <i>The 287-bp panhandle derived from the HIV-1 env gene is sufficient for the efficient inhibition of HIV expression from pNL4-3.</i>	77
3) <i>A 287bp EGFP panhandle in an EGFP AS RNA promotes efficient inhibition of EGFP expression.</i>	78

Table of Contents (continued)

4) *Potential Mechanism of Inhibition*.....79

Chapter 3 – Analysis of ADAR editing during AS RNA-mediated inhibition

.....81

Results

1) *DNA sequencing analysis of the panhandle region of cDNA made from AS RNA shows a high level of A-to-G changes when the AS RNA traffics on the Rev/RRE pathway.*82

2) *Overexpression of huADAR1L or huADAR2s impairs RRE-driven AS inhibition.*89

3) *Analysis of potential ADAR editing in AS-RNA and target HIV mRNAs with and without ADAR1L or ADAR2s overexpression*92

4) *The 5' and 3' nearest-neighbor preferences pattern of edited As in the AS RNA compared to the reported patterns associated with RNA editing by a particular ADAR enzyme*.....99

5) *RRE-driven AS RNA-mediated inhibition is improved when ADAR1 expression is reduced.*
.....102

6) *Analysis of evidence of ADAR editing in AS-RNA and target HIV mRNAs with and without ADAR1 or ADAR2 shRNA*.....105

7) *RRE-driven antisense env RNA-mediated inhibition is less effective when cellular huTudor-SN (SND1) is depleted.*108

8) *AS RNA-derived miRNAs are specifically detected in total RNA isolated from cells expressing RRE-driven AS RNA*.....110

Discussion

1) *The analysis of AS-RNA mediated inhibition of HIV points to specific interactions between the ADAR RNA editing machinery and the HIV-1 Rev/RRE pathway*.....116

2) *Natural levels of ADAR editing benefits Rev/RRE-mediated AS RNA inhibition of HIV-1 replication, whereas too much editing impairs it.*118

3) *In the presence of HIV 1 target mRNAs, ADAR RNA editing is seen throughout the 937-nt antisense AS region, whereas only a few of the target mRNAs are edited.*120

Table of Contents (continued)

<i>4) Lower expression level of huADAR1 enzymes improves the RRE-driven AS RNA-mediated inhibition of HIV-1 replication. The improved inhibition is associated with lower level of A-to-G changes (%G) in both the AS RNA and the target RNA.....</i>	122
<i>5) Human Tudor-SN (SND1) may be involved in the inhibition of HIV-1 mediated by the RRE-driven AS-RNA vector.</i>	122
<i>6) 19-23 nt AS RNA-derived miRNAs/siRNAs can be specifically detected in RNA from cells where HIV replication is inhibited by RRE-driven AS RNA.</i>	124

Chapter 4 – Concluding remarks and future directions127

<i>Antisense RNA-based therapy for non-infectious diseases.....</i>	137
---	------------

Chapter 5 – Materials and Methods139

A. Construction of antisense plasmids and nomenclatures.....	139
B. Proviral clones used in the study.....	142
C. shRNA in vitro transcription and shRNA treatment of 293T cells	142
D. Cell lines and transfections.....	145
E. p24 ELISA.....	145
F. Western blot analysis.....	145
G. Total RNA extraction and cytoplasmic RNA fractionation.....	147
H. cDNA synthesis and PCR amplification of antisense target sequence and antisense RNA vector.....	150
I. NGS library preparation and sequencing.....	151

Table of Contents (continued)

J. Small RNA library preparation.....	154
K. NGS data analysis.....	159
References.....	161

List of Figures

Figure 1: An overview of HIV-1 replication cycles and potential targets for antiretroviral intervention.....	16
Figure 2: An overview of HIV-1 DNA synthesis (reverse transcription).....	19
Figure 3: Alternative splicing of HIV-1.....	22
Figure 4: HIV-1 maturation and viral particle.....	25
Figure 5: The HIV-1 Rev/RRE nuclear transport pathway.....	27
Figure 6: MPMV Constitutive Transport Element (CTE).....	29
Figure 7: mRNA export.....	32
Figure 8: Retroviral mRNA exports and the two major RNA export pathways operating in the cells	33
Figure 9: Engineering of Lentiviral vector systems and improvements in performance and biosafety.....	40
Figure 10: Inhibition of particle production from HIV-1 proviruses by antisense RNA.....	46
Figure 11: Target RNAs associated with polyribosome.....	47

List of Figures (continued)

Figure 12: GagPol Pr160 and Nef production in the presence of sense or antisense (AS) vectors	48
Figure 13: RNAi Pathway.....	51
Figure 14: RNase III proteins and their mechanisms.....	52
Figure 15: Schematic diagram of human Argonaute 2 (AGO2), GW182, and Poly(A)-Binding Protein (PABP).....	55
Figure 16: Schematic representation of potential mechanisms for miRNA-mediated translational repression.....	56
Figure 17: Schematic representation of potential mechanisms for miRNA-mediated mRNA decay.	57
Figure 18: ADAR family proteins.....	59
Figure 19: Deamination of Adenosine (A) to Inosine (I) by ADAR.....	60
Figure 20: Recoding events caused by A-to-I Editing on RNA.	61
Figure 21: Interaction between RNA editing and RNA-interference pathways	63
Figure 22: Diagrams showing (A) HIV-1 gene organization and antisense env RNA vectors used in the experiments; (B) RRE-driven antisense env RNA vector; (C) CTE-driven antisense env RNA vector; and (D) antisense env RNA vector with no nuclear exporting element.....	68
Figure 23: Diagrams showing hypothetical structure of antisense <i>env</i> RNA molecules and nuclear exporting elements.	69
Figure 24. A diagram demonstrating the construction of antisense RNA vector lacking a panhandle.	69

List of Figures (continued)

Figure 25: A deletion of the panhandle region of an RRE-driven antisense RNA (AS RNA) vector leads to a ten-fold decrease in the inhibition of HIV replication.....	70
Figure 26: Diagrams of antisense RNA vectors used in an experiment shown in Figure 27	72
Figure 27: A 287-nt env panhandle alone is sufficient to an effective inhibition.....	73
Figure 28: Diagrams showing a proviral clone pHIG and its transcripts.....	75
Figure 29: Inhibition of p24 expression from pHIG with AS vectors expressing EGFP AS RNA with and without a panhandle.....	75
Figure 30: High level of A-to-G changes (%G) in the 287-bp panhandle when the RRE-driven antisense env RNA vector traffics on the Rev/RRE pathway even in the absence of the target. No A-to-G changes were observed in the loop region.....	85
Figure 31. No A-to-G changes were observed in the loop region.	86
Figure 32: Next Generation Sequencing (NGS) data of the antisense repeat (AR) and loop region of cDNA made from RRE-driven antisense env RNA (AS-RRE) in the presence of Rev.....	87
Figure 33: The quantification of A-to-G changes observed in the antisense repeat (AR) and loop region of cDNA made from RRE-driven antisense env RNA (AS-RRE) in the presence of Rev as shown in Figure 32.....	88
Figure 34: An overexpression of wild type huADAR1 p160, huADAR1 p110, or huADAR2s significantly impairs RRE-driven antisense RNA-mediated inhibition of HIV replication.....	90
Figure 35: Overexpression of huADAR2s wt or an N terminus-deleted mutant impairs the inhibition, but this is not observed with huADAR2s lacking dsRNA binding domains or catalytic activity.....	92

List of Figures (continued)

Figure 36: With very effective inhibition, relatively high levels of A-to-G changes are observed throughout the antisense env sequence of the RRE-driven antisense RNA (AS-RRE).....	94
Figure 37: Low-level A-to-G changes were observed in the target region (AST) of the AS-RNA trafficking through the HIV-1 Rev/RRE pathway.....	95
Figure 38: The distribution of individual ADAR-edited (A-to-G changes) clones of the cDNA of AS-RRE RNA target region from total and cytoplasmic HIV mRNA with and without ADAR2s overexpression.....	97
Figure 39: The 5' and 3' nearest-neighbor preferences pattern of ADAR editing in the antisense repeat (AR) of the AS RNA vector (Figure 30).....	100
Figure 40: The 5' and 3' nearest-neighbor preferences pattern of ADAR editing in the sense repeat (SR) of the AS RNA vector (Figure 30).	101
Figure 41: RRE-driven antisense env RNA-mediated inhibition is improved when the producer cells are treated with shRNA against huADAR1.....	103
Figure 42: shRNA against ADAR1 significantly reduces ADAR1 p160 expression and moderately reduce ADAR1 p110 expression.	104
Figure 43: shRNA against huADAR2 in 293T cells overexpressing FLAG-tagged and HIS-tagged huADAR2 (huADAR2-FLIS) are able to significantly reduce ADAR2 expression.....	104
Figure 44: A-to-G changes in the antisense env sequence of the RRE-driven or No RRE antisense RNA (AS-RRE) in the absence or presence of shRNA against ADAR1 or ADAR2....	106
Figure 45: A-to-G changes in the target env sequence in the absence or presence of shRNA against ADAR1 or ADAR2.....	107

List of Figures (continued)

Figure 46: RRE-driven antisense env RNA-mediated inhibition is less effective when the producer cells are treated with shRNA against huTudor-SN.....	109
Figure 47: Alignment of small RNAs present in transfected cells to the HIV AS target region.	114
Figure 48: The diagram demonstrates the hypothetical mechanism of the Rev/RRE-pathway dependent antisense RNA (AS RNA)-mediated inhibition of HIV-1 replication.	130
Figure 49: The diagram demonstrates the hypothetical mechanism explaining that ADAR overexpression impairs the RRE-driven antisense RNA-mediated inhibition of HIV-1 replication.	135

List of Tables

Table 1: The number of adenosine (A) in both the antisense and sense repeats that show A-to-G changes as seen in Figure 30.....	86
Table 2: Analysis of A to G changes in the AS-RRE RNA target region in cDNAs from total and cytoplasmic HIV mRNA with and without ADAR2s overexpression as shown in Figure 38.....	98
Table 3: A table showing the “old” 5’ and 3’ nearest-neighbor preferences of edited Adenosine (A).....	99
Table 4: A table showing the “new” 5’ and 3’ nearest-neighbor preferences of edited Adenosine (A).....	99
Table 5: NGS analysis of small RNAs present in total RNA from cells transfected with pNL4-3 in the presence or absence of AS vectors.	112
Table 6: A list of proviral clones, antisense RNA vectors, and expression plasmids used in this study.....	140

Chapter 1 – Introduction

Human Immunodeficiency Virus (HIV)

Clinical observations of Acquired Immune Deficiency Syndrome (AIDS) were reported in the early 1980s and, in 1983, Human Immunodeficiency Virus (HIV) was isolated from a lymph node biopsy of a patient with lymphadenopathy¹. Rapid molecular cloning of HIV and genome sequencing enabled scientists to identify viral diversity and determine the origin and evolution of HIV. A major finding, based on viral genomic information, showed that there are two types of HIV (HIV-1 and HIV-2). Both are the result of cross-species transmission of Simian Immunodeficiency Viruses (SIV)². Since the mid-1980s, numerous advancements in HIV research have led to effective treatments of HIV infection.

The understanding of HIV-1 replication has led to the discovery of many new antiretroviral therapies. The HIV-1 replication cycle (see **Figure 1**) can be divided into two stages: an early and a late stage. The early stage of HIV-1 replication includes the events from viral entry to the integration of viral DNA into the host genome (**Figure 1**, steps **1** to **6**). The late stage encompasses the events from viral gene expression to the release and maturation of new viruses (**Figure 1**, step **7** to **13**). HIV-1 replication can be summarized as follows.

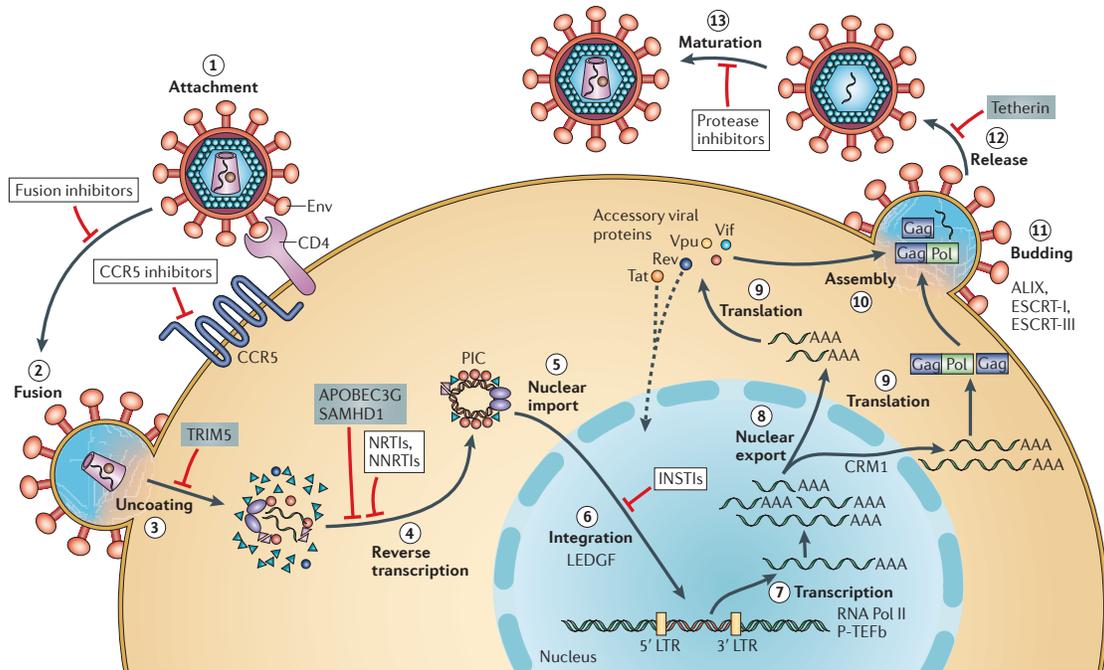


Figure 1. An overview of HIV-1 replication cycles and potential targets for antiretroviral intervention

(1) An HIV-1 virus infects its host cells using the Envelope (Env) glycoprotein spikes to bind to the receptor CD4 and either the membrane-spanning co-receptor CC-Chemokine Receptor 5 (CCR5) or C-X-C Chemokine Receptor type 4 (CXCR-4), which is not show here. (2) The resulting fusion between the viral and cellular membranes leads to viral entry into the cell. (3) In cytoplasm, the viral core shell partially uncoats. (4) Subsequent reverse transcription (5) leads to the formation of a viral Pre-Integration Complex (PIC) which is imported into the nucleus. (6) The HIV-1 integrase in the PIC integrates proviral dsDNA into the host genome, aided by the chromatin-binding protein Lens Epithelium-Derived Growth Factor (LEDGF). (7) Cellular RNA Polymerase II (RNA Pol II) and Positive Transcription Elongation Factor b (P-TEFb) mediate the transcription of HIV-1 mRNAs. Alternative splicing yields 3 species of HIV-1 mRNAs, i.e. ~9-kb unspliced (genomic), ~4-kb singly spliced and ~2-kb multiply spliced mRNAs. (8) The multiply spliced HIV-1 mRNAs are exported to cytoplasm and serve as coding mRNA for Rev, Tat, and Nef. The unspliced and singly spliced species are exported via the HIV-1 Rev/RRE nuclear export pathway which is CRM1-dependent. (9) HIV-1 mRNAs are translated for viral assembly. (10) The ~9-kb unspliced genomic mRNA is incorporated into viral particles with protein components. (11) Cellular ESCRT (Endosomal Sorting Complex Required for Transport) complexes and ALIX mediate viral budding and (12) viral release. (13) HIV-1 protease immediately cleaves HIV-1 gag or gagpol polyproteins to yield mature and infectious virus particle. The sites of action of clinical inhibitors (white boxes) and cellular restriction factors (blue boxes) are indicated. INSTI, INtegrase Strand Transfer Inhibitor; LTR, Long Terminal Repeat; NNRTI, Non-Nucleoside Reverse Transcriptase Inhibitor; NRTI, Nucleoside Reverse Transcriptase Inhibitor. Figure adapted from Engelman A. and Cherepanov Nature Reviews Microbiology 2012³.

Viral entry

Viral entry begins when HIV-1 envelope protein gp120 binds to the cell surface receptor CD4 (**Figure 1**; steps 1 and 2). The binding induces the formation of a bridging sheet between the inner and outer domains of the envelope protein gp120 monomer, exposing the binding site for a co-receptor, which is either the CC-chemokine Receptor 5 (CCR5) or the α -chemokine receptor CXCR4. The binding of the co-receptor leads to insertion of the fusion peptide, located at the amino terminus of the transmembrane envelope protein gp41, into the cell membrane, ultimately resulting in membrane fusion³. Based on the co-receptor usage, HIV-1 can be differentiated into three types of viruses: the X4 viruses which use CXCR4 (predominantly expressed on the cell surface of T-cells); the R5 viruses which use CCR5 (predominantly expressed on the cell surface of macrophages); and the R5X4 viruses which can use both of these co-receptors⁴. A small percentage of Caucasians have a homozygous *CCR5* mutation that makes them resistant to HIV infection. This *CCR5* Δ 32 mutation has been shown to prevent viral entry of the R5 viruses, indicating that this co-receptor is crucial to viral infection².

Uncoating

Following viral entry, the viral core enters the host cell cytoplasm (**Figure 1**; step 3). The HIV-1 viral core is converted to a Reverse Transcription Complex (RTC) and, later, the Pre-Integration Complex (PIC). During the uncoating, HIV-1 Capsid proteins (CA) disassociate from the core. However, some of the Matrix protein (MA), Nucleocapsid (NC), Reverse Transcriptase (RT), Integrase (IN),

and Vpr remain associated with the complex⁴. Interestingly, Rhesus TRIM5 α inhibits this uncoating process in monkey cells. This protein has been shown to work as a restriction factor by promoting rapid uncoating of the incoming HIV-1 Capsid (CA) and negatively affecting PIC formation. TRIM5 α interaction with CA leads to the degradation of CA in a proteasome-independent manner^{5,6}. A chimera of human TRIM21-rhesus TRIM5 α has also been shown to recognize retroviral capsid-like complexes and restrict viral replication⁷.

Reverse Transcription

A major defining feature of retroviruses is their ability to convert their RNA genomes into dsDNA in the cytoplasm of the infected cells through the use of the Reverse Transcriptase (RT) enzyme (**Figure 1**; step 4). HIV-1 RT is a heterodimer of two subunits, p66 and p51. p66 and p51 are derived from the same region of the Pr160^{GagPol} precursor⁴. The process of HIV-1 reverse transcription is shown in **Figure 2**.

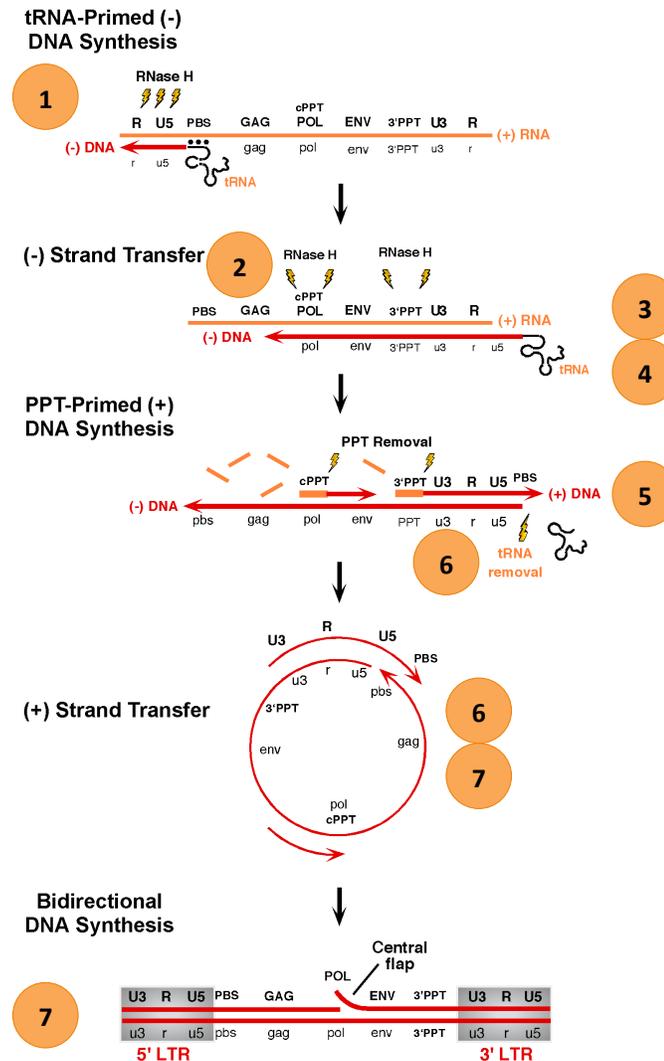


Figure 2. An overview of HIV-1 DNA synthesis (reverse transcription)

Reverse transcription is described stepwise as follows: **(1)** HIV dsDNA synthesis starts with the use of a tRNA, a primer that binds to the primer binding site (pbs). The synthesis proceeds to the 5' end of the RNA molecule and generates a DNA/RNA hybrid. **(2)** HIV RT's RNase H degrades the RNA portion of this DNA/RNA hybrid and generates a DNA fragment known as the minus-strand strong stop DNA. **(3)** The first strand transfer takes place when the minus-strand strong stop DNA jumps from the 5' to the 3' end of the genome by using a short region of homology called "R". **(4)** Minus-strand DNA synthesis continues, using the 3' end of the minus-strand strong stop DNA as a primer. **(5)** Plus-strand DNA synthesis occurs with the use of RNA fragments remaining from minus-strand synthesis as primers. These are purine-rich sequences known as the PolyPurine Tracts (PPTs). **(6)** The tRNA bound to the pbs is removed by RT's RNaseH, thereby allowing second-strand transfer to take place. **(7)** Plus-strand DNA synthesis proceeds to the end of the minus strand. For HIV, an additional termination site (the Central Termination Signal or CTS) is located near the center of the genome. Since the CTS is 3' of the central PPT, approximately 100 nucleotides of plus-strand DNA is displaced, resulting in the formation of a DNA "flap". The central flap is known to play a role in the import of the viral PIC to the nucleus. Figure adapted from LeGrice S.F.J. JBC 2012⁸.

Two classes of antiretroviral drugs targeting RT are Nucleoside and Non-Nucleoside RT Inhibitors (NRTIs and NNRTIs, respectively). Both drugs inhibit DNA polymerization and are core components of highly active antiretroviral therapy (HAART). Non - Nucleoside RT Inhibitors (NNRTIs) are allosteric inhibitors, whereas Nucleoside RT Inhibitors (NRTIs) are chain terminators and competitive inhibitors. However, HIV-1 can develop resistance to both classes of drugs through drug exclusion and other mechanisms. HIV-1 reverse transcription can also be influenced by a cellular restriction factor called Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3G (APOBEC3G)⁹⁻¹¹. This cytidine deaminase is incorporated into viral particles and introduces mutations during reverse transcription by converting cytidines in the viral cDNA to uracils¹²⁻¹⁴. HIV-1 uses a special viral protein Vif to counteract APOBEC3G. HIV-1 Vif antagonizes APOBEC3G by inducing the degradation of APOBEC3G via ubiquitylation^{15,16}. The understanding of this process has led to the development of small molecule drugs that can inhibit Vif-mediated degradation of APOBEC3G^{17,18}. However, these have not yet been tested in human trials.

Integration

The HIV-1 Pre-Integration Complex (PIC), harboring the dsDNA provirus, is imported into nucleus (**Figure 1**; steps 5 and 6). The HIV-1 Integrase (IN) then integrates the provirus into the host genome. During the integration, IN cleaves each end of proviral Long Terminal Repeat (LTR) to generate recessed 3' termini. IN then cuts chromosomal DNA across a major groove while joining the viral DNA end to target DNA 5'-phosphates. The integration is completed when

the single-strand gaps, bordering the unjoined viral DNA 5' ends, are repaired^{3,19}. The integration sites of Lentiviruses, such as HIV-1, have been shown to mainly be within active genes due to the interaction between IN and the host cell chromatin-binding protein Lens Epithelium-Derived Growth Factor (LEDGF or p75)^{3,20-22}.

HIV gene expression and viral mRNA export

The HIV-1 Long Terminal Repeat (LTR) that is generated by reverse transcription (**Figure 2**) contains three regions that are referred to as U3 (for Unique, 3' end), R (for Repeated) and U5 (for Unique, 5' end). It also contains the *cis*-acting elements required for RNA synthesis. Transcription initiates at the U3/ R junction of the LTR. The U3 region contains binding sites to several host cell transcription factors that direct the binding of RNA Polymerase II (RNA pol II) to the DNA template. Then, the host cell Transcription Factor IID (TFIID) binds to a TATA box, located approximately 25 nucleotides upstream of the transcription start site. Three Sp1 and two NF- κ B binding sites are located 5' upstream of the TATA box in the U3 region⁴. Other transcription factor binding sites (i.e. including LEF, Ets, and USF binding sites) also are located in the U3 region. It is known that HIV-1 basal transcription activity is relatively low. However, the transcription rate is greatly increased in the presence of HIV-1 transcriptional transactivator protein Tat. The HIV-1 Tat promotes efficient transcriptional elongation⁴. Three "classes" of HIV-1 mRNA are expressed due to alternative splicing²³ (see step 7 and 8 of **Figure 1** and **Figure 3**).

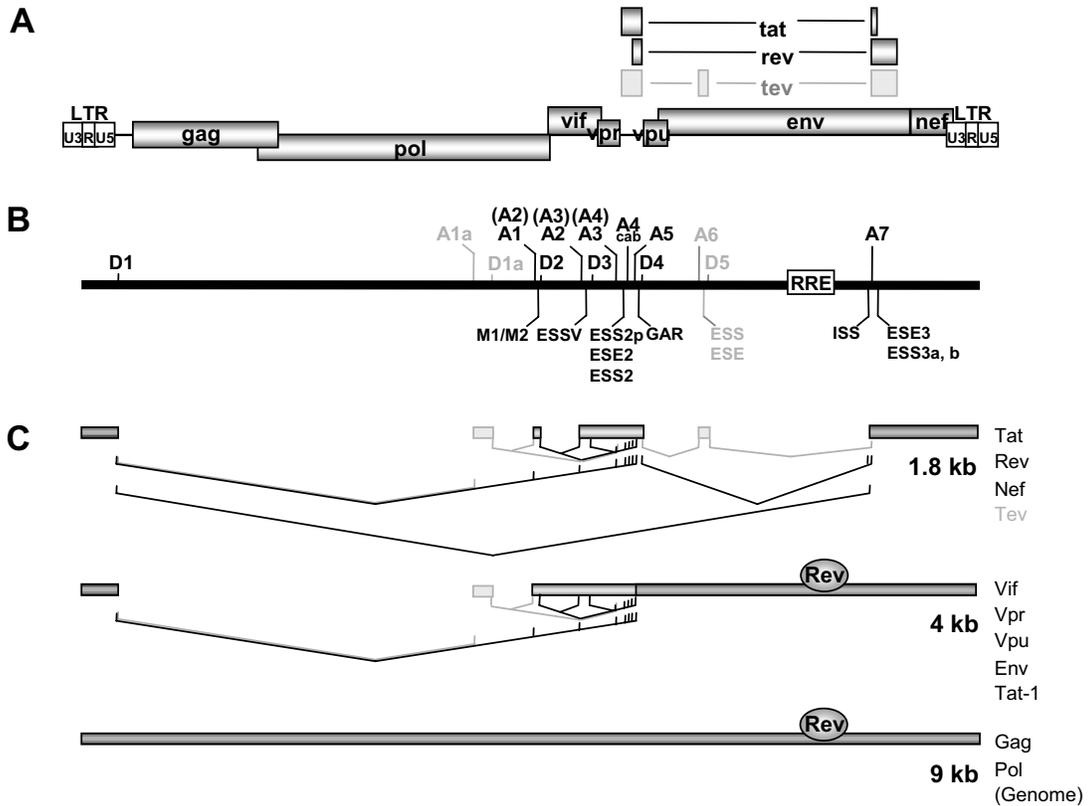


Figure 3. Alternative splicing of HIV-1

(A) Organization of the HIV-1 genome. Filled boxes indicate Open Reading Frames (ORFs) of HIV genes in all isolates, light grey boxes indicate the isolate-specific Tev ORF. **(B)** Splice sites, splicing regulatory elements and the Rev Responsive Element (RRE). 5' splice sites: D1a-5; 3' splice sites: A1-7. Splice sites A6/D5 are isolate specific and not functional in the isolate NL4-3. Splicing regulatory elements: M1, M2 (heptameric sequences); ESSV; ESS2p; ESE2/ESS2; GAR; ESS/ESE; ISS; ESE3; ESS3a, b. **(C)** Splicing pattern and proteins encoded by the different mRNA classes. The 1.8 and 4 kb mRNAs contain obligatory sequences (dark grey) as well as alternative sequences (light grey) due to alternative splicing. The nuclear export of the 4 kb mRNAs and the genomic full-length (genome) 9 kb mRNA is dependent on Rev binding. Figure adapted from Kammler et al. *Retrovirology* 2006²³.

The recent development of the Primer ID-tagged deep sequencing to quantify HIV-1 splicing²⁴ and older data²⁵ has revealed that HIV-1 mRNA undergoes splicing using four 5' splice sites (donors) and 10 3' splice sites (acceptors) to generate over 50 physiologically relevant transcripts in two size classes (~2 kb and ~4 kb). The smaller HIV-1 transcripts (~2 kb) are readily exported to the cytoplasm using the host cell machinery. However, the large unspliced (~9 kb) and singly spliced (~4 kb) HIV-1 mRNAs require the viral Rev protein for nuclear export via the REV/RRE pathway. After the translation, the Rev protein is imported into the nucleus, where it recognizes, binds, and multimerizes on the Rev Response Element (RRE), found in both the unspliced and singly-spliced viral transcripts. Subsequently, Rev acts as an adaptor for the host nuclear export factor CRM1 (or XPO1). The details of the HIV-1 Rev/RRE nuclear export pathway are discussed later in "Nuclear RNA Export Pathways".

Viral assembly and release

The viral assembly process begins at the plasma membrane following viral protein synthesis, as shown (**Figure 1**; steps 9 and 10). The HIV-1 Gag precursor polyprotein Pr55^{Gag} interacts with plasma membrane via its phosphatidylinositol-4,5-bisphosphate binding domain and myristoyl moiety²⁶. HIV-1 Pr55^{Gag} promotes Gag-Gag and Gag-Env interactions, encapsulation of the viral RNA genome, and viral release from the host cell⁴. In addition, the host cell's class E Vacuolar Protein Sorting (VPS) machinery helps to orchestrate viral release. Most class E VPS proteins are subunits of ESCRT (Endosomal Sorting Complex Required for Transport) complexes. Both ESCRT-I and ESCRT-II

function during membrane budding. ESCRT-III is required for membrane scission and the releasing of nascent viral particles²⁷ (**Figure 1**; steps 11 and 12). Another cellular protein, Tetherin (CD317 and BST2), inhibits viral release by retaining a nascent viral particle to the plasma membrane^{28,29}. However, the HIV-1 transmembrane protein Vpu can circumvent Tetherin-mediated restriction^{30,31}.

Viral maturation

This last step of HIV replication is mediated by the viral Protease (PR) (see **Figure 1**; step 13). Viral maturation occurs concomitant with viral release. This process converts immature viral particles to infectious viral particles via the proteolytic cleavage of HIV-1 Gag and GagPol to yield the individual structural components Matrix (MA), p24 Capsid (CA) and Nucleocapsid (NC), and the enzymes Reverse Transcriptase (RT), Integrase (IN) and Protease (PR)^{32,33} (see viral maturation and diagrams in **Figure 4**). Several antiretroviral drugs have been developed to inhibit HIV protease and many are approved for clinical use³⁴.

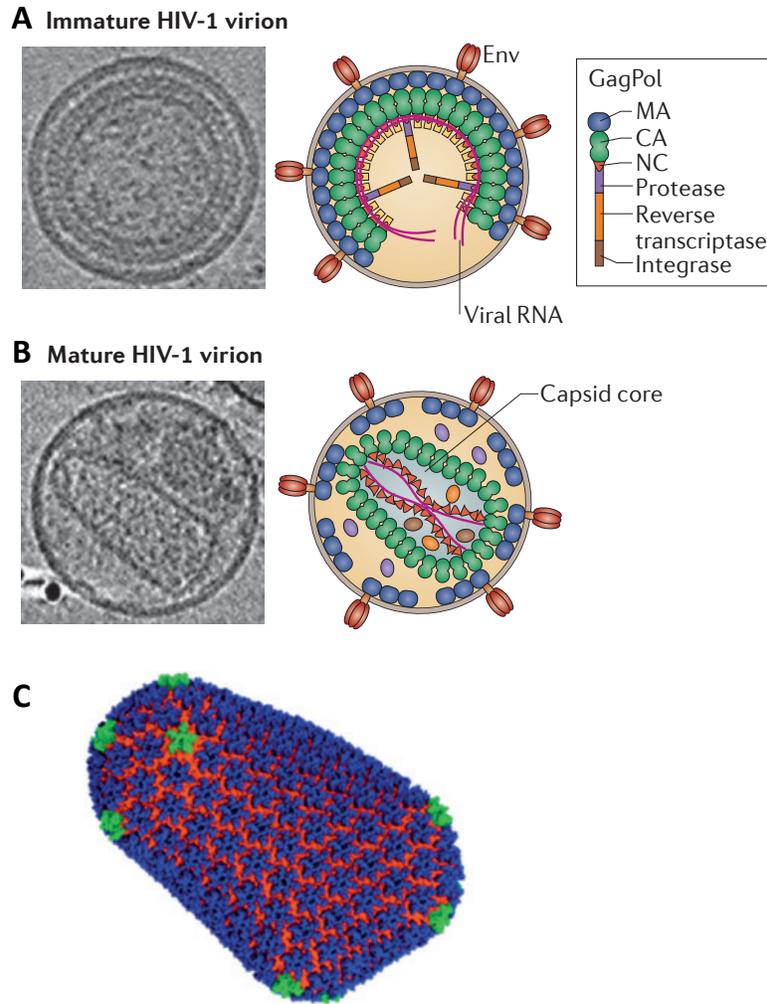


Figure 4. HIV-1 maturation and viral particle

During virus release, the viral protease cleaves a number of sites in both the Gag and GagPol polyproteins to initiate virus maturation. This leads to major changes in virion morphology, including the generation of the conical capsid core. **(A)** A cryoelectron tomogram (left panel) and an illustration (right panel) of the immature HIV-1 virion. **(B)** A cryoelectron tomogram (left panel) and an illustration (right panel) of the mature HIV-1 virion. **(C)** The all-atom structure of an HIV-1 capsid core, with CA pentamers (green) in the otherwise hexameric lattice. Env, Envelope; MA, Matrix; NC, Nucleocapsid. Figure adapted from Freed E.O. *Nature Reviews Microbiology* 2015³⁵.

RNA Export Pathways

The HIV-1 Rev/RRE Export Pathway

HIV-1 replication requires the expression of unspliced, incompletely spliced, and fully spliced transcripts. Analysis of the different HIV-1 mRNAs has demonstrated that the fully spliced mRNAs encode the viral regulatory proteins Tat, Rev, and Nef, whereas incompletely and unspliced HIV-1 mRNAs encode the viral structural proteins and the accessory proteins Vif, Vpr, and Vpu. In the absence of Rev function, the unspliced and incompletely spliced HIV-1 mRNAs are retained in the nucleus, whereas fully spliced HIV-1 mRNAs are normally exported and expressed^{36,37}. Nuclear export of the Rev-dependent mRNAs requires a structured *cis*-acting RNA target in these RNAs, the Rev Response Element (RRE), which serves as a binding site for Rev³⁶. Thus the HIV-1 Rev/RRE pathway is characterized by its ability to overcome nuclear retention of mRNAs with retained introns to allow their nuclear export³⁶⁻³⁹.

The HIV-1 Rev protein contains two important functional domains: an N-terminal sequence required for RRE binding and Rev multimerization; and a ~10-amino acid leucine-rich domain closer to the C terminus that acts as a nuclear export signal (NES)^{40,41}. The Rev NES directly interacts with Xpo1 (Crm1), a member of the karyopherin family of nucleocytoplasmic-transport factors^{42,43}. The Rev/RRE/Xpo1 complex also recruits RanGTP. After nuclear export, hydrolysis of the bound GTP to GDP allows for conformational changes that

result in cargo release⁴⁴ (see **Figure 5 and 8** for a schematic representation of the HIV-1 Rev/RRE nuclear transport pathway).

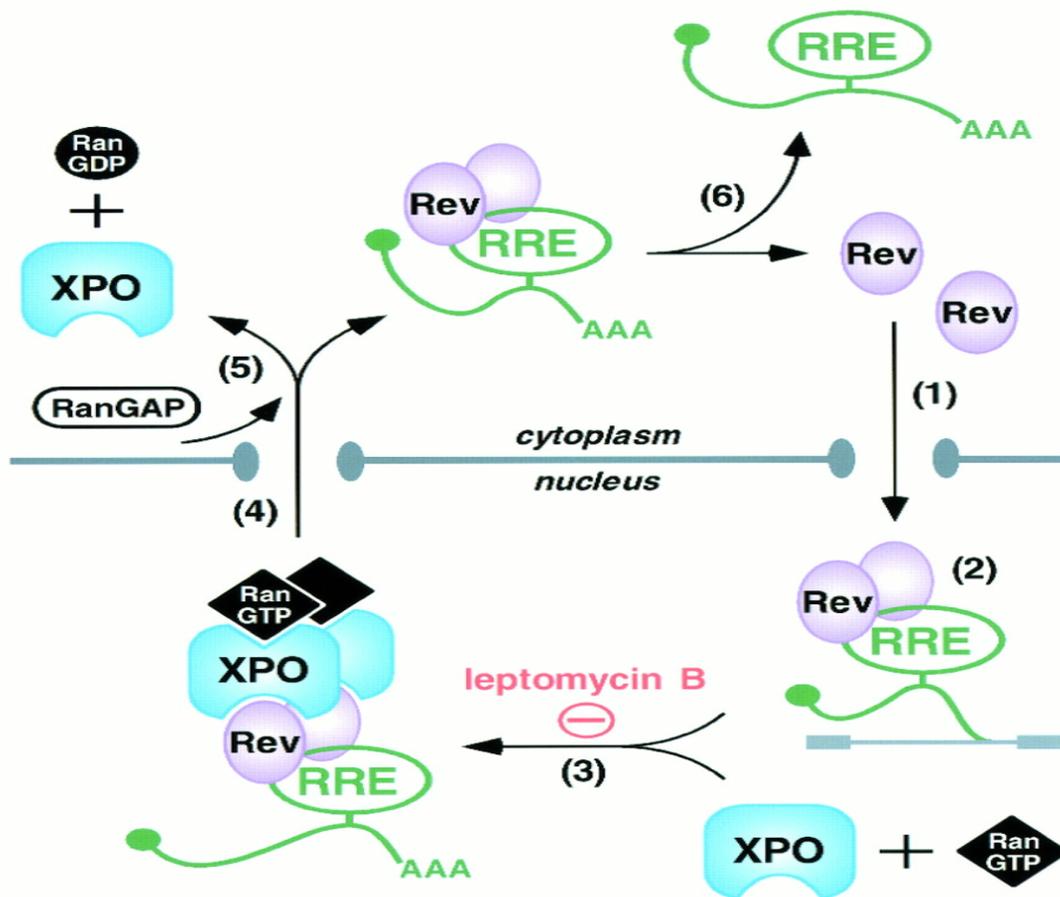


Figure 5. The HIV-1 Rev/RRE nuclear transport pathway

Upon entry into the nucleus (1), Rev proteins bind and multimerize on the Rev Response Element (RRE) of unspliced and incompletely spliced HIV-1 transcripts (2). Xpo1 (or Crm1) and Ran GTP bind cooperatively to Rev-RRE complexes, which can be antagonized by leptomycin B, (3) and mediate nuclear export (4). In the cytoplasm, Ran GTPase-Activating Protein (RanGAP) stimulates GTP hydrolysis and triggers the dissociation of Xpo1 and Ran GDP from Rev (5). Finally, Rev shuttles back to the nucleus (1), and the exported RNA is either translated or packaged into virions (6). Figure from Emerman M and Malim MH Science 1998⁴⁵.

The MPMV CTE-mediated nuclear export pathway

The Mason-Pfizer Monkey Virus (MPMV) is a simple retrovirus, and its genome encodes only Gag, Pol, and Env proteins. Despite its simple genomic organization, MPMV has to export unspliced as well as spliced mRNAs to the cytoplasm of the host cells. The unspliced mRNA serves as the MPMV genome and the mRNA for the Gag and Pol proteins, whereas the fully spliced mRNA encodes the Env proteins. Like HIV-1, MPMV thus has to overcome nuclear retention of intron retaining genomic RNA. A region mapped to a 154-nt region at the 3' end of the MPMV genome has been shown to provide a function similar to HIV-1 Rev/RRE⁴⁶⁻⁴⁹. This element was given the name Constitutive Transport Element (CTE), since it functions in the absence of any viral protein. Unlike HIV-1, which utilizes the virus-encoded Rev protein for nuclear export, the MPMV uses a host cell protein, Nxf1 (or Tap), to achieve nuclear export of the intron retaining genomic RNA (see **Figure 6 and 8** for a schematic representation of the MPMV CTE pathway). The Nxf1 protein has been shown to make a direct interaction with the Mason-Pfizer Monkey Virus (MPMV) CTE^{50,51}. Additionally, Nxt1 and Nxf1 proteins acting *in trans* and the MPMV CTE *in cis* have been shown to functionally replace the HIV Rev protein and the HIV-1 RRE in HIV replication^{46,49}. Unlike Xpo1/Crm1, Nxf1 (Tap) mediates nuclear export in a Ran-independent manner⁵².

Interestingly, intron 10 of the human NXF1 gene contains a sequence exhibiting striking primary sequence and secondary structure homology with the MPMV CTE⁵³. The presence of the human "CTE" in intron 10, together with the

human Nxf1 protein and its cofactor, the human Nxt1 protein, has been shown to facilitate the nucleocytoplasmic export and expression of an alternatively spliced NXF1 mRNA that retains intron 10. This NXF1 mRNA, with retained intron 10, is translated into a short isoform of the Nxf1 protein (small Nxf1 or sNxf1) in both human and rodent cells^{53,54}. The sNxf1 protein has recently been detected in several normal mammalian tissues, for example rodent hippocampal and cortical neurons. The sNXF1 protein is also observed in cytoplasmic granules in some neurons in the rodent neocortex. Furthermore, the sNxf1 has been shown to function as an alternative Nxf1 cofactor in export and translation of CTE-containing mRNAs found colocalized with Staufen2 proteins in cytoplasmic granules or polyribosomes⁵⁴.

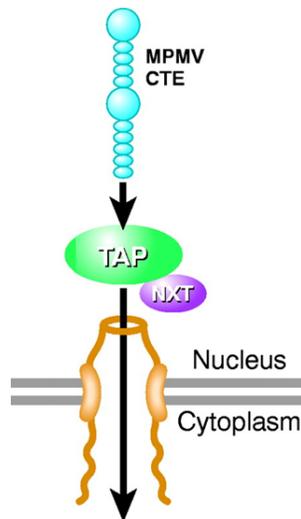


Figure 6. MPMV Constitutive Transport Element (CTE)

Nuclear export of retrovirus Mason Pfizer Monkey Virus (MPMV) mRNAs is independent of the Ran system and karyopherins. Incompletely spliced MPMV mRNAs contain a cis-acting RNA sequence named the Constitutive Transport Element (CTE). The key mediator of export function is a heterodimer of Nxf1/Tap (here called Tap) and a small cofactor termed Nxt/p15. In yeast cells, the Nxf1/Tap ortholog is Mex67p, whereas Nxt1/p15 is Mtr2p. The host cell's Nxf1/Nxt heterodimer binds to the CTE tp mediate nuclear export. Figure from Cullen BR. Journal of Cell Science 2003²³².

The NXF CTE element is conserved in many mammalian NXF genes as well as in teleost fish, including *Latimeria chalumnae* (a coelacanth) and *Danio rerio* (zebrafish). The CTE element of zebrafish is present in the same intron as intron 10 of the human NXF1 gene, and the intron-retained NXF1 mRNA is stably

expressed. A functional study has also demonstrated that an HIV GagPol reporter construct containing the CTE of zebrafish functions efficiently in human cells in the presence of the zebra fish Nxf1 and Nxt proteins⁵⁵.

An Overview of host cell mRNA Export

Prior to nuclear export, nascent transcripts are processed through four different events, namely 5' capping, splicing, 3'end cleavage, and polyadenylation, in order to yield mature mRNAs. Each of the RNA processing events has links to nuclear export. Unprocessed or incompletely processed mRNAs are subjected to degradation in the nucleus. Additionally, each processing event has been proposed to trigger the recruitment of protein factors that are necessary for export (see reviews^{56,57}).

Capping and splicing are both essential to the recruitment of the Transcription-Export (TREX) complex, which has been proposed as a requirement for mRNA nuclear export^{58,59} in both yeast and mammalian cells, although this remains controversial. In *S. cerevisiae*, the TREX complex consists of the THO complex and nuclear export factors; for instance Sub2 (an ATP-dependent DEAD-box RNA helicase), Yra1, and Tex1. In human cells, the TREX complex contains Uap6 and Aly, proteins analogous to Sub2 and Yra1, respectively. It has been demonstrated that the TREX complex is poorly recruited to transcripts lacking either the 5' cap or the Exon Junction Complex (EJC). 3' end cleavage and polyadenylation of premature mRNAs yield mature mRNAs whose 3' ends are the sites for the assembly of protein complex that

includes Nuclear Pore Complex (NPC) components (see review by Carmody et al. 2009⁵⁷).

In metazoa, nuclear export of many mature mRNAs appears to be mediated via the non-karyopherin heterodimer of Nxf1 (Tap) and Nxt1 (p15). Interestingly, Nxf1/Nxt1-mediated nuclear export does not depend on the RanGTP gradient for directionality. During nuclear export complex assembly, ATP-bound UAP56 recruits Aly to the mRNP. ATP hydrolysis by Uap56 then allows for the transfer of the Aly to the mRNA. Nxf1-Nxt1 heterodimers bind to Aly, and the RNA-binding affinity of Nxf1 is increased in the presence of Aly; see review by Köhler et al. 2007⁵⁶, Carmody et al. 2009⁵⁷, and Björk and Wieslander 2017⁶⁰. Though many mRNAs appear to be exported in an Nxf1/Nxt1-dependent manner, nuclear export of other mRNAs is Crm1-dependent. Since Crm1 lacks direct RNA-binding properties, Crm1 relies on a set of adaptor proteins for exporting its cargo. HuR is known to be one of these adaptor proteins and has been shown to be involved in, for example, nuclear export of *CD83* and *FOS* mRNAs. Another example of an adaptor protein is the eukaryotic translation initiation factor 4E (eIF4E), which mediates *CYCLIN D1* mRNA nuclear export in human cells.^{61,62,63}

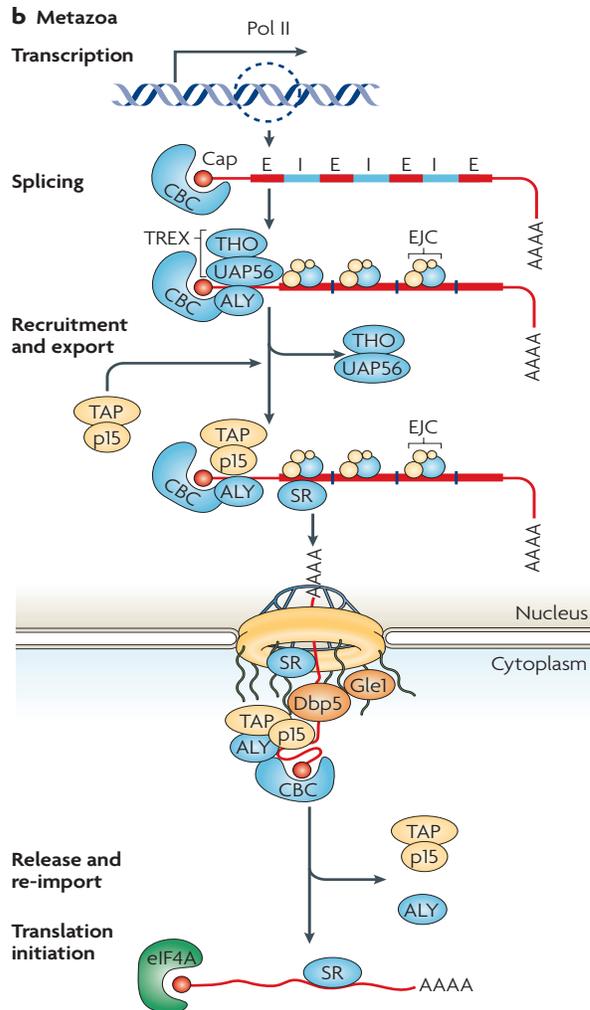


Figure 7. mRNA export

Depicted are the splicing and cap-dependent modes of human TREX recruitment to the mRNP. The downstream events in mRNA export, including the recruitment of the Tap-p15 (Nxf1/Nxt1) mRNA export receptor, are similar for metazoa and yeast. Export receptors are indicated in yellow and adaptors in blue. For simplicity, the Exon-Junction Complex (EJC) on the translocating and cytoplasmic mRNP is not shown. E, Exon; I, Intron; TREX, Transcription-coupled Export complex. Figure from Alwin Köhler and Ed Hurt *Nat Rev Mol Cell Biol* 2007⁵⁶.

Following the assembly of nuclear export complex on mRNAs, the export-competent mRNPs are targeted to the Nuclear Pore Complex (NPC). The export receptors interact with a particular class of NPC proteins known as FG-Nups. When emerging at the cytoplasmic site of the nuclear pore complex, the mRNPs encounter an NPC fibril-associating complex consisting of ATP-dependent RNA helicase Dbp5 and its activator Gle1, and the signaling molecule inositol hexakisphosphate (InsP₆). InsP₆-bound Gle1 stimulates the ATPase activity of Dbp5, thereby converting Dbp5 from ATP- to ADP-bound state. The resulting conformational change brings about the removal of a set of exporting complexes,

for instance Nxf1/Nxt1 and Aly. The mRNAs then advance to translation in cytoplasm; see review by Köhler et al. 2007⁵⁶ and Carmody et al. 2009⁵⁷; see **Figure 7** for a schematic representation of cellular mRNA export. Figure 8 provides a summary of the two retroviral mRNA export pathways for the export of RNA with retained introns and their relationship to the normal cellular mRNA export pathways.

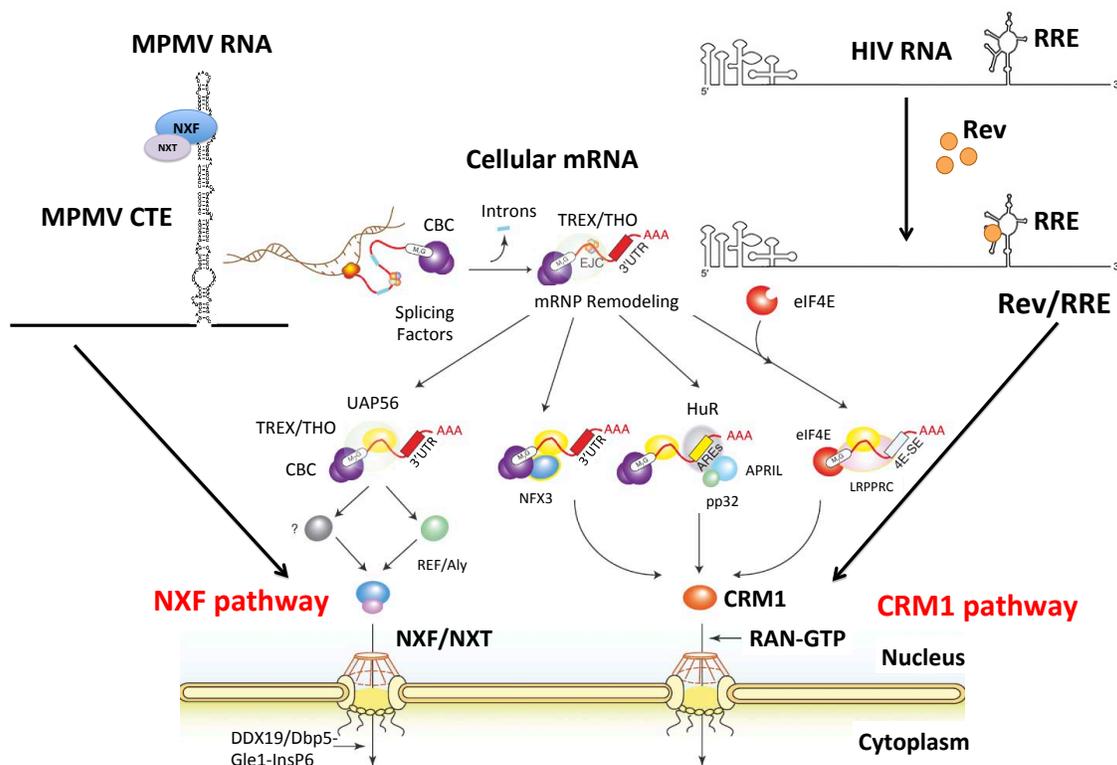


Figure 8. Retroviral mRNA export and the two major RNA export pathways operating in the cells

A simple retrovirus, Mason Pfizer Monkey Virus (MPMV), exports intron-containing mRNAs through the NXF pathway. A lentivirus, human immunodeficiency virus 1 (HIV-1), uses HIV-1 Rev to export intron-containing viral mRNAs through the CRM1 export receptor. Figure adapted from Siddiqui N and Borden KL. WIREs RNA 2012⁶⁴.

Naturally occurring antisense RNAs (AS RNAs)

The presence and functional importance of antisense RNAs (AS RNAs) have been demonstrated in studies in many species⁶⁵⁻⁷². Several AS RNAs have been shown to regulate gene expression via various mechanisms. For example, they have been shown to play important roles in X-chromosome inactivation (Tsix), genomic imprinting (Air), and trans-acting regulation of sense RNA expression (HOTAIR and ANRIL)^{65,72-78}. Another example of naturally occurring AS RNA-mediated regulation of gene expression is that of inhibition of BCMA, a tumor necrosis factor receptor homologue, important in B-cell development⁷⁹. The antisense RNA of this gene is transcribed from the same locus as the sense RNA and has typical mRNA features, including an open reading frame. Overexpression experiments of the AS RNA leads to a reduction of BCMA protein levels, but not the mRNA levels⁷⁹. RNA editing is also observed in the BCMA mRNA. This suggests an involvement of double-stranded RNA-specific adenosine deaminases (ADARs) in inhibition.⁷⁹ With the discovery of a large number of new AS RNAs through unbiased transcriptome sequencing, the overall importance of AS RNA regulation is just starting to be realized. However, abnormal expression of AS RNAs has already been proposed to contribute to several illnesses, including alpha-thalassemia, cardiac disease, and Alzheimer's disease^{73,80,81}.

Naturally occurring retroviral antisense RNAs

Several studies have shown evidence of naturally occurring antisense transcription in many retroviruses, including lentiviruses, delta retroviruses, gamma retroviruses, and beta retroviruses. Naturally occurring HIV-1 antisense RNA (AS RNA) was originally reported to include a region complementary to the envelope gene (*env*)⁸². Additional investigations identified the promoter for this AS-RNA in the 3' LTR (U3-R) region of the provirus^{83,84}. This HIV-1 AS RNA is constitutively expressed in various cells chronically infected with HIV-1, as well as in HIV-1 infected human PBMCs⁸⁴⁻⁸⁸. The Asp-1 protein, translated from an ORF in the HIV-1 AS RNA, has been shown to induce autophagy. Human T-Lymphotropic Virus 1 (HTLV-1) expresses an AS RNA known as HBZ, which has been reported to be involved in viral RNA regulation and leukemogenesis (see review by Matsuoka and Jeang⁸⁹). Other studies have reported the expression of viral AS RNAs by feline immunodeficiency virus (FIV)⁹⁰ and Moloney Murine Leukemia Virus (M-MLV)⁹¹. Retroviral AS RNAs may thus play integral roles during viral replication.

The transcription of naturally occurring antisense HIV-1 RNA

It is known that the HIV-1 3' LTR promoter is essential for driving AS RNAs transcription. Transcription initiates at multiple positions in the U3 region of the 3' LTR^{82,86,92}. Kobayashi-Ishihara et al.⁹³ reported that four HIV-1 antisense transcripts can be expressed in 293T cells, transiently transfected with an expression plasmid containing proviral HIV-1 NL4-3 DNA with a heterologous

promoter driving antisense transcription⁸⁶. Furthermore, several host transcription factors have been shown to induce HIV-1 AS RNAs, including Specificity Protein-1 (Sp1), Upstream Stimulating Factor (Usf), and Nuclear Factor-kappa B (NF-κB)^{86,94–96}.

The HIV-1 3' LTR contains several NF-κB binding sites, known to be involved in the transcription of AS-RNA. The results of dose–response effects of TNF-α on the antisense LTR confirmed the involvement of NF-κB in the regulation of HIV-1 3' LTR promoter activity in the antisense orientation⁸⁶. Only mutations in the NF-κB site and not mutations in the TATA box in the HIV-1 3' LTR promoter disrupted promoter activity in the antisense orientation⁸⁶. A separate study reported that the lack of a TATA box in the HIV-1 3'LTR results in staggered multiple transcription initiation sites⁸⁷. Also, the presence of initiator (INR) motifs (YYANWYY) could serve as alternative transcription initiation sites of HIV-1 AS RNA^{84,97}. A further analysis by Manghera et al suggested additional HIV-1 AS-RNA promoters⁹². It is possible that the multiple transcription factor binding sites in HIV-1 3' LTRs allow for differential, regulated expression of HIV-1 AS-RNAs to enhance viral fitness, infectivity and spread.

HIV-1 AS RNA and its potential role in viral latency

HIV-1 AS RNA has also been proposed to also play a role in establishing viral latency. AS RNA was detected in the acutely or chronically infected cell lines, as well as acutely infected human peripheral blood mononuclear cells. For example, it was detected in Molt-4 acutely infected with HIV-1 NL4-3, as well as

HIV-1_{IIIB}-chronically infected ACH-2, and OM10.1 cells lines, and HIV-1-infected PHA-activated PBMCs⁸⁶. Further investigation also identified two new forms of AS RNAs: ASP RNA-Long variant 1 and 2 (ASP-L1 and ASP-L2) [GenBank: JQ866626]. The major variant, ASP-L1, is transcribed from nucleotide position 9451 in the 3' LTR U3 region of HIV DNA, and terminates at nucleotide position 6878 in the env region. As for the minor variant ASP-L2, the transcription is terminated at nucleotide position 6783⁸⁶. The majority of HIV-1 AS RNA (ASP-L) could be found in the nuclei of various HIV-1 infected cells, as well as in primary PHA-activated PBMCs⁸⁶.

Kobayashi-Ishihara et al. have also shown that HIV-1 AS RNA (ASP-L) inhibits HIV-1 replication *in vitro* in chronically infected cells and PBMCs⁸⁶. Viral production, measured by Reverse Transcriptase-activity assays (RT assays), was reduced in the ASP-L-expressing cells. A repression of viral replication in ASP-L stably expressing cells infected with HIV-1 could be achieved, and lasted for more than 30 days. Furthermore, only four days after infection, a 5-fold reduction of HIV-1 gag and tat transcripts was shown in ASP-L-expressing cells. When the specific shRNAs directed against the antisense HIV RNA (ASP-L) were used, the repression of viral replication was relieved. These findings suggest the possibility that HIV-1 AS RNA could play a key role in establishing viral latency. Understanding the mechanism of AS RNA regulation may thus, eventually, lead to the development of novel strategies for prevention of disease, and ultimately, a cure.

Lentiviral vectors

Gene therapy treats diseases or genetic disorders by introducing therapeutic genetic material into target cells. Retroviruses stably integrate their genome into the genome of the host cell and modified retroviruses can be used as vectors to permanently alter gene expression. Many different retroviral vectors have been developed. The first retroviral vectors were based on gamma-retrovirus vectors⁹⁸⁻¹⁰⁰. The first approved clinical trial in 1991 used a gamma-retroviral vector to treat Severe Combined Immunodeficiency (SCID). Peripheral blood CD34+ cells from patients were transduced with a gamma-retrovirus driving the expression of adenosine deaminase⁹⁸. In 2000, X-linked Severe Combined Immunodeficiency (X-SCID) was corrected in 11 children using a murine leukemia virus-based retroviral vector that drives the common interleukin receptor c-chain in bone marrow⁹⁹. Unfortunately, several cases of leukemia directly caused by the gene therapy were later reported due to retrovirus-induced insertional mutagenesis¹⁰⁰. Alternatively, Lentiviral vectors have been developed for gene therapy. Although Lentiviral vectors can induce insertional mutagenesis, it seems to occur at a frequency lower than that of the gamma retroviral vectors, due to different integration patterns¹⁰¹. Lentiviral vectors can also transduce quiescent cells¹⁰². Several modifications have been made in order to improve the efficacy and safety of Lentiviral vectors¹⁰³.

Three generations of Lentiviral vectors

Lentiviral vectors are divided into three generations according to the packaging system used during viral production. The first-generation packaging plasmid contains all gag and pol sequences, the viral regulatory genes tat and rev, and the accessory genes vif, vpr, vpu, and nef. The second-generation packaging plasmid lacks all four accessory genes (i.e. vif, vpr, vpu, and nef) to improve its safety. The safety is further improved in the third generation. The third generation Lentiviral vector system consists of a split-genome packaging system, in which the rev gene is expressed from a separate plasmid and the 5' LTR of the transfer vector is replaced by a strong tat-independent constitutive promoter¹⁰⁴. These different vector systems are shown in **Figure 9**¹⁰⁵. Recently, non-integrating Lentiviral vectors (NILV) have been developed to avoid insertional mutagenesis by mutating the viral integrase or the integrase attachment sites in the 5' U3 region of the transfer vector plasmid¹⁰⁶. The resulting vectors remain in the nucleus as an episome. NILVs can be used in post-mitotic cells (i.e. retina, brain, and muscle) where persistent gene expression is not required¹⁰⁶.

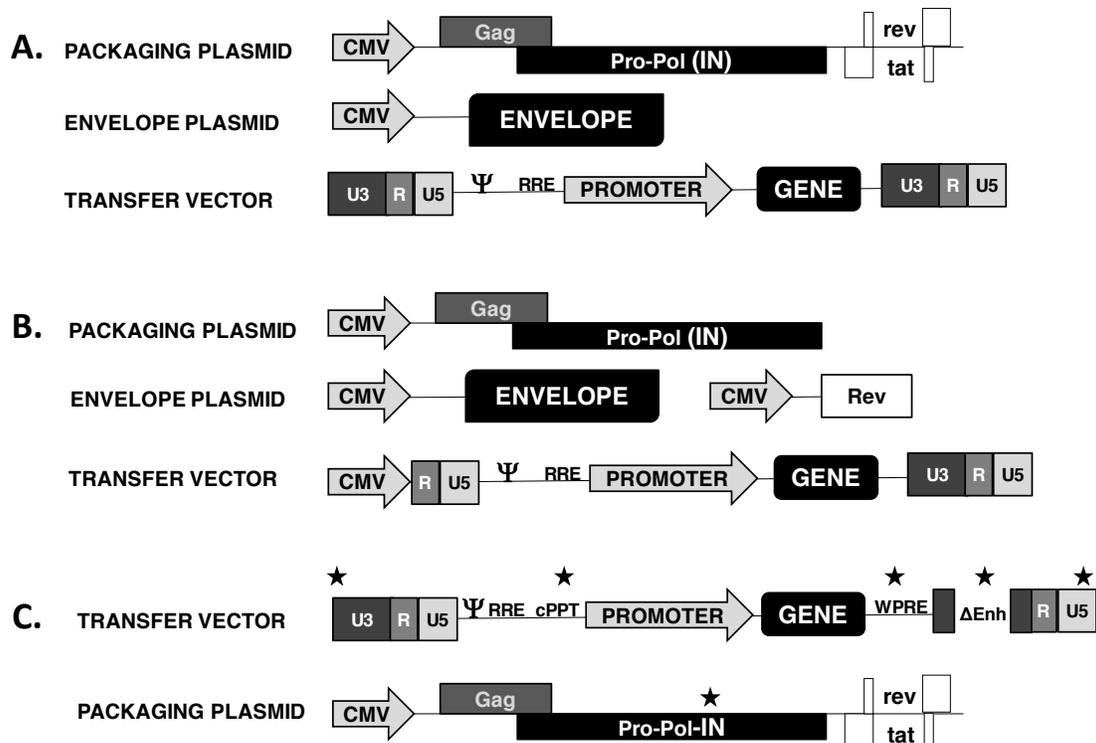


Figure 9. Engineering of Lentiviral vector systems and improvements in performance and biosafety.

(A) The second-generation Lentiviral vector system consists of three plasmids. In many cases the “helper” plasmids contain a strong constitutive promoter, such as cytomegalovirus promoter (CMV) to drive expression of the inserted genes. The packaging plasmid expresses the Gag and Gag/Pol proteins required for assembly and reverse transcription/integration (structural plasmid), as well as Rev and Tat. The envelope plasmid expresses the VSV-G envelope glycoprotein for vector pseudotyping. The transfer vector plasmid lacks all HIV-genes, but provides all the necessary *cis*-acting sequences. The transfer vector contains two Long-Terminal Repeats (LTR) and an expression cassette. The gene cassette contains an internal promoter of choice controlling the expression of the gene of interest. The packaging signal is represented as Ψ and the Rev Response Element as RRE. **(B)** The third-generation Lentiviral vector system is shown. This system is similar to the second generation. However, the REV and TAT genes have been removed from the packaging plasmid. The transferred genome expression is controlled by a strong constitutive promoter (i.e. the cytomegalovirus immediate early promoter or CMV), by the replacement of the 5' HIV U3 region. **(C)** Further improvements on Lentiviral vector biosafety and performance are highlighted in this figure by stars. These include the addition of the central polypurine flap (cPPT), the Woodchuck Post-transcriptional Regulatory Element (WPRE), and the removal of enhancers from the 3' HIV U3, generating self-inactivating lentivectors (Δ Enh). Non-Integrating Lentiviral Vector (NILVs) can be engineered by introducing point mutations and deletions within the integrase attachment sites in the 5' U3 region of the transfer vector plasmid, or in the integrase ORF. Figure adapted from Liechtenstein et al. *Cancers* 2013¹⁰⁵.

Improvement of the transfer vectors

A Lentiviral transfer vector consists of an expression cassette and the HIV *cis*-acting elements for packaging, reverse transcription, and integration (see **Figure 9**). Self-inactivating transfer vectors have been engineered by deleting the enhancer/promoter region in the U3 region of the 3' LTR. This further minimizes the risk for insertional mutagenesis and the risk of formation of replication-competent lentivirus. It also reduces the potential for promoter interference^{107–109}. The use of tissue-specific promoters in the transfer vectors has enabled high levels of expression in specific cells. Several studies have reported the successful use of cell-specific promoters in retroviral vectors. The enolase promoter, SYN promoter, synapsin 1 promoter, CD44 promoter, glial fibrillary acidic protein promoter, and vimentin promoter have been shown to express well in neuronal and glial cells (see review by Escors and Breckpot 2010¹⁰³). The Alpha-fetoprotein promoter has been used to specifically drive the expression of a suicide transgene in hepatocarcinoma cells. Also, a PSA promoter-based lentiviral vector, delivering the diphtheria toxin A gene, has been shown to successfully eradicate prostate cancer cells in both tissue culture and a mouse tumor model. For “gene immunotherapy”, an HLA-DRA promoter-containing lentiviral vector has been used to specifically drive the expression of human MHC class II⁺ on the surface of dendritic cells in the non-obese diabetic/SCID mouse engraftment model (see review by Escors and Breckpot 2010¹⁰³).

Pseudotyping of Lentiviral vectors

The tropism of packaged Lentiviral vectors is determined by the viral envelope glycoproteins that are required for cell entry. Lentiviral particles can be formed with different envelope proteins in their membranes, by packaging them in cells expressing the envelope protein of interest. This process is called pseudotyping. Lentiviral particles are usually “pseudotyped” with the envelope protein of Vesicular Stomatitis Virus (VSV-G). VSV-G is a glycoprotein which broadly interacts with cellular receptors and thus can be used in many different cell types^{110,111}. Furthermore, the VSV-G pseudotyped particles are highly stable, making it possible to concentrate the virus to obtain higher titers¹¹². In “gene immunotherapy”, a mutated envelope protein of Sindbis virus has been used to pseudotype Lentiviral vector particles, allowing them to specifically bind to DC-SIGN of dendritic cells¹¹³. Lentiviral vector pseudotyping has been extensively described by Cronin et al. 2005¹¹¹.

The use of Lentiviral vectors in immunotherapy

A main goal of cancer immunotherapy is to break tolerance to tumor-associated antigens since they are mostly either self or quasi-self antigens. To achieve this, mature immunogenic dendritic cells have to present tumor-derived antigens to specific CD4 or CD8 T cells in order to drive an efficient antitumor response. Lentiviral vectors have been selected for transducing monocyte-derived DCs, which would present tumor-derived peptides to both CD4 and CD8 T cells¹⁰³. Also, Lentiviral vectors have been shown to elicit cytotoxic T-

lymphocytes responses against transgene-encoded proteins^{114–116}. Recently, a Lentiviral vector, which expresses a chimeric antigen receptor with specificity for the B-cell antigen CD19, coupled with CD137 and CD3-zeta signaling domains, has been used to transduce autologous T-cells in order to generate chimeric antigen receptor–modified T cells (CAR-T)¹¹⁷. This strategy has been successfully used to treat Chronic Lymphocytic leukemia (CLL) and B-cell malignancies, respectively^{117,118}. A review by Johnson and June offers a discussion on the further development of CAR-T cells for cancer immunotherapy¹¹⁹.

Antisense RNA and gene therapy

A principle of antisense-mediated inhibition of gene expression is that nucleic acid sequences (antisense) complementary to its targeted mRNA (sense) can hybridize. This leads to the disruption of normal RNA processing and interference with nuclear export, RNA stability, or translation¹²⁰. Exogenously added synthetic antisense oligonucleotides have been used to interfere with the viral replication of Rous Sarcoma Virus (RSV), as well as virus transformation in tissue culture^{121,122}. A similar approach was used many years ago to inhibit HIV-1 replication¹²³. The antiviral property of antisense oligonucleotides was achieved by inhibiting viral replication and/or viral packaging via the formation of a sense-antisense duplex between viral genomic RNA and complimentary antisense oligonucleotides.

These early experiments led to the idea that a Lentiviral vector expressing antisense RNA against HIV transcripts could be delivered into CD4+ T cells or CD34+ stem cells of patients to potentially achieve efficient HIV inhibition. The genetically engineered cells would stably express antisense RNA and become resistant to viral infection. This type of strategy has come to be known as 'intracellular immunization'¹²⁴. However, an early attempt at HIV gene therapy using CD4+ T cells constitutively expressing antisense RNA to HIV-1 Tat was not successful¹²⁵. In this strategy, CD4+ T cells were transduced with a version of antisense RNA vector consisted of HIV 5' and 3' LTRs, a Neo^R gene ORF, and an HIV-1 TAT antisense RNA driven by a parvovirus P6 promoter. The *de novo* HIV infection assay of transduced cells showed no difference in cell viability between engineered CD4+ T cell clones and control clones. Furthermore, it was shown that even very high amounts of TAT antisense RNA molecules failed to interfere with HIV Tat expression¹²⁵.

Later, Lu et al. described a new vector system for antisense RNA-mediated inhibition of HIV replication. Transduction of HIV-1 natural host cells, CD4+ lymphocyte with an HIV-based vector, which contains an antisense *env* RNA sequence, was demonstrated to result in very efficient inhibition of HIV replication in transduced cells¹²⁶. However, the initial report did not include experiments to address the potential mechanism for the potent antisense RNA-mediated inhibition. Ward et al., later reported that the efficient inhibition of HIV-1 replication by HIV-1 derived Lentiviral antisense *env* RNA vectors requires trafficking through the HIV-1 Rev/RRE pathway¹²⁷. As described above, the HIV-

1 Rev/RRE pathway is best known for its ability to surmount nuclear retention of HIV-1 unspliced and incompletely-spliced transcripts and mediate their nuclear export^{36,38,39}.

Specifically, the study showed that the antisense HIV-1 *env* RNA has to traffic through the Rev/RRE pathway in order to achieve efficient inhibition (see **Figure 10**). The inhibition was significantly reduced when Rev was replaced by RevM10-Tap, a mutant form of Rev that directs RNA trafficking to the NXF/NXT pathway, used by many cellular mRNAs. The mechanistic study of the Rev/RRE-mediated antisense inhibition reported no involvement of nuclear retention of HIV-1 target mRNA. It was demonstrated that the targeted mRNA was exported to the cytoplasm and remained associated with polyribosome complexes, as shown in **Figure 11**. However, HIV-1 capsid protein (p24) levels were greatly reduced. Furthermore, the studies showed that the reduction of p24 protein levels was not due to general protein degradation since the production of HIV-1 Nef protein, translated from a completely spliced HIV-1 transcript containing no target sequence, was unaffected (see **Figure 12**).

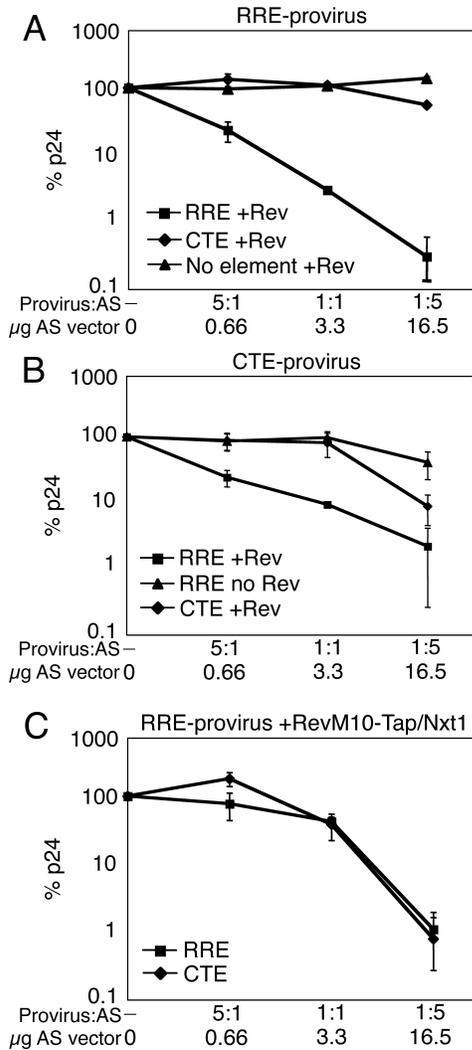


Figure 10. Inhibition of particle production from HIV-1 proviruses by antisense RNA

(A) Inhibition of particle production from an RRE-driven provirus. 293T cells were transfected with a pNL4-3 provirus lacking a functional *rev* gene. Increasing amounts of No element, CTE-driven or RRE-driven antisense constructs (AS) were added with a plasmid making Rev. After 48 hours, p24 expression was assayed and plotted as the percentage of p24 expression from the provirus alone. The molar ratio of provirus to antisense plasmid and microgram amount of antisense plasmid (AS) is indicated on the X-axis. **(B)** Inhibition of particle production from a Rev-independent CTE-driven provirus. 293T cells were transfected with the pNL4-3 provirus lacking a functional RRE and *rev* gene and the MPMV CTE cloned into *nef*. Increasing amounts of RRE- or CTE-driven antisense constructs (AS) were added with or without Rev co-expression. **(C)** Inhibition of particle production from an RRE-driven provirus in the presence of RevM10-Tap and Nxt1. 293T cells were transfected with the pNL4-3 proviral clone lacking a functional *rev* gene, pCMV-RevM10-Tap, pCMV-Nxt1 and increasing amounts of RRE or CTE-driven antisense construct (AS). RevM10-Tap exports the RNA using the Tap/NXF1 pathway. Figure from Ward et al. J. Virol. 2009¹²⁷.

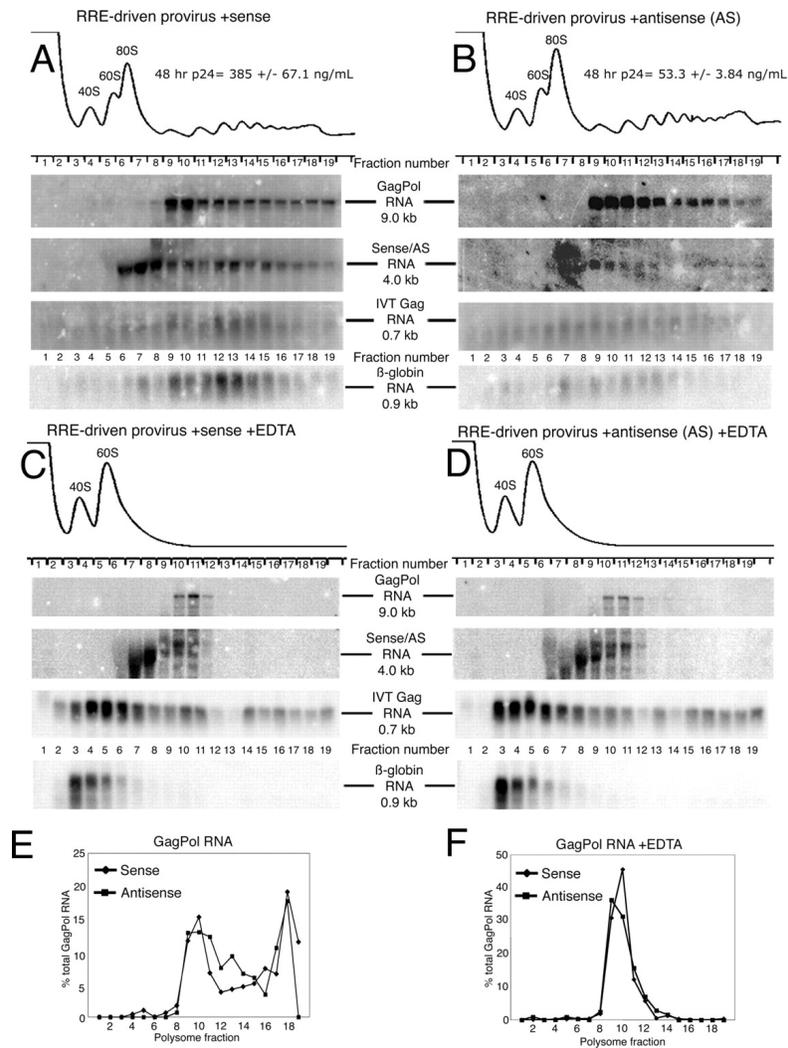


Figure 11. Target RNAs associate with polyribosome

(A and B) A total of 10^7 293T cells were transfected with 5 μ g of the pNL4-3 proviral clone lacking a functional *rev* gene, 500 ng of pCMV-Rev, and 660 ng of RRE-driven sense (A) or antisense (B) vector at a 5:1 molar ratio of provirus and vector. Also, 5 μ g of CMV- β -globin was cotransfected as a control. At 48 h posttransfection, cytoplasmic extracts were prepared and separated by centrifugation through a sucrose gradient. Gradients were fractionated while monitoring the UV absorbance at 254 nm. RNA was purified from fractions after the addition of in vitro-transcribed Gag (IVT Gag) RNA for a recovery control. RNA from fractions were analyzed by Northern blotting and probed with specific radiolabeled DNA probes to GagPol, sense/antisense, β -globin, and IVT Gag RNA. Blots were analyzed by using a Molecular Dynamics PhosphorImager and ImageQuant software. (C and D) Polyribosome analysis of GagPol and sense/antisense RNA after EDTA treatment and sucrose gradient centrifugation. Transfected cells were harvested and processed as described in panels A and B except that prior to sucrose gradient centrifugation, cytoplasmic extracts were treated with 15 mM EDTA. (E) Quantitation of GagPol RNA localization in a sucrose gradient in the presence of sense/antisense RNA. GagPol RNA levels were normalized to IVT Gag and plotted as a percentage of the total GagPol RNA present in the gradient. (F) Quantitation of GagPol RNA localization in a sucrose gradient in the presence of sense/antisense RNA after treatment with 15 mM EDTA. Analysis was carried out as in panel E. AS, antisense. Figure from Ward et al. J. Virol. 2009¹²⁷.

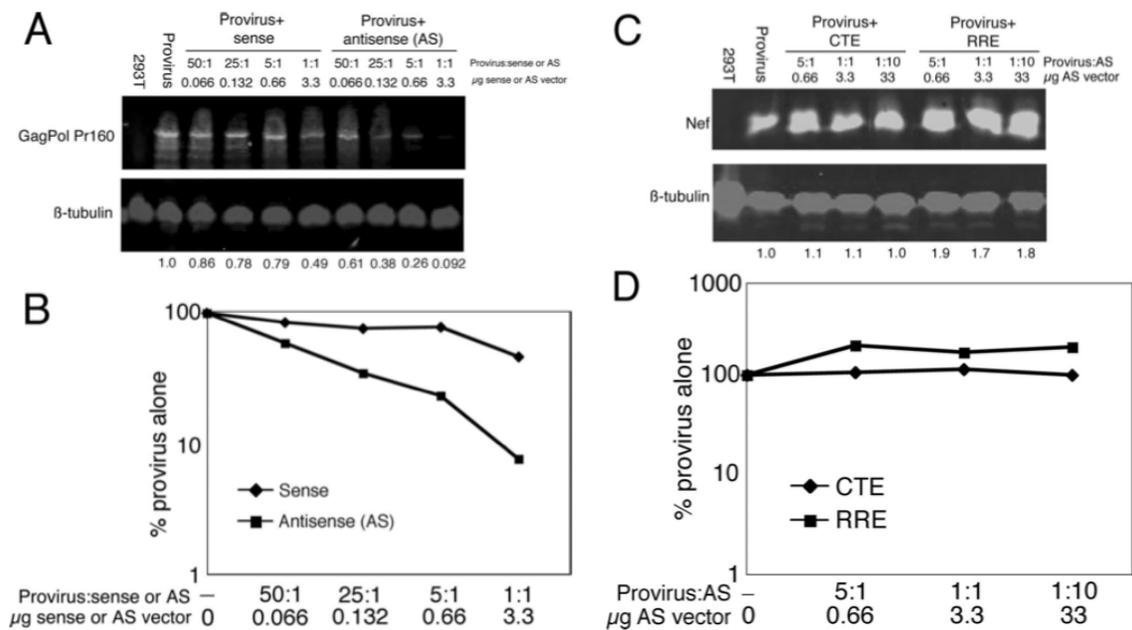


Figure 12. GagPol Pr160 and Nef production in the presence of sense or antisense (AS) vectors

(A and B) Pr160 production in the presence of RRE-driven sense or antisense (AS) vector. 293T cells transfected with a proviral clone and RRE-driven sense or antisense constructs. The proviral clone contains a nonfunctional myristylation signal, as well as GagPol frameshift and protease mutations and makes non-secreted Pr160. Cells were harvested at 48 hours and extracts were Western blotted with a monoclonal antibody to p24, and a commercial polyclonal antibody to β-tubulin (Abcam). GagPol Pr160 expression levels relative to provirus alone is indicated below each lane. **(C and D)** HIV-1 Nef production in the presence of CTE- or RRE-driven antisense (AS) constructs. 293T cells were transfected with a proviral clone and increasing amounts of the CTE- or RRE-driven antisense (AS) constructs. Cells were harvested, blotted and analyzed as described in **(A)** except that an HIV-1 Nef monoclonal antibody (SN20) was used in place of the p24 monoclonal antibody. Figure from Ward et al. J. Virol. 2009¹²⁷.

This antisense-mediated inhibition of HIV-1 replication demonstrated a novel mechanism in which the HIV-1 Rev/RRE pathway was essential for antisense RNA-mediated reduction of HIV-1 capsid protein (p24). The inhibition did not inhibit nuclear export or association of the targeted mRNA with the translation machinery. This led the authors of the study to suggest that perhaps there was an involvement of RNAi-mediated translational repression mechanisms in the inhibition. Thus a reduction in the rate of translational

initiation, a decrease in the rate of translational elongation, or even a rapid degradation of produced GagPol protein could be the cause of the reduction of GagPol protein levels (p24 protein levels) that was seen in these studies.

Recently, a report from a clinical trial using the antisense vector to modify autologous CD4 T cells in 17 individuals living with HIV has been published. This trial was registered at www.clinicaltrials.gov as number NCT00295477. The Lentivirus vector used (VRX496-T or Lexgenleucel-T™) expresses a 937-nt long antisense sequence complementary to HIV-1 *env*¹²⁸. The antisense RNA sequence is similar to that used by Lu et al¹²⁶. Antiviral effects were studied during analytic treatment interruption (ATI) in a subset of 13 patients. VRX496-T was reported to reduce viral load in 6 of 8 subjects (P = 0.08). VRX496-T modified cells were also identified in a population of intraepithelial lymphocytes, which were isolated from rectal biopsies at baseline and post-infusion before and after ATI¹²⁸. Interestingly, VRX496-T engrafted cells were shown to have a half-life of approximately five weeks in the blood. The success of this clinical trial has shown a potentially promising future for the use of a long antisense RNA in gene therapy using Lentiviral vectors.

RNA interference (RNAi)

RNA interference (RNAi) mechanisms were first discovered in *C. elegans*. Homologous pathways have now also been found in *D. melanogaster*, mammals, plants, and protozoa^{129–131}. The pathways are also exploited by DNA viruses¹³². Several lines of evidence suggest the involvement of RNAi in a number of cellular processes, such as development, differentiation, growth, homeostasis, stress responses, apoptosis, and immune activations (see review by Skalsky and Cullen¹³²). RNA interference (RNAi) is mediated by ~19-24 nt-long double-stranded RNA molecules. RNAi mediators can be categorized into either siRNAs or miRNAs. The antisense strands of the siRNAs anneal with perfect complementarity to their target sequence, while microRNA only forms partial complementary base-pairing (see review by Kim et al.¹³³).

The RNAi pathway begins with the generation of small RNA (**Figure 13**). In brief, ~80 nucleotide (or longer) RNA polymerase II-transcribed, hairpin-structured RNAs (known as primary-microRNAs or pri-miRNAs) are 5' capped and polyadenylated, like regular mRNAs. Later, DiGeorge syndrome Critical Region 8 (DGCR8) proteins bind to the pri-miRNAs and recruit the RNase III enzyme, Drosha, to form a ~650 kDa multiprotein complex named the Microprocessor (see review by Czech et al.¹³⁴). DGCR8 orients the catalytic RNase III domain of Drosha to release hairpins from pri-miRNAs by cleaving RNA about eleven nucleotides away from the ssRNA-dsRNA junction (**Figure 14**). The resulting double-stranded RNA hairpins (known as precursor-miRNAs or pre-miRNAs), contain a unique two-nucleotide overhang at the 3' end (see

review by Kim et al.¹³³). Interestingly, debranching of lariat-structures can also give rise to pre-miRNAs, named mirtrons. Many lines of evidence also suggest that pre-miRNAs can be generated from alternative folding of transfer RNAs (tRNAs) or small nucleolar RNAs (snoRNAs), or tRNAse Z cleavage of pri-miRNAs containing tRNA-like structures linked to pre-miRNAs stem-loops), see review by Skalsky and Cullen¹³².

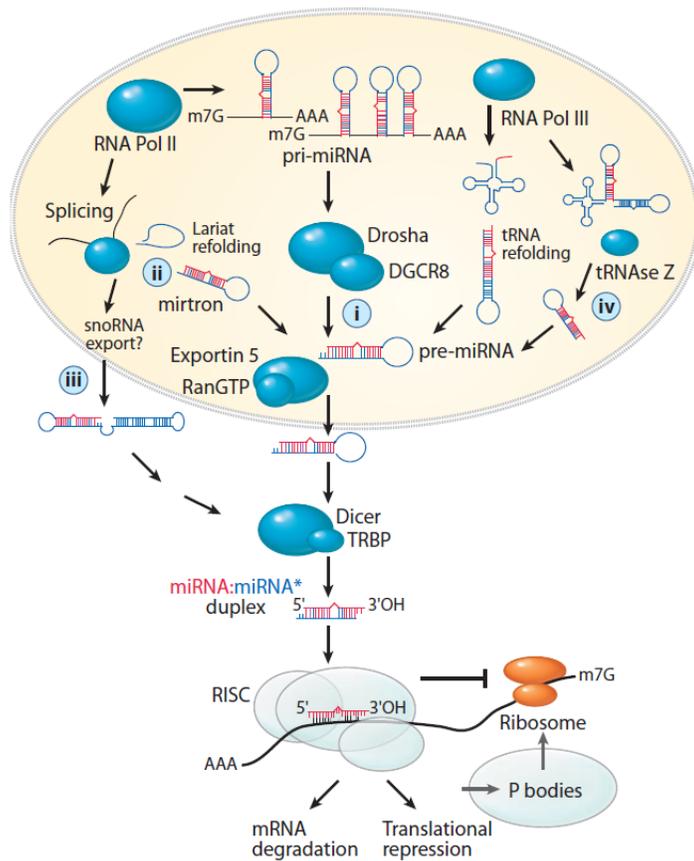


Figure 13. RNAi Pathway

Precursor miRNAs (pre-miRNAs) can be generated (i) by Drosha/DGCR8 cleavage of long primary miRNAs (pri-miRNAs) or independent of Drosha/DGCR8, (ii) following debranching of lariat-structures known as mirtrons, (iii) through alternative folding of transfer RNAs (tRNAs) or small nucleolar RNAs (snoRNAs), or (iv) by tRNAse Z cleavage of pri-miRNAs containing tRNA-like structures linked to pre-miRNAs stem-loops. Once exported to the cytoplasm, pre-miRNAs are cleaved by Dicer to generate a miRNA duplex, one strand of which is incorporated into the RNA-induced silencing complex (RISC) to target messenger RNAs (mRNAs). Figure from Skalsky RL and Cullen BR *Annu Rev Microbiol* 2010¹³².

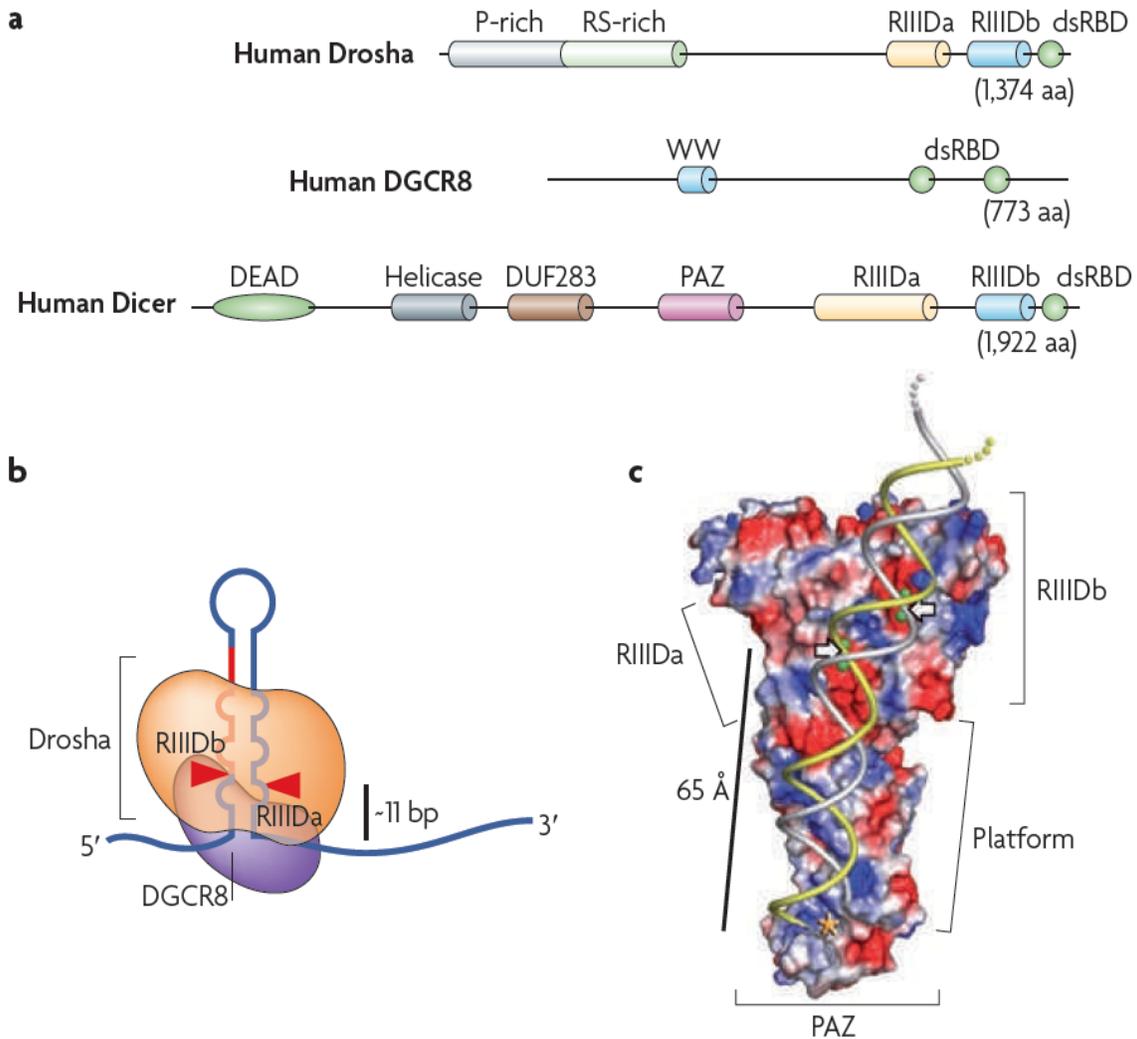


Figure 14. RNase III proteins and their mechanisms

(a) Drosha is a 130-160 kDa-nuclear RNase III enzyme. It contains two tandem RNase III domain (RIIIDa and b), as well as a double-stranded RNA-binding domain (dsRBD). DiGeorge syndrome critical region gene 8 (DGCR8) (or Pasha in *Drosophila melanogaster* and *Caenorhabditis elegans*) is ~120 kDa and found localized in the nucleoplasm and the nucleolus. DGCR8 (Pasha) contains two dsRBDs that recognize the ssRNA–dsRNA junction of a substrate. The WW domain of DGCR8 is required for oligomerization and pri-miRNA processing. Dicer is a ~200-kDa cytoplasmic RNase III enzyme. The middle region of Dicer contains a PAZ domain, which binds to the 3' protruding end of RNAs. The DEAD-box RNA helicase domain is not necessary for Dicer activity *in vitro*, and its role remains unknown. Two RNase III domains of Drosha and Dicer interact with each other to make an intramolecular dimer. The first RIIID cleaves the 3' strand of dsRNA, whereas the second RIIID cleaves the 5' strand, generating a 2-nt 3' overhang. (b) Drosha interacts with a DGCR8 (or Pasha), through its middle region to form the Microprocessor. The Microprocessor complex processes primary microRNA into premature microRNA, which are exported for further processing in cytoplasm by Dicer. (c) A structure of *G. intestinalis* Dicer containing dsRNA docked to the PAZ domain and the RNA stem interacts with the flat, positively charged extension to reach the catalytic centre. The distance between the PAZ domain and the RNase III catalytic site (65 Å) approximately matches the length of the product of its product, a 25-nt miRNA. Figure from Kim et al. Nature Reviews 2009¹³³.

The double-stranded RNA hairpins (or pre-miRNAs) are then exported into cytoplasm by the export receptor Exportin-5. Exportin-5 recognizes their short 3' overhang and mediates nuclear export through the Exportin-5/RanGTP pathway¹³⁵. In the cytoplasm, the terminal loops of the small RNA hairpins are removed by Dicer, which is an RNase III enzyme, acting in association with the Human Immunodeficiency Virus (HIV) trans-activating response (TAR) RNA binding protein (TRBP). Recently, Park et al. has demonstrated that human Dicer anchors to both the 3'- and 5'-ends of the dsRNA hairpin, and the cleavage site is determined by the distance, approximately 22 nucleotides, from the 5' end. The cleavage requires a 5'-terminal phosphate group, and it is conserved in *Drosophila*¹³⁶. This generates ~22-bp short RNA duplex intermediates bearing 2-nt 3' overhangs at each end; see **Figure 13 and 14**. One strand of the duplex (shown as miRNA) is incorporated into the RNA induced silencing complex (RISC) complex. The remaining passenger strand (shown as miRNA*) is degraded. Strand selection depends on the degree of base pairing at the duplex 5' ends—the strand less stably base paired at its 5' end is preferentially incorporated into RISC, see review by Skalsky RL and Cullen BR, 2010¹³². Interestingly, the HIV-1 TAR RNA binding protein (TRBP) has been implicated in the recruitment of Ago2 to the small 22-nt RNAi bound by Dicer¹³⁷.

The RNA-induced silencing complex (RISC) is a ribonucleoprotein complex consisting of two core protein components, namely Argonaute (AGO) proteins and the Trinucleotide repeat-containing gene 6A protein (TNRC6 or GW182) protein (**Figure 15**). In humans, one of four different Argonaute proteins (Ago1–4) is present in the RISC complex. Only Ago2 exhibits endonuclease activity and has the ability to cleave bound target mRNAs^{138,139}. AGO proteins escort small RNAi molecules to their mRNA targets. TNRC6A (GW182) protein, interacting with AGO proteins via their GW repeats, subsequently brings about RNAi-mediated posttranscriptional regulation¹⁴⁰.

Mature miRNAs have been shown to bind to complementary sequence in the 3'UTRs of their target transcripts, through 5'nucleotides 2–7 of the mature miRNAs, named the seed sequence. Imperfect complementarity between miRNAs and their targets mostly results in translational repression and/or deadenylation and subsequent degradation of mRNA targets, see review by Fabian et al. 2010¹⁴⁰. On the other hand, a perfect complementarity between seed sequence of siRNAs and their targets often results in mRNA cleavage and destabilization (**Figure 16** and **Figure 17**).

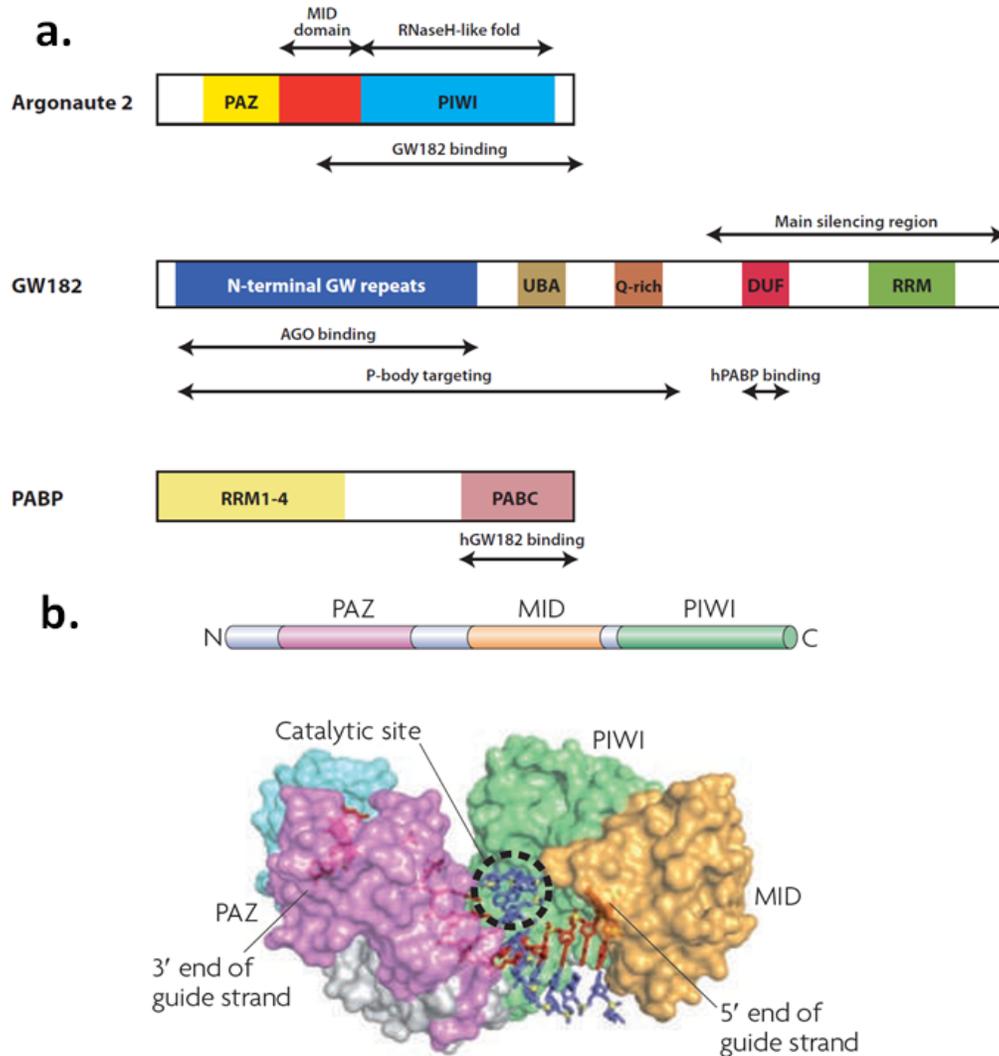


Figure 15. Schematic diagram of human Argonaute 2 (AGO2), GW182, and poly(A)-binding protein (PABP)

(a) Argonaute proteins: The Argonaute (Ago) family can be classified into two subclades: the Ago subfamily and the Piwi subfamily. The Ago proteins are expressed ubiquitously, interact with microRNAs (miRNAs) or small interfering RNAs (siRNAs), and function as post-transcriptional regulators. The Piwi proteins are abundantly expressed in germ cells and function in transposon silencing, together with Piwi-interacting RNAs (piRNAs). There are four human Argonaute proteins, only Ago2 is enzymatically functioning in RNAi. Ago2 contains an enzymatically competent RNaseH-like PIWI domain, which endonucleolytically cleaves perfectly complementary RNA targets. The PAZ domain serves as a docking site for the 3' end of small RNA, whereas the MID domain anchors the 5' terminal nucleotide. GW182 protein: There are three human GW182 paralogs (TNRC6A, -B, and -C), whereas *Drosophila* contains only one GW182 protein (dGW182 or Gawky) sharing a similar domain organization. *C. elegans* contains two proteins, AIN-1 and -2, which differ substantially from GW182s but perform analogous functions. The N-terminal region of GW182 contains glycine-tryptophan (GW) repeats, which interacts with AGO proteins. The GW-rich, ubiquitin-associated (UBA), and glutamine-rich (Q-rich) domains are responsible for targeting GW182 proteins to P bodies. The C-terminal region contains DUF (domain of unknown function) motifs and RNA recognition motifs (RRMs). The RRM is a major effector domain, mediating translational repression and deadenylation of mRNA. The Poly(A)-binding protein

(PABP): The protein contains four RRM domains and a conserved C-terminal domain, PABC. Mammalian PABP binds directly to the silencing region of human GW182 proteins via PABC. **(b)** The structure of *Thermus thermophilus* Ago binding to a guide strand and target strand duplex. The PIWI domain has a structure that is similar to RNase H, which cuts the RNA strand of an RNA–DNA hybrid. The PIWI domain of some Ago proteins can cleave the target RNA bound to small RNA: this is called slicer activity. Only AGO2 has slicer activity, whereas in *Drosophila melanogaster* all Ago and Piwi proteins possess slicer activity. Apart from the endonucleolytic cleavage that is mediated by the PIWI domain, the Ago proteins can induce translational repression and exonucleolytic mRNA decay through interaction with other protein factors. Figure from Kim et al. Nature Reviews 2009¹³³.

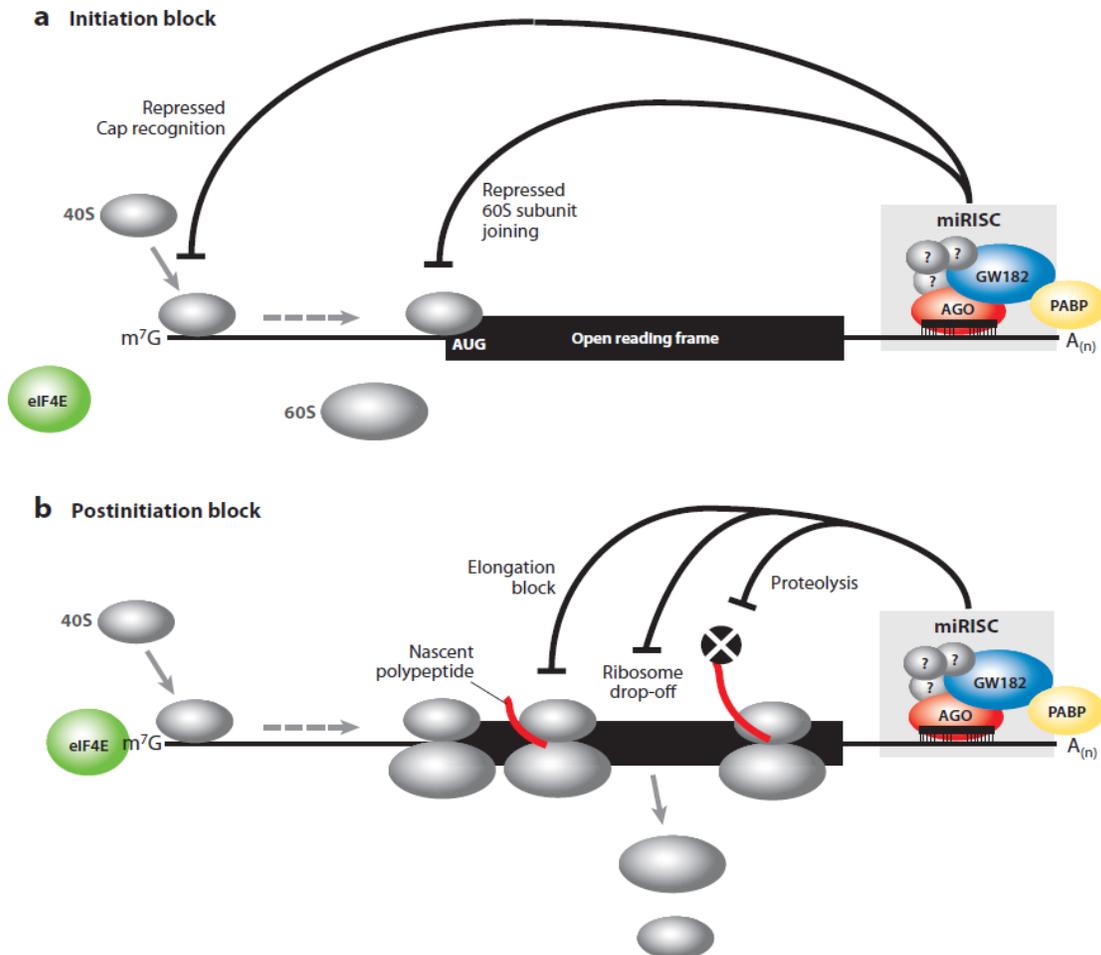


Figure 16. Schematic representation of potential mechanisms for miRNA-mediated translational repression

(a) Initiation block: The RISC inhibits translation initiation by interfering with eIF4F-cap recognition and 40S small ribosomal subunit recruitment or by antagonizing 60S subunit joining and preventing 80S ribosomal complex formation. The reported interaction of the GW182 protein with the poly(A)-binding protein (PABP) might interfere with the closed-loop formation mediated by the eIF4G-PABP interaction and thus contribute to the repression of translation initiation. **(b)** Post-Initiation block: The RISC complex is hypothesized to inhibit ribosome elongation, inducing ribosome drop-off, or facilitating proteolysis of nascent polypeptides. Figure from Fabian et al. Aun Rev Biochem 2010¹⁴⁰.

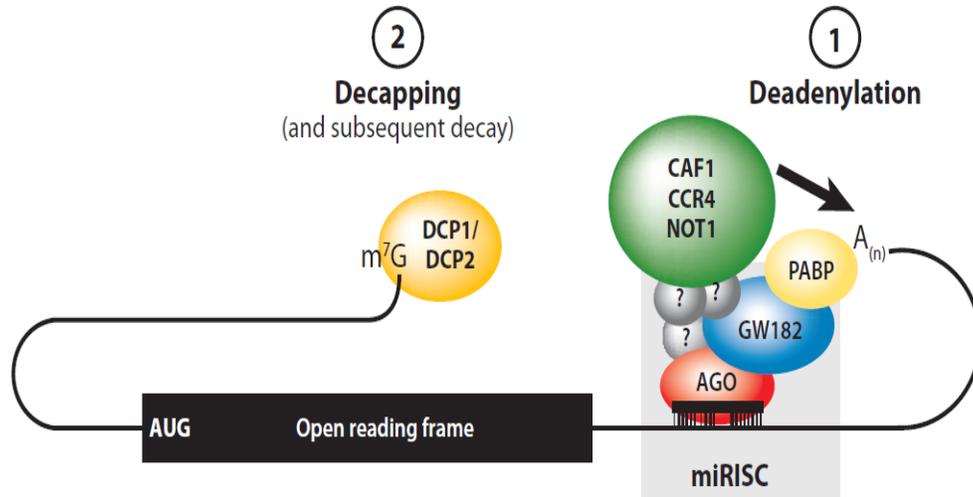


Figure 17. Schematic representation of potential mechanisms for miRNA-mediated mRNA decay

(1) The RISC complex interacts with the CCR4-NOT1 deadenylase complex to facilitate deadenylation of the poly(A) tail. RISC-mediated deadenylation requires direct interaction between GW182 protein and the poly(A)-binding protein (PABP). (2) Following deadenylation, the 5-terminal cap (m⁷G) is removed by the decapping DCP1-DCP2 complex. Abbreviations: AGO, Argonaute; CAF1, CCR4-associated factor; CCR4, carbon catabolite repression 4 protein; NOT1, negative on TATA-less. Figure from Fabian et al. *Annu Rev Biochem* 2010¹⁴⁰.

ADAR editing

The human genome contains three different adenosine deaminase acting on dsRNA (ADAR) genes: ADAR1 (encoding 2 ADAR1 isoforms), ADAR2 and ADAR3, (see review by Bass BL 2002¹⁴¹ and Nishikura K. 2010¹⁴²). The ADAR family catalyzes Adenosine-to-Inosine (A-to-I) editing in either partially or perfectly double-stranded RNAs. ADARs are expressed in most metazoan species¹⁴³. The human enzymes share a conserved C-terminal catalytic deaminase domain and a N-terminal double-stranded RNA-binding domain¹⁴⁴, see top of **Figure 18**. ADAR1 has three double-stranded RNA binding domains. Alternative promoters in the ADAR1 genes result in two forms of the enzyme. A

short 110-kDa ADAR1 (ADAR1S) contains 2 N-terminal Z-DNA binding motifs, 3 double-stranded RNA binding domains (dsRBD), and a C-terminal catalytic deaminase domain. A longer 150-kDa ADAR1 (ADAR1L) contains an additional Z-DNA-binding motif and a nuclear export signal (NES). ADAR1S (p110) is constitutively transcribed from two promoters, whereas ADAR1L (p150) is interferon-inducible in response to cellular stress or viral infection. ADAR1L (p150) localize primarily in cytoplasm. ADAR1S is found primarily in the nucleus. ADAR2 contains two double-stranded RNA-binding domains and a catalytic adenosine deaminase domain. ADAR2R, an alternatively spliced variant of ADAR2, contains additional 49 amino acid residues in the N-terminal, resulting from an insertion of a coding, in frame Alu exon, (see review by Nishikura K. 2010¹⁴²). This form is not shown in **Figure 18**. However, **Figure 18** also shows the single ADAR isoform (most similar to ADAR2) that is found in *Drosophila*, as well as two *C. elegans* ADAR proteins.

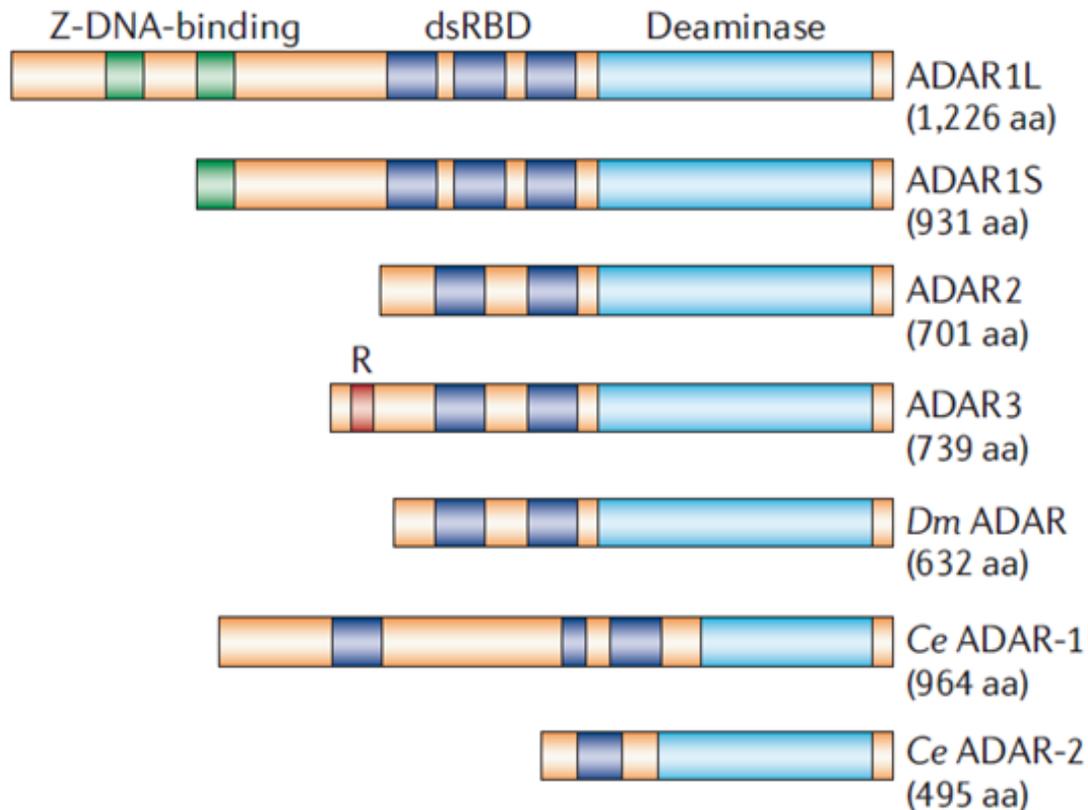


Figure 18. ADAR family proteins

Three human ADAR (adenosine deaminase acting on RNA)-family members (ADAR1–3), *Drosophila melanogaster* (Dm) ADAR and two *Caenorhabditis elegans* (Ce) proteins, ADAR-1 and ADAR-2, share common functional domains: 2 or 3 double-stranded RNA Binding Domain (dsRBD) and a catalytic deaminase domain. Certain structural features, such as Z-DNA-binding domains and the Arg-rich (R) domain, are unique to particular ADAR members. Binding of ADAR to doublestranded RNA (dsRNA) substrates is mediated through dsRBDs, whereas Z-DNA-binding domains might increase the affinity of ADAR1L specifically for short dsRNAs such as siRNAs. Binding of the R domain to single-stranded RNAs has been reported, but its biological significance is currently unknown. Two ADAR1 translation products, the isoforms ADAR1L and ADAR1S, result from transcription from different promoters followed by alternative splicing. Figure from Nishikura K. *Nat Rev Mol Cell Biol* 2006¹⁴⁵.

ADAR1 and 2 are expressed in most tissues and can homodimerize. ADAR3, however, is expressed primarily in post-mitotic cells in certain parts of the central nervous system. Enzymatic function of ADAR3 has not been demonstrated. However, an *in vitro* study demonstrated that ADAR3 can

decrease the efficiency of ADAR1 and 2 editing through the use of both the dsRNA binding domain and the ss-RNA binding domain (R-domain) for sequestering substrates of ADAR1 and 2. ADAR3 may thus specifically control editing mediated by ADAR1 and 2 in the central nervous system^{142,143}. The product of ADAR deamination of adenosine residues (A) in the RNA is inosine (I); see **Figure 19**. Thus ADARs introduce changes into the primary sequence information in RNAs, as well as in the structure due to I-U base-pair mismatch¹⁴¹; see **Figure 20**. When the transcripts are translated, inosine (I) residues are read as if they are guanosine (G), leading to potential amino acid changes.

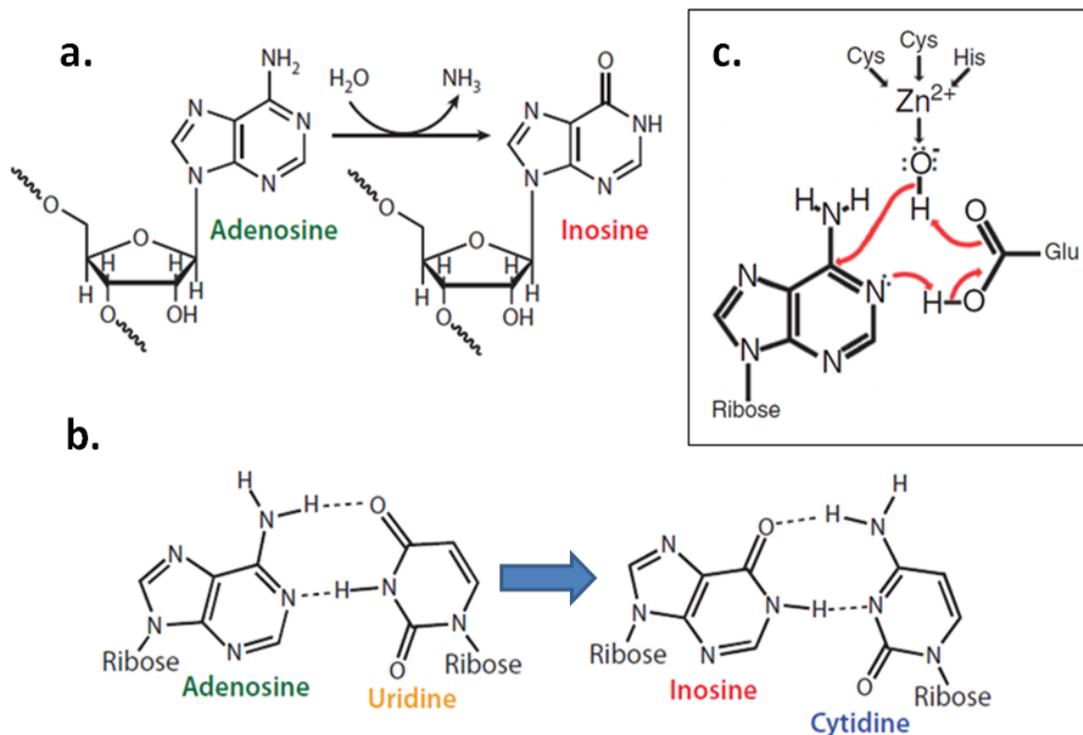


Figure 19. Deamination of Adenosine (A) to Inosine (I) by ADAR

(a) A hydrolytic deamination reaction converts adenosine to inosine. (b) Adenosine base pairs with uridine, whereas inosine base pairs, as if it were guanosine, in a Watson-Crick-bonding configuration with cytosine. (c) The hypothesized mechanism underlying catalysis by ADAR, based on an ADAR crystal. Figure from Wulff et al. Wiley Interdiscip Rev RNA. 2010¹⁴³.

		Second letter				
		U	C	A	G	
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA } Stop UAG }	UGU } Cys UGC } UGA } Stop UGG } Trp	Third letter U C A G
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	
	A	AUU } AUC } Ile AUA } AUG } Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	

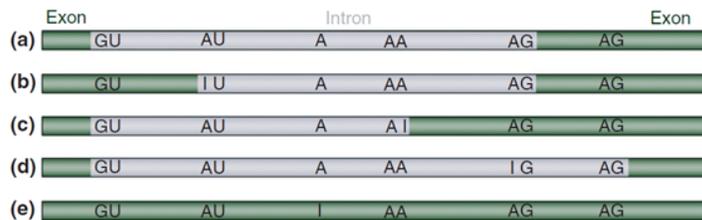


Figure 20. Recoding events caused by A-to-I Editing on RNA

A summary of recoding events that can be caused by A-to-I editing, including the creation of a Met start codon and the destruction of stop codons (Top panel). A summary of splice site modifications possible upon editing of canonical splice site and branch point motifs. The original splice pattern (a) can be modified by the creation of a new 5 splice site (b), by the creation of a new 3 splice site (c), by the destruction of a 3 splice site (d), or by the destruction of the branch point (e) (Bottom panel). Figure from Wulff et al. Wiley Interdiscip Rev RNA. 2010¹⁴³.

ADAR editing may potentially be involved in creating diversity within the HIV-1 genome, for example if HIV-1 natural antisense RNAs are expressed from cellular promoters downstream of an integrated provirus. Furthermore, Roger H. Miller originally proposed that the HIV provirus contains an additional open reading frame (ORF) region that is complementary to the envelope gene (*env*). The proposed antisense ORF is capable of encoding a ~19-20 kDa protein, predicted to be highly hydrophobic⁸². As discussed above, further investigation identified a promoter for HIV-1 antisense RNA located in 3' LTR or U3-R region of the provirus^{83,84}. HIV-1 antisense transcripts are constitutively expressed in

various cell lines chronically infected with HIV-1 and, interestingly, antibodies to its predicted encoded protein have been detected in HIV-1 patients⁸⁵.

Several lines of evidence have suggested that ADAR editing in the HIV-1 genome enhances HIV-1 replication and infectivity¹⁴⁶⁻¹⁴⁹. It is plausible that base-pairing between HIV-1 sense and naturally occurring antisense RNAs results in dsRNAs, a substrate for ADAR-mediated RNA editing. The presence of I:U mismatch could disrupt the duplex and release HIV mRNAs carrying new mutations. However, A-to-I hyperedited dsRNAs have been demonstrated, by the Carmichael group, to be generally retained in the nucleus by nuclear matrix¹⁵⁰. However, the study showed that Rev/RRE could overcome this retention in an oocyte model system. However whether ADAR enzymes can lead to new mutations, and, subsequently new strains of HIV virus remains unproven.

As described above, we hypothesize that small regulatory RNAs could be generated from the double-stranded RNA (dsRNA) of sense:antisense RNAs. During the inhibition, ADAR enzyme(s) could modify adenosines of the dsRNA and turn them into inosines. Due to the HIV-1 Rev/RRE pathway, the edited dsRNAs may escape nuclear retention and the A-to-I edited dsRNAs would be exported via the HIV-1 Rev/RRE pathway, thereby leading to the export of edited RNAs into the cytoplasm. Cytoplasmic type III RNase enzyme Dicer, could then process the dsRNAs into several molecules of miRNA since the dsRNAs have partially lost their sequence identity to HIV-1 *env* gene due to ADAR editing. Though analysis of ADAR editing is an objective of this investigation, the possibility of dsRNA turnover or degradation should also be taken into account in

this analysis. A-to-I hyperedited long dsRNAs can be rapidly degraded by Tudor staphylococcal nuclease (Tudor-SN)^{142,151}, see **Figure 21**. This nuclease contains five staphylococcal/micrococcal nuclease domains and a tudor domain¹⁵². Tudor-staphylococcal nuclease (Tudor-SN) has been reported as a component of the RISC in *Caenorhabditis elegans*, *Drosophila* and mammals.

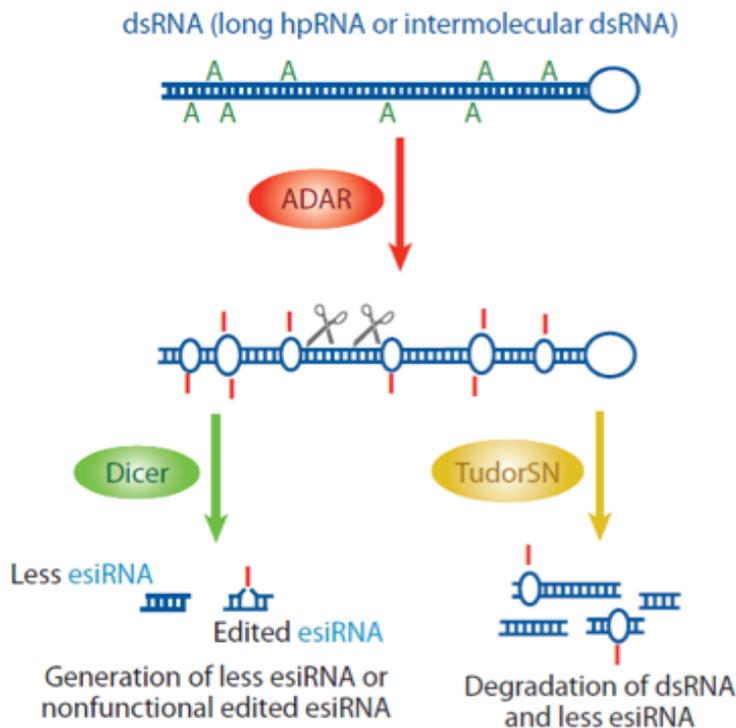


Figure 21. Interaction between RNA editing and RNA-interference pathways

The presence of numerous inosine-uridine mismatches and the alteration of the double-stranded RNA (dsRNA) structure by ADAR (adenosine deaminase acting on RNA) leads to the generation of fewer siRNAs due Dicer resistance of edited dsRNAs. Alternatively, extensively edited dsRNAs may be rapidly degraded by Tudor staphylococcal nuclease (Tudor-SN), resulting in the generation of fewer siRNAs. Figure from Nishikura K. Annu Rev Biochem 2010¹⁴².

Project Rationale

Ward et al.¹²⁷ reported that the Rev/RRE pathway was essential for efficient antisense RNA (AS RNA)-mediated inhibition of HIV-1 replication. The levels of the targeted HIV-1 GagPol mRNA, encoding p24 (capsid protein) and structural core proteins, however, were not reduced. Furthermore both GagPol mRNA and the AS RNAs were found associated with polysomes¹²⁷. This project was undertaken to further explore the mechanism for the efficient AS inhibition observed with HIV vectors compared to other vector systems.

Chapter 2 – A double stranded RNA panhandle promotes antisense RNA-mediated inhibition of HIV viral production in a Rev/RRE dependent manner.

Several studies describing the use of antisense RNAs targeting either lentiviral mRNAs or mRNAs for receptors required for cell entry have been previously published^{153–163}. However, those studies did not demonstrate enough efficacy for further use in clinical trials due to poor transduction efficiency or minimal effects of the antisense RNA in the cells. More recently, a lentiviral vector, VRX496, that delivers a long antisense (AS) RNA directed at the HIV envelope (*env*) region was shown to be able to efficiently inhibit HIV replication^{126–128}. The original vector has been used in Phase I clinical trials after *ex vivo* delivery to CD4 cells and CD34 cells and has been well tolerated to date. A subsequent study showed evidence of A-to-G mediated changes in the *env* target region in HIV produced in cells expressing the AS RNA. Some of these changes resulted in nonfunctional Env proteins and were reported to be consistent with ADAR-mediated editing of RNA hybrids between AS and target RNAs¹²⁶. However, the mechanism for the very efficient inhibition was not further explored.

In a later study in our laboratory, using a similar HIV-1 vector system, it was reported that the efficient inhibition of HIV-1 replication by HIV-1 derived lentiviral antisense *env* RNA vectors required trafficking through the HIV-1 Rev/RRE pathway¹²⁷. The HIV-1 Rev/RRE pathway is best known for its ability

to overcome nuclear retention of HIV-1 unspliced and incompletely spliced transcripts and mediate their nuclear export^{36,38,39}. The inhibition was significantly reduced when Rev was replaced by RevM10-Tap (NXF1), a mutant form of Rev, which traffics through the NXF/NXT pathway, used by many cellular mRNAs. A mechanistic study of the Rev/RRE-mediated antisense inhibition reported no involvement of nuclear retention of HIV-1 target mRNA. The targeted mRNAs were exported to the cytoplasm and remained associated with polyribosome complexes. However, HIV-1 capsid protein, (p24) levels were greatly reduced. The p24 protein is expressed from a full length, unspliced HIV mRNA that is a target of the expressed AS RNA. The studies confirmed that the reduction of p24 protein levels was not due to non-specific effects, since the production of HIV-1 Nef protein, translated from a completely spliced HIV-1 transcript that lacks the AS target, was unaffected. Furthermore, the antisense RNAs affected neither particle assembly, nor release. Taken together, the data from these studies indicated that the AS RNA-mediated inhibition was not simply a result of editing and retention of the target mRNA in the nucleus mediated by double-stranded RNA (dsRNA) formation, as was suggested in the previous studies.

In the course of further studies with the vector described in Ward et al. 2009¹²⁷, we discovered that the constructs used in that study expressed an AS RNA which contained a 287 bp dsRNA panhandle formed by sequences in the env region that creates an inverted repeat. This panhandle was not present in the original vector used in the Lu et al. 2004 experiments¹²⁶. Here we show that

the vector containing the panhandle shows a significantly enhanced ability to inhibit HIV replication compared to the original AS RNA lacking a panhandle. In addition, we show that the 287 bp panhandle region is sufficient to achieve significant inhibition. In all cases, efficient inhibition requires the RNA to traffic on the Rev/RRE pathway.

Results

An AS RNA that contains a 287-bp panhandle is able to inhibit HIV replication more efficiently than an AS RNA that lacks it.

In the construction of the HIV vectors that were used in the Ward et al study¹²⁷ (**Figure 22**), a region that was complementary to part of the envelope (*env*) AS sequences were retained in the vectors. This region is located immediately downstream of the 937-bp long antisense *env* sequence and leads to the formation of a 287 bp panhandle in the RNA that is expressed from the antisense vectors (**Figure 23**). To be able to directly compare the inhibition obtained with these vectors to the vectors used Lu et al 2004 study¹²⁶, where the AS RNA lacked a panhandle, we removed the 287 nucleotide region 3' to the AS *env* insertion, to create three new vectors. The new vectors express an RNA that contains the 937 nt AS region, but lack the double-stranded RNA panhandle. (**Figure 24 and Materials and Methods** for the construction of the antisense RNA vector lacking a panhandle).

To directly compare the inhibition that could be achieved with these vectors, constructs with and without panhandles were co-transfected into 293T

cells together with the pNL4-3 proviral plasmid and inhibition was measured as a reduction in p24 levels, as previously described¹²⁷. The result from the analysis showed that RRE-driven antisense vector with a panhandle was about 10-fold more effective in inhibiting HIV than the corresponding vector lacking a panhandle at a molar ratio of 1:1 (**Figure 25A**). As before, very little inhibition was observed in the case of both the CTE-driven antisense vector and the antisense vector with no exporting element (**Figure 25A**). In the case of these vectors, the presence of a panhandle did not affect the inhibition. Furthermore, an experiment using different amounts of the AS vector with the RRE (25:1, 5:1, 1:1 and 1:5 molar ratios of target vector to AS vectors) showed that the vector expressing the AS RNA with the panhandle resulted in a significant inhibition of HIV p24 expression even at a ratio of 5:1, (**Figure 25B**). Thus this experiment confirmed that the RRE-containing AS RNA that contains a panhandle shows significant enhanced inhibition compared to the AS RNA without the panhandle.

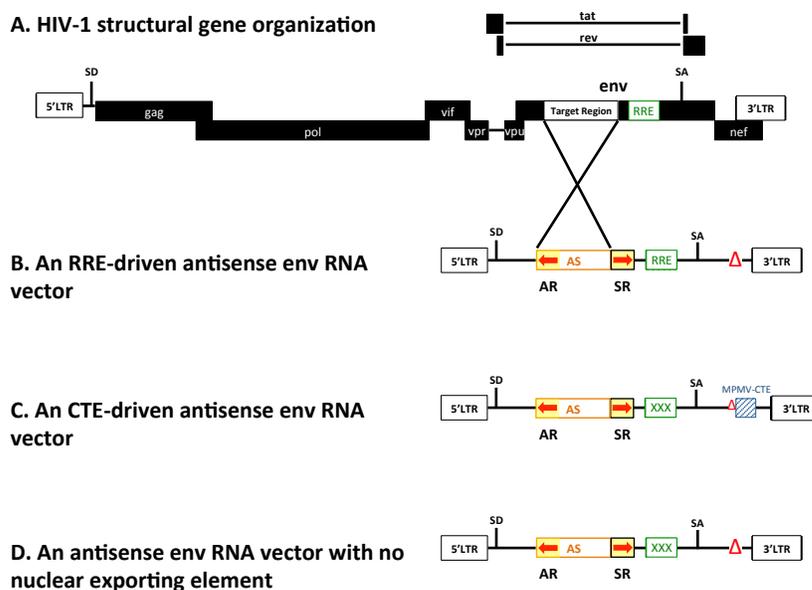


Figure 22. Diagrams showing (A) HIV-1 gene organization and antisense env RNA vectors used in the experiments; (B) RRE-driven antisense env RNA vector; (C) CTE-driven antisense env RNA vector with no nuclear exporting element.

Red arrows in yellow boxes represent the inverted repeats, which lead to the formation of the double-stranded RNA panhandle.

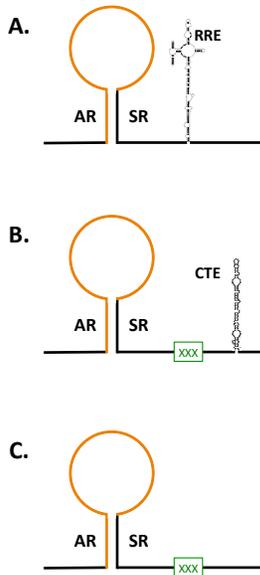


Figure 23. Diagrams showing hypothetical structure of antisense *env* RNA molecules and nuclear exporting elements

(A) RRE-driven antisense *env* RNA vector, (B) CTE-driven antisense *env* RNA vector, and (C) antisense *env* RNA vector with no nuclear exporting element. Orange line represents a 937-nt antisense *env* RNA sequence. AR and SR represent 287-nt antisense and sense *env* RNA repeats, respectively. RRE and CTE represent HIV-1 Rev Response Element and Mason Pfizer Monkey Virus (MPMV)'s Constitutive Transport Element. XXX is the mutated RRE.

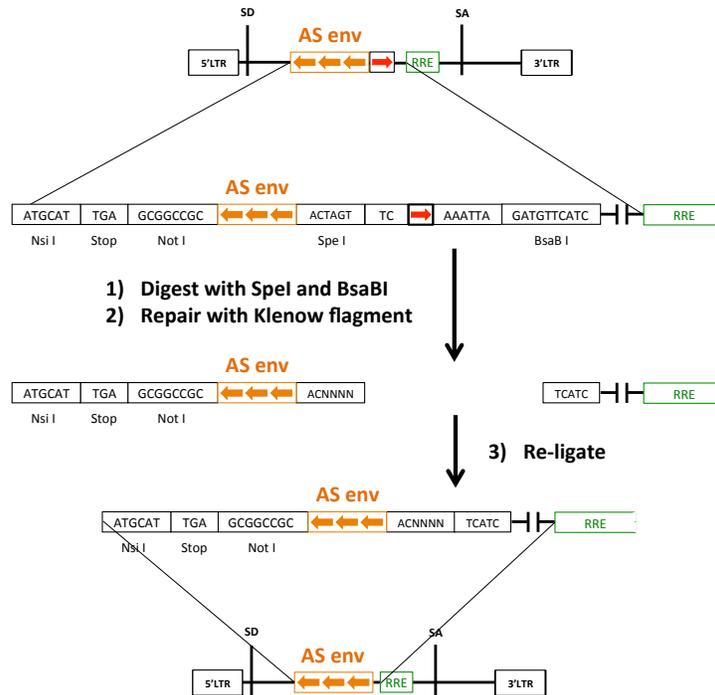


Figure 24. A diagram demonstrating the construction of antisense RNA vector lacking a panhandle

The RRE-driven antisense *env* RNA vector containing the panhandle is shown at the top the diagram. The orange box represents a 937-nt antisense *env* RNA sequence (AS *env*). The red arrow represents the 287-nt sense *env* RNA repeat. RRE represent HIV-1 Rev Response Element.

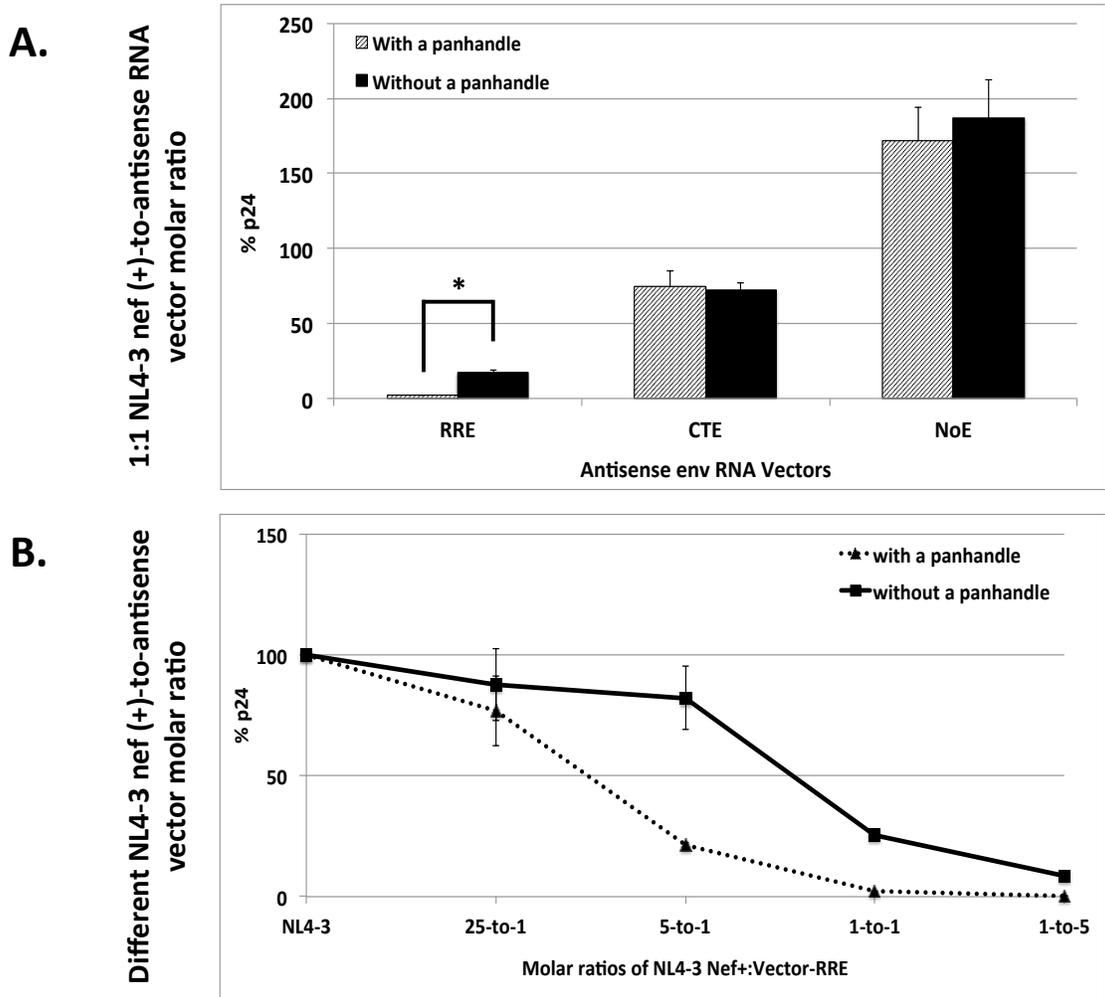


Figure 25. A deletion of the panhandle region of an RRE-driven antisense RNA (AS RNA) vector leads to a ten-fold decrease in inhibition of HIV replication

A total of 6.6 million 293T cells were co-transfected by calcium phosphate co-precipitation with an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the antisense vector constructs containing the RRE, CTE or no element (NoE). **(A)** an increasing amount of the RRE containing antisense vectors **(B)**. The antisense vectors with no panhandle have their sense repeat of the panhandle deleted. As a positive control, 293T cells were transfected with pNL4-3 Nef (+) alone. Cultured supernatants were collected 48 hours post-transfection then assayed for viral production by p24 ELISA. Data are presented as mean \pm standard deviation of two independent experiments. The % p24 is calculated from a ratio of a p24 value of a cultured supernatant of the 293T cells expressing NL4-3 and the antisense vector to that of a positive control. Statistical significance was assessed using a two-tailed unpaired Student's *t*-test. Error bars indicate Standard deviation means ($n=2$). *P*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances. **P* < 0.05.

An AS RNA containing a double-stranded RNA panhandle with only a 287-nt HIV specific AS sequence is sufficient for efficient inhibition.

The vectors expressing the AS RNA with a panhandle contains a 937-nucleotide region that is antisense to the HIV env region. To test whether an AS RNA containing a panhandle could inhibit HIV-1 p24 production, when only the 287-nucleotide “panhandle” region is antisense to HIV, we replaced the 650-nt HIV env sequences in the predicted loop region of all 3 antisense vectors with a 729-nt antisense EGFP sequence (pHR5247, pHR5253, and pHR5259). There was no target for the antisense EGFP in pNL4-3 nef (+) (pHR1145). For the transfection experiments, a total of 6.0×10^6 293T cells were transfected with a 1:1 molar ratio of the HIV target plasmid to the antisense vector constructs containing different targeting regions and export elements (pHR3476, pHR3477, pHR3478, pHR5284, pHR3756, pHR3757, pHR5151, pHR5153, pHR5155, pHR5247, pHR5253, and pHR5259, see Materials and Methods). The antisense RNA vectors are shown in **Figure 26**.

The RRE-driven antisense vectors (AS-RRE) containing an HIV-1 env panhandle, with either an original antisense HIV env sequence (AS env) loop or an antisense EGFP (AS EGFP) loop, effectively inhibited the p24 production, as shown in **Figure 27A** and **27C**, respectively. In contrast, inhibition with the vectors containing the CTE and no element was poor, although the CTE-driven antisense vector (AS-CTE) containing an env panhandle and an antisense EGFP (AS EGFP) loop showed increased inhibition of p24 production by 2 fold when compared to the original AS-CTE with an HIV-1 env loop (**Figures 27A** and

27B). As expected, no inhibition was observed when the AS RNA only contained GFP sequences (**Figure 27D**) and the antisense RNA-mediated inhibition obtained with the original vectors with and without panhandles (**Figure 27A and 27B**), were similar to the previous data shown in **Figure 25A**. Thus the results of this experiment clearly show that efficient inhibition was observed with the Rev/RRE-driven antisense RNA vector, even when the vector contained only the EGFP AS loop region and the 287 nt panhandle sequences that were derived from env (**Figure 26C and 27C**).

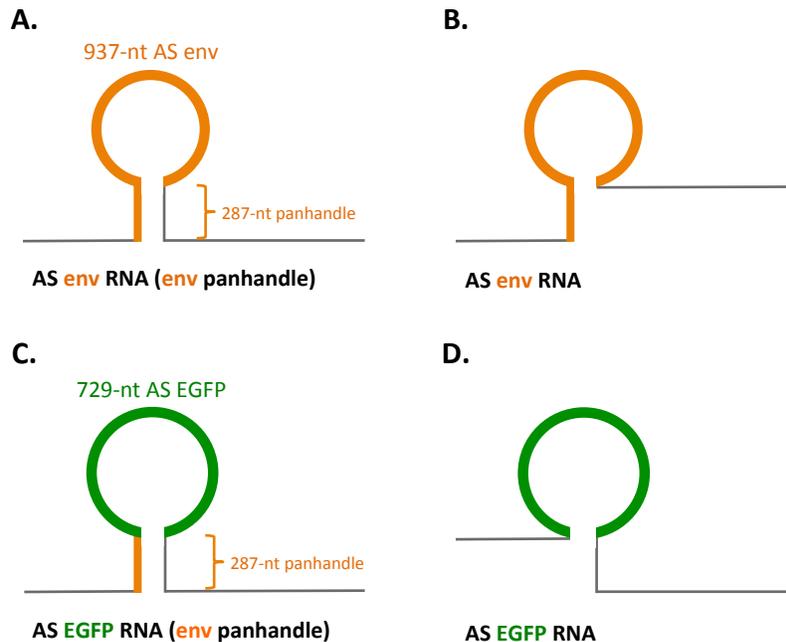


Figure 26. Diagrams of antisense RNA vectors used in an experiment shown in Figure 27

(A) A diagram representing the targeting region of the original antisense env RNA vectors containing HIV-1 env 287-nt dsRNA panhandle and 3 different export elements. **(B)** A diagram representing antisense env RNA vectors lacking an env panhandle. **(C)** A diagram representing antisense EGFP RNA vectors containing HIV-1 env 287-nt dsRNA panhandle. **(D)** A diagram representing the antisense vector expressing only EGFP AS RNA. See **Table 6** in Material and Methods for more detail.

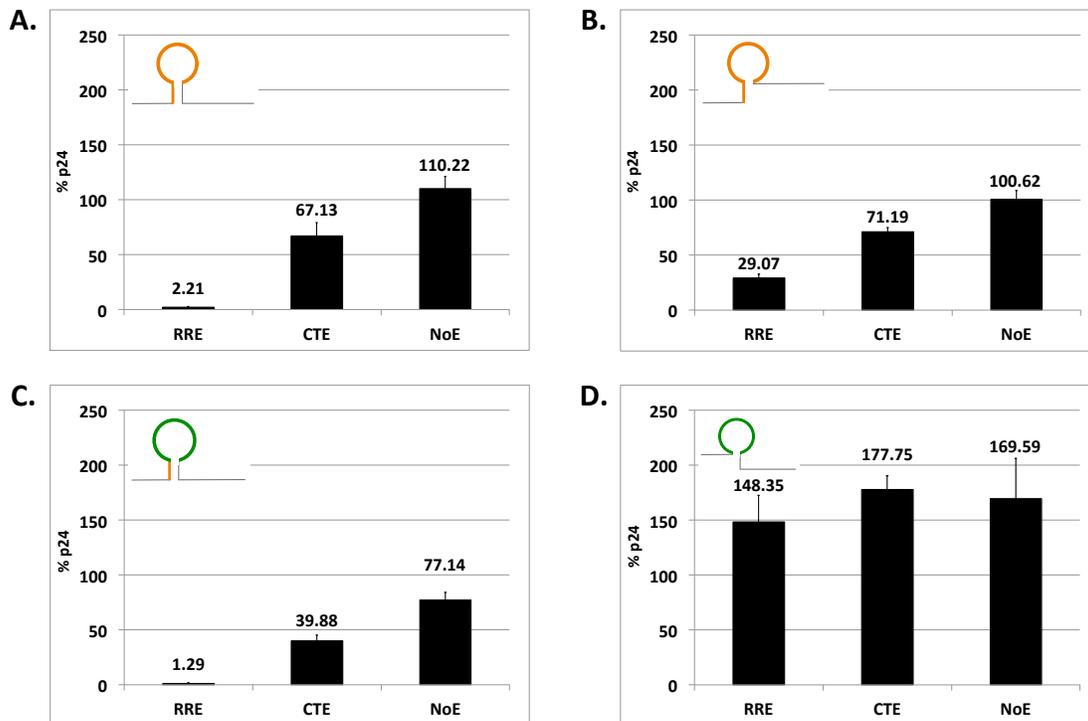


Figure 27. A 287 nt env panhandle alone is sufficient to an effective inhibition

A total of 6 million 293T cells were co-transfected using calcium phosphate co-precipitation with an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and one of the antisense vector constructs. (A) to (D) show the inhibition of HIV replication using 4 different sets of AS-RNA vectors (see Figure 4). Diagrams representing the AS-RNAs used in each experiment are shown in the upper left corner of the graphs. As a positive control, 293T cells were transfected with pNL4-3 Nef (+) alone. Culture supernatants were collected 48 hours post-transfection and assayed for viral production by p24 ELISA. Data are presented as mean \pm standard deviation of four independent experiments. The % p24 is calculated from a ratio of a p24 value of a culture supernatant of the 293T cells expressing NL4-3 and the antisense vector to that of a positive control.

An antisense vector that delivers an AS EGFP RNA with a panhandle is able to efficiently inhibit expression of a GFP containing Lentiviral vector, in an HIV Rev/RRE pathway-dependent manner.

In this study, we used an NL4-3 derived proviral construct, named pHIG (pHR4138) as a new target (Figure 28A). pHIG contains a HSA-IRES-EGFP sequence cassette sequence in the HIV Nef region¹⁶⁴. The murine heat stable antigen (HSA) and EGFP are translated from all classes of viral transcripts:

unspliced mRNA (**Figure 28B**), singly spliced mRNAs and multiply spliced mRNAs (**Figure 28C**). The HIV GagPol polyproteins, which will ultimately give rise to the p24 capsid proteins, are only produced from the unspliced mRNAs. To target pHIG, we constructed new AS vectors (containing RRE, CTE or no element) where the AS RNA is targeted to only EGFP (see Materials and Methods). The vectors were constructed with a 287 nt panhandle derived from EGFP sequences (**Figure 29**). To test whether these new antisense EGFP vectors could mediate the inhibition of expression from pHIG, 293T cells were transfected with a 1:1 molar ratio of the target plasmid, pHIG (pHR4138), to one of the antisense EGFP vector constructs. The previously described vectors expressing AS RNA to EGFP, but lacking a panhandle (see **Figure 26**) were also used in these experiments. The results are shown in **Figure 29A** and **29B**. As shown in **Figure 29**, the RRE-driven antisense EGFP vectors expressing the EGFP AS RNA containing an EGFP panhandle was able to efficiently inhibit p24 expression from pHIG. The inhibition was similar to that obtaining with the RRE-driven antisense env RNA vector with the env panhandle in conjunction with pNL4-3 (see **Figure 25**). However, as seen in **Figure 29B**, the original antisense EGFP vectors lacking the panhandle inhibited less well than when the panhandle was present. Better inhibition was observed when the AS RNA trafficked on the CTE pathway, compared to the experiments using pNL4-3, however with the CTE there was no difference between the AS RNA with and without the panhandle.

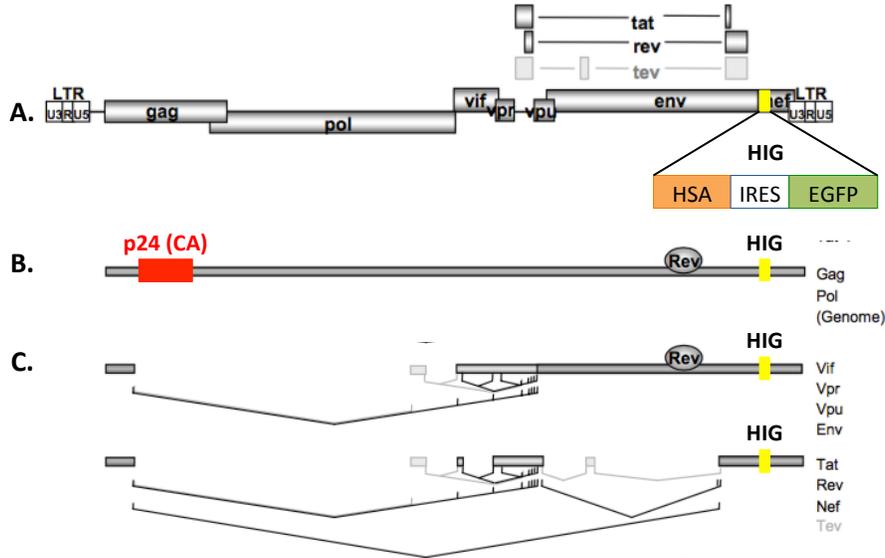


Figure 28. Diagrams showing a proviral clone pHIG and its transcripts

(A) A diagram of an pNL4-3-derived proviral clone pHIG (pHR4138; pNL4-3.HSA.R+.E- HSA-IRES-EGFP) containing a HIG (HAS-IRES-EGFP) sequence cassette in the HIV Nef region¹⁶⁴. **(B)** A diagram showing an unspliced pHIG transcript expressing all reporters: p24, HSA, and EGFP. **(C)** A Diagram showing both singly spliced and multiply spliced pHIG transcripts expressing only HSA and EGFP. Figures adapted from Kammler et al. *Retrovirology* 2006²³.

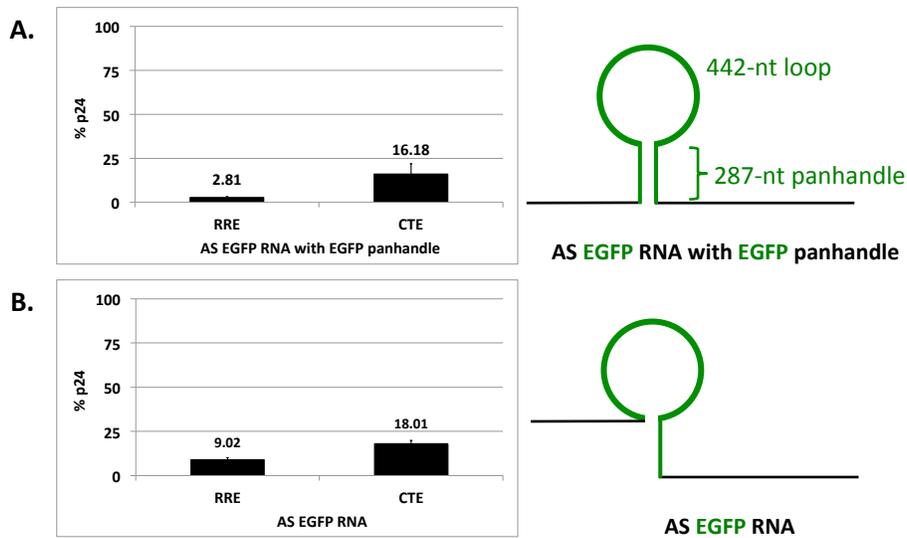


Figure 29. Inhibition of p24 expression from pHIG with AS vectors expressing EGFP AS RNA with and without a panhandle

A total of 6.0×10^6 293T cells were transfected with a 1:1 molar ratio of the target plasmid, pHIG (pHR4138), to **(A)** antisense EGFP vector constructs with an EGFP panhandle (i.e. pHR5249, pHR5255, and pHR5261). **(B)** antisense EGFP vectors with no panhandle (i.e. pHR5284, pHR3756, and pHR3757). The top panel shows % p24 levels. The right panel shows schematic diagrams of the antisense EGFP vectors (N = 4; 100% p24 = 4.36×10^5 pg/ml).

Discussion:

A 287-bp panhandle enhances the AS RNA-mediated inhibition of HIV-1 replication.

The initial experiments performed in this study showed that the AS RNA vector constructs (pHR3476, pHR3477, and pHR3478), previously used in the study by Ward et al., contain 287-bp inverted repeats in the AS RNA that is expressed from these vectors, i.e. AR and SR, as shown in **Figure 22**¹²⁷. The presence of inverted repeats generates a long double stranded RNA panhandle in the AS RNA, as shown in **Figure 23**. This panhandle leads to significantly increased viral inhibition, but only when the AS RNA traffics on the Rev/RRE pathway. The clinical VRX496 vector-mediated inhibition of HIV replication in natural host cells, i.e. CD4⁺ lymphocytes, showed efficient inhibition in the absence of the panhandle, as shown in **Figure 23**¹²⁶. For this reason, we decided to analyze whether removing the 287-bp dsRNA panhandle from the AS RNA vector constructs would increase or decrease the inhibition.

The deletion of the 287-bp SR inverted repeat (**Figure 22** and **Figure 23**), in the vectors should theoretically give rise to a longer single stranded antisense env RNA. We initially hypothesized that this might lead to better base pairing between the HIV-1 sense (mRNAs) and the antisense RNA, resulting in more effective inhibition of viral production from pNL4-3 Nef (+) (pHR1145). However, when the panhandle was removed in the RRE-driven AS RNA, the efficiency of the inhibition was decreased approximately 10-fold. The panhandle made no difference when the AS env RNA contained either a CTE or no export element.

(NoE) (**Figure 25**). This finding was thus contradictory to our original expectation that a longer perfect base pairing between the sense RNA (a targeted viral mRNA) and the RRE-driven AS RNA would yield more efficient inhibition.

The 287-bp panhandle derived from the HIV-1 env gene is sufficient for the efficient inhibition of HIV expression from pNL4-3.

Our experiments have clearly demonstrated that an RRE-driven AS env RNA containing a 287-bp panhandle mediates inhibition of HIV replication much more efficiently than an AS env RNA that lacks it. We thus also investigated whether Rev/RRE-dependent inhibition of HIV-1 replication could be maintained using a vector that contained a 287-bp panhandle derived from the HIV-1 env gene, but an EGFP antisense sequence in the predicted loop region (**Figure 26C**).

Our experiments also showed that the RRE-driven AS EGFP RNA vector with the panhandle derived from HIV-1 env sequences was able to still efficiently inhibit viral production (**Figure 27C**). The inhibition was similar to that mediated by the original RRE-driven AS RNA vector (**Figure 27A**). This finding confirms that the panhandle is sufficient to achieve efficient inhibition of viral replication, but only when the AS RNA traffics on the Rev/RRE pathway. When an AS EGFP RNA vector that lacked the env panhandle was used (**Figure 27D**), no inhibition was observed. This was an important control to show that no inhibition was observed when the AS RNA could not target the HIV mRNA. In conclusion, our experiments clearly demonstrate that the presence of a panhandle in an AS

RNA that traffics on the Rev/RRE pathway enhanced the inhibition of viral replication.

A 287bp EGFP panhandle in an EGFP AS RNA promotes efficient inhibition of EGFP expression.

In order to confirm that a panhandle can enhance AS mediated inhibition in a Rev/RRE dependent manner, a new set of AS EGFP RNA vectors containing an EGFP panhandle was used to target the pNL4-3 derived proviral clone pHIG (pHR4138), which contains an EGFP sequence in the Nef region. The presence of the EGFP panhandle in the AS RNA was shown to result in a significantly better inhibition of expression from pHIG, compared to an EGFP AS RNA that lacked a panhandle (**Figure 29A**). Again, the best inhibition was observed when the AS RNA trafficked on the Rev/RRE pathway. Interestingly, the CTE-driven AS EGFP RNA vectors with or without an EGFP panhandle were able to rather efficiently inhibit expression from pHIG (**Figure 29B**). Although we have not investigated the relative abundances of the different mRNA species of pHIG in the transfected 293T cells by a Northern blot analysis, we predict that the relative abundance of pHIG transcripts would be similar to that of a HIV provirus or a proviral clone of pNL4-3. However, all transcripts expressed from pHIG will be expected to contain the EGFP target sequence (**Figure 28**). It is, therefore, possible that the base pairing between the multiply spliced Rev/RRE independent pHIG sense RNAs and the AS EGFP RNAs would yield more dsRNA substrates for ADAR editing, These could potentially be further

processed to yield interfering small RNAs, resulting in better inhibition of viral replication, as shown in **Figure 29**. Nonetheless, the RRE-driven antisense RNA vector with an EGFP panhandle (**Figure 29A**) still gave the most effective inhibition, consistent with the experiments targeting HIV env.

Potential Mechanism of Inhibition

In the original studies of the VRX496 AS RNA effects, it was hypothesized that the inhibition of HIV replication was related to the fact that double-stranded RNAs can be targeted by RNA editing enzymes, known as Adenosine Deaminases Acting on RNA (ADARs)¹²⁶. ADAR enzymes bind to dsRNAs and convert adenosines to inosines by deamination. It was further hypothesized that the edited RNA would be retained in the nucleus and degraded. However, Ward et al. demonstrated that the target mRNA was exported to the cytoplasm, where it was stably associated with the translation machinery.

The very efficient inhibition, even with sub-stoichiometric amounts of the AS RRE-RNA, occurs when the AS-RRE RNA contains a panhandle. It is possible that this region forms dsRNA, which is then edited by ADAR enzymes, since this would only require an intra-molecular base pairing. This could lead to editing in most of the AS RNA molecules co-transcriptionally, soon after transcription or anywhere along the Rev-RRE pathway, independent of interactions with the target HIV mRNA. Further processing of this RNA could then lead to the formation of small RNAs (e.g. miRNAs) that could inhibit target RNA translation. In the absence of a panhandle, dsRNA would only form if the

AS-RNA base pairs with the target env region in the HIV mRNAs. This will occur post-transcriptionally and is likely to be rather inefficient, explaining the reduction in inhibition seen with the AS-RNA lacking a panhandle. However, this could still result in ADAR editing and processing to small RNAs that could inhibit HIV production, albeit not as efficiently.

The potential relationship between the panhandle, the HIV-1 Rev/RRE pathways, and ADAR editing will be further discussed in Chapter 3. However, two major conclusions that can be drawn from these experiments are that the presence of a panhandle in the AS RNA targeting HIV-1 env sequences or EGFP sequences enhances the inhibition and that efficient inhibition requires that the antisense RNA traffics on the HIV-1 Rev/RRE pathway. Taken together, these results suggest that efficient AS RNA HIV vectors could be developed to target almost any gene, simply by including a double stranded region in the AS RNA.

Chapter 3 – Analysis of ADAR editing during AS RNA-mediated inhibition

The previous published study by Ward et al., as well as the results presented in Chapter 2, suggest that RNAi mechanisms may be involved in the inhibition, since very small amounts of AS RNA were required for efficient inhibition and the inhibition appeared to occur at the level of mRNA translation¹²⁷. Lu et al. 2004 reported that DNA sequencing analysis of HIV episomal DNA, isolated from infected SupT1 cells stably expressing AS RNA, showed many adenosine-to-guanosine (A-to-G) transitions in the targeted region¹²⁶. They proposed that these changes may have been the result of editing by ADAR enzymes (ADAR)¹²⁶.

As described in the introduction, ADARs are expressed in most metazoan species¹⁴³. In humans, ADAR enzymes are expressed in most tissues¹⁶⁵. The human genome contains three different ADAR genes: ADAR1 (encoding 2 ADAR1 isoforms: ADAR1 long form p160 and ADAR1 short form p110), ADAR2 and catalytically inactive ADAR3^{141,142}. The expression of ADAR1 long form (ADAR1L) is stimulated by type I and II interferon responses, and this isoform localizes to the cytoplasm^{166–168}. ADAR1 p110 and ADAR2 are constitutively expressed, and both localize in nucleus. ADAR3, which appears to lack catalytic activity, is mainly expressed in the brain. ADAR 1 and 2 catalyzes Adenosine-to-Inosine (A-to-I) editing in either partially or perfectly double-stranded RNAs^{141,169,170}. To date, ADAR enzymes have been shown to have both pro- and anti-viral activities^{146–148,171,172}.

The experiments presented in Chapter 2 demonstrated that the AS RNA vector containing the panhandle shows a significantly enhanced ability to inhibit HIV replication compared to an AS RNA lacking a panhandle. We also showed that the 287 bp panhandle region was sufficient to achieve significant inhibition. In both cases, efficient inhibition required trafficking of the AS RNA on the Rev/RRE pathway. We thus decided to perform additional experiments to investigate if ADAR enzymes are involved in the inhibition and if these enzymes edit AS and/or the targeted HIV mRNAs.

Results

DNA sequencing analysis of the panhandle region of cDNA made from AS RNA shows a high level of A-to-G changes when the AS RNA traffics on the Rev/RRE pathway.

To test whether the AS panhandle was potentially edited by ADAR enzymes and how this was affected by trafficking on either the Rev/RRE or Nxf1 pathway, a total of 10 millions 293T cells were co-transfected with 3.3 µg of the RRE-containing vectors expressing the env AS RNA with the env panhandle ((pHR3476 or pHR3478, see diagram in **Figure 30**) and either a plasmid expressing HIV-1 Rev, or a plasmid expressing a HIV-1 RevM10-Nxf1 fusion protein together with a plasmid expressing Nxt1 protein. HIV-1 RevM10-Nxf1 lacks the Rev NES but still binds to the RRE. It thus directs the RRE-containing RNA down the Nxf1 export pathway. As controls, 293T cells were transfected with either the RRE-driven AS vector or an AS RNA vector with no export

element alone. 500 ng of the expression plasmid pCMVTat was also included in all transfections in order to promote the transcription from the HIV-LTR containing antisense RNA vectors. Calcium phosphate coprecipitation was used for transfection.

At forty-eight hours post-transfection, total RNA was extracted. The specific primers were, then, used to synthesize and amplify cDNAs representing the sense and antisense repeats in the AS RNAs. Bulk PCR products were initially sequenced using Sanger DNA sequencing analysis (**Figure 30**). Data are presented as %G, which indicates the percent A to G changes at each adenosine in the antisense and sense repeats of the panhandle. %G is calculated from a ratio of an area under the curve of a guanosine chromatogram to that of an adenosine at the same position. As can be seen in **Figure 30A**, the cDNA made from HIV AS RNA that trafficked on the Rev/RRE pathway (**Figure 30A**) showed the most A-to-G changes in both the antisense and sense repeats (see also **Table 1**) and highest level of changes (i.e. % G). Although the cDNA made from the AS RNA that trafficked on the Nxf1 pathway showed some A to G changes (**Figure 30B**), there were significantly fewer sites edited and lower levels of editing at those sites. The same was true when no Rev was present or when the AS RNA had no export element (**Figure 30C** and **30D**). No significant A-to-G changes were observed in the single strand region (the loop) in cDNA from any of the antisense RNA vectors (see **Figure 31**).

The high level of A to G changes in cDNA from the AS repeat region of the AS RNA that trafficked on the Rev/RRE pathway (**Figure 30A**) was confirmed by

NGS analysis using an Illumina miSeq instrument, followed by analysis using Geneious Software (see **Figure 32A**). Again, no changes were found in the non-repeat regions. In these experiments, we also co-transfected cells with the AS-RRE plasmid (and the plasmid expressing Rev) with plasmids that express either human ADAR1L or ADAR2S (these were kind gifts from Mary O'Connell. Overexpression of ADAR1L increased the level of A-to-G changes in the antisense repeat (AR), and some changes were now also seen in the non-repeat region, see **Figure 32B**. Meanwhile, overexpression of ADAR2S led to increased A-to-G changes only in the antisense repeat (AR), see **Figure 32C**. **Figure 33** presents the quantification of A-to-G changes shown in **Figure 32**. These results indicate that the AS RNA interacts with the ADAR machinery and that it can be edited by both ADAR1L and ADAR2S.

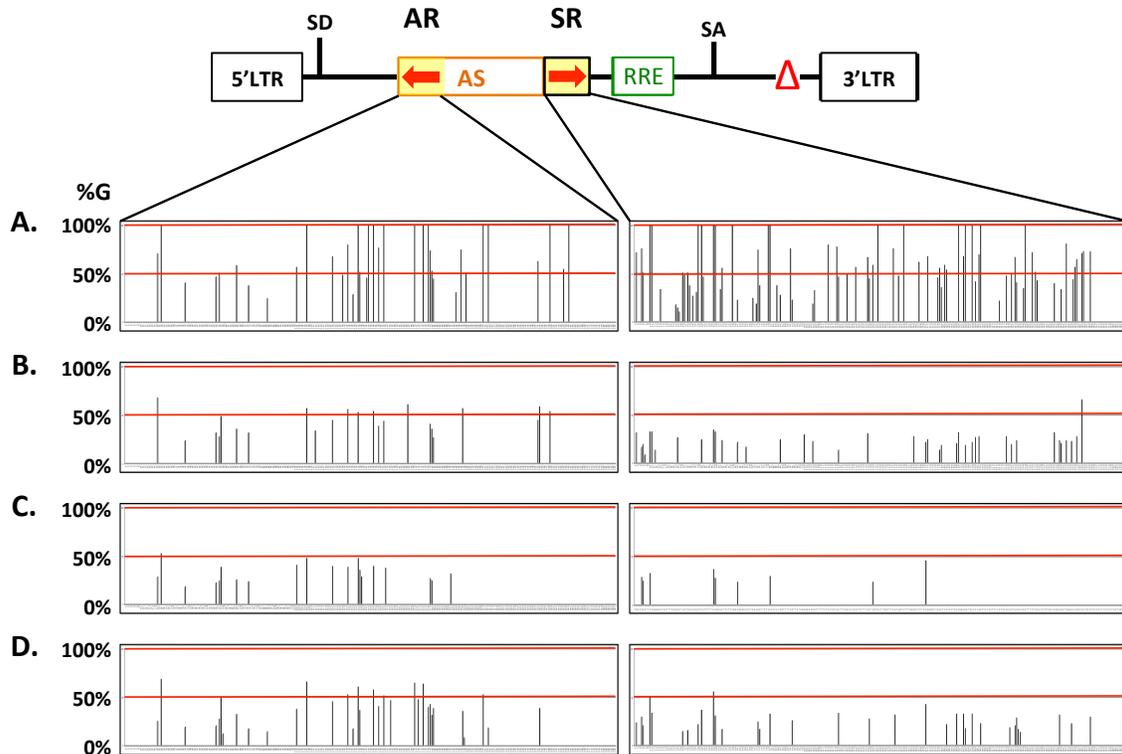


Figure 30. High level of A-to-G changes (%G) in the 287-bp panhandle when the RRE-driven antisense env RNA vector traffics on the Rev/RRE pathway even in the absence of the target. No A-to-G changes were observed in the loop region.

(A) Histograms of the changes (%G) observed in antisense and sense repeats in cDNA from RRE-driven antisense env RNA trafficking through the HIV-1 Rev/RRE pathway in the presence of HIV-1 Rev wt. **(B)** Histograms of the changes (%G) observed in antisense and sense repeats in cDNA from RRE-driven antisense env RNA trafficking through the Nxf1 pathway in the presence of HIV-1 RevM10-Nxf1 and Nxt1. **(C)** Histograms of the changes (%G) observed in cDNA from antisense and sense repeats of an RRE-driven antisense env RNA in the absence of Rev. **(D)** Histograms of the changes (%G) observed in antisense and sense repeats in cDNA from antisense env RNA with no export element expressed in 293T cells. A schematic diagram of the AS-RNA vector with the HIV-1 RRE is shown at the top of the figure.

IDs	Number of A-to-G changes in Antisense Repeat	Number of A-to-G changes in Sense Repeat
30A	36	81
30B	23	41
30C	20	9
30D	33	35

Table 1. The number of adenosine (A) in both the antisense and sense repeats that show A to G changes as seen in Figure 30

The total number of adenosines in the antisense and the sense repeats of the panhandle are 73 and 109, respectively.

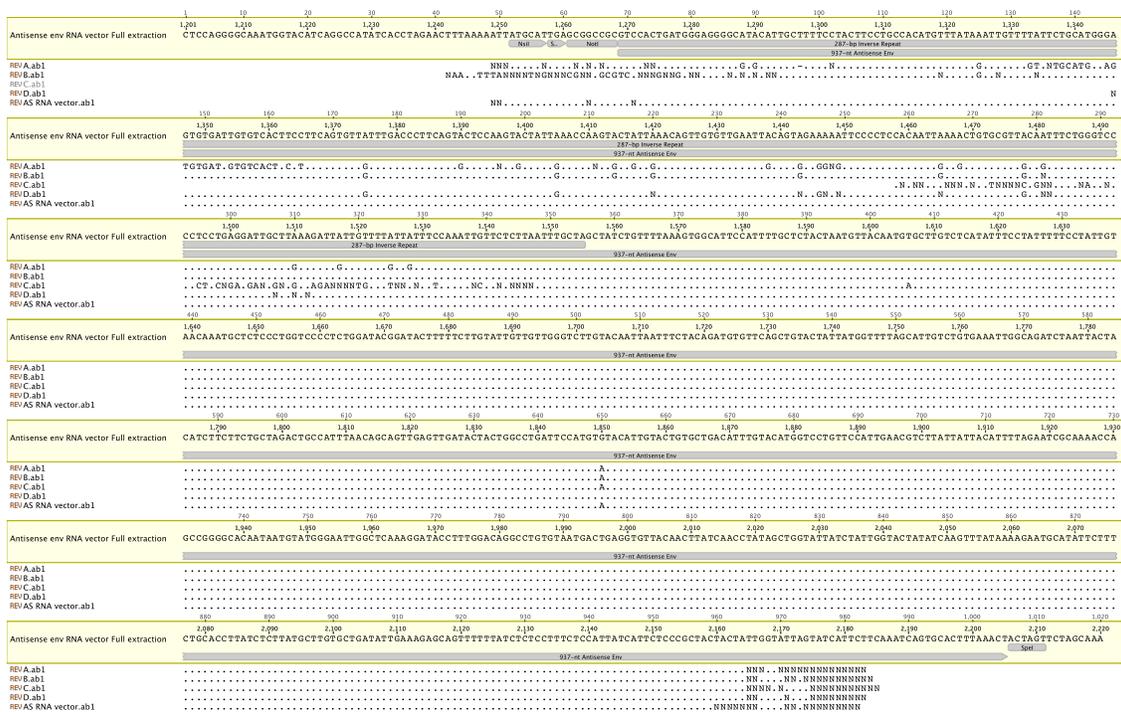


Figure 31. No A-to-G changes were observed in the loop region

Sanger DNA sequencing of bulk PCR products showing nucleotide changes observed in antisense repeat and the loop region in cDNAs from (A) RRE-driven antisense env RNA trafficking through the HIV-1 Rev/RRE pathway in the presence of HIV-1 Rev wt, (B) RRE-driven antisense env RNA trafficking through the Nxf1 pathway in the presence of HIV-1 RevM10-Nxf1 and Nxt1, (C) RRE-driven antisense env RNA in the absence of Rev, and (D) antisense env RNA with no export element expressed in 293T cells. (AS RNA vector) The bulk PCR product of the RRE-driven antisense env RNA vector (pHR3476) was sequenced. In the diagram, the loop region locates downstream of the 287-bp inverse repeat.

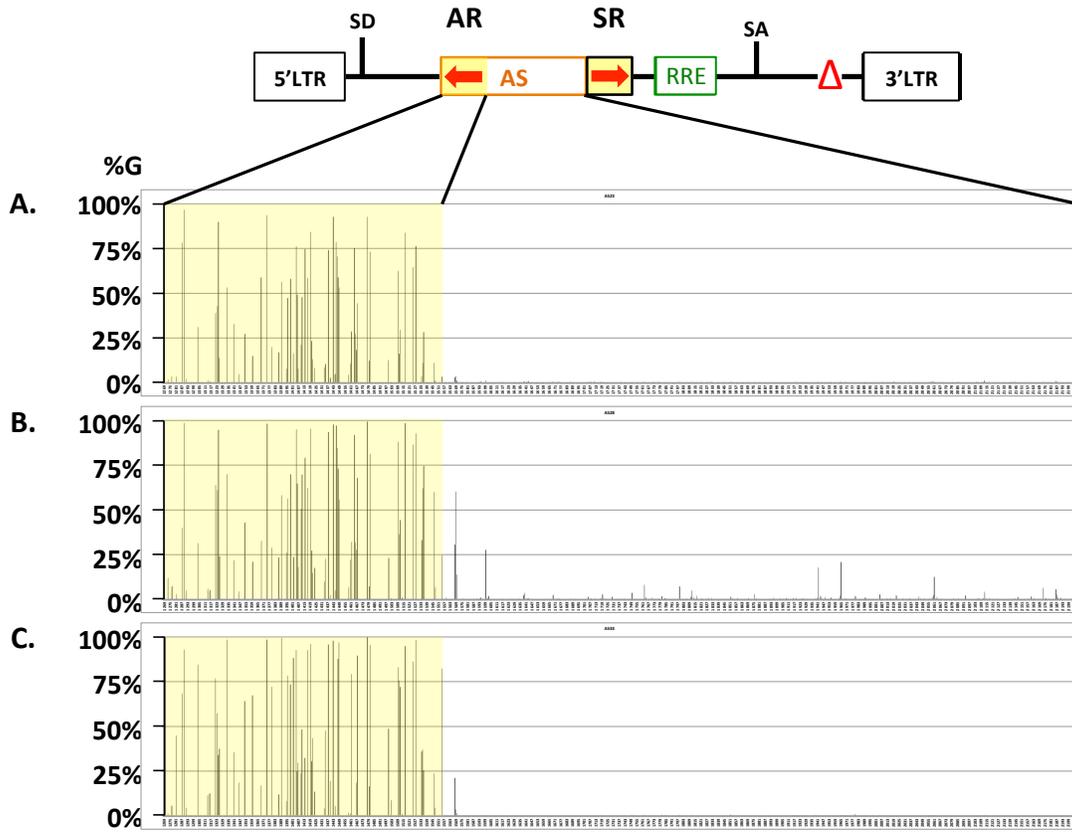
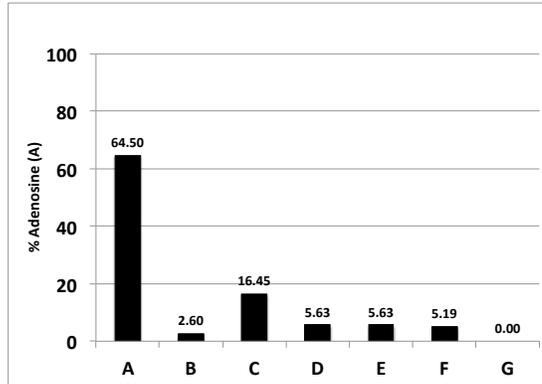
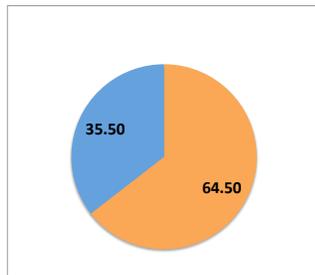


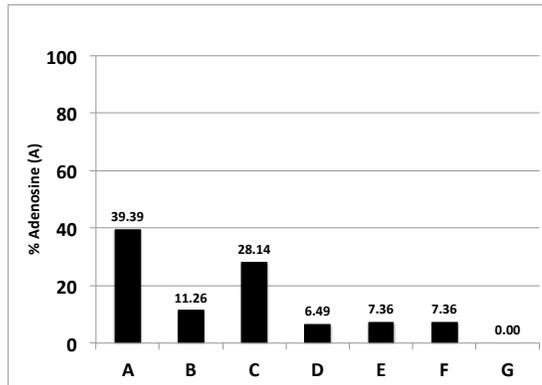
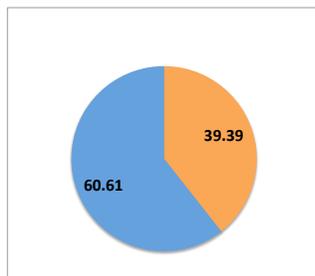
Figure 32. Next Generation Sequencing (NGS) data of the antisense repeat (AR) and loop region of cDNA made from RRE-driven antisense env RNA (AS-RRE) in the presence of Rev

(A) A histogram of % A-to-G changes (%G) in cDNA from AS-RRE RNA in the presence of Rev. **(B)** A histogram of % A-to-G changes (%G) in cDNA from AS-RRE RNA in the presence of Rev and ADAR1L over-expression. **(C)** A histogram of % A-to-G changes (%G) in cDNA from AS-RRE RNA in the presence of Rev and ADAR2S over-expression.

A. NO ADAR overexpression



B. ADAR1L overexpression



C. ADAR2s overexpression

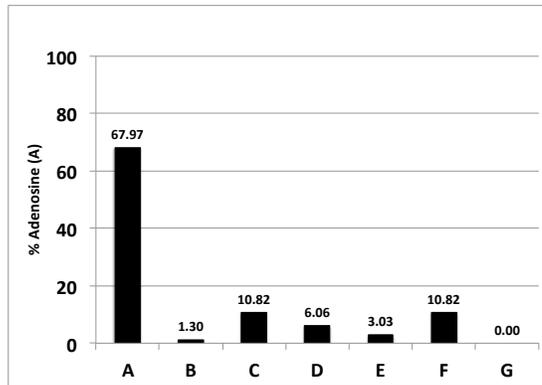
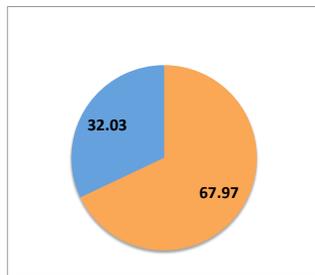


Figure 33. The quantification of A-to-G changes observed in the antisense repeat (AR) and loop region of cDNA made from RRE-driven antisense env RNA (AS-RRE) in the presence of Rev as shown in Figure 32

The Pie charts show the % of adenosines that show ADAR-mediated editing (blue) or no editing (orange) from (A) AS-RRE RNA in the presence of Rev, (B) AS-RRE RNA in the presence of Rev and ADAR1L over-expression, and (C) AS-RRE RNA in the presence of Rev and ADAR2S over-expression. The adjacent histograms show different levels of editing observed in cDNAs The x-axis of the histograms show the different levels of ADAR editing at any particular residue. i.e. A: No editing; B: editing was < 1.0%; C: editing was between 1% and 25.0%; D: editing was between 25.0% and 50.0%; E: editing was between 50.0% and 75.0%; F: editing was between 75.0% and 100 %; G: editing was 100%

Overexpression of huADAR1L or huADAR2s impairs RRE-driven AS inhibition.

The results presented in the first part of this chapter demonstrated that, a significant level of A-to-G changes was observed in a cDNA obtained from the panhandle regions of the antisense vector trafficking on the HIV-1 Rev/RRE pathway, in the absence of an HIV target mRNA. Furthermore, overexpression of either ADAR1 or 2 increased A to G changes, suggesting that the changes were a result of ADAR editing. When the HIV target is present, the viral mRNA and the antisense RNA would potentially form a double stranded RNA, which would also be expected to be a substrate for ADAR editing.

To further investigate the potential role of human ADAR enzymes in the antisense RNA-mediated inhibition of HIV-1 viral particle production, we performed experiments in which human ADAR1L p160 (pHR4039), ADAR1 p110 (pHR4922), or ADAR2s (pHR4044) were cotransfected with plasmids expressing the HIV-1 env panhandle-containing AS RNAs with an RRE (pHR3476) or without an RRE (pHR3478) and pNL4-3 (pHR1145) (see **Figure 34**). These experiments showed that overexpression of all three ADAR enzymes resulted in a significant loss of inhibition when the AS RNA trafficked on the Rev/RRE pathway. In contrast, ADAR expression had no significant effect when the AS RNA lacked the RRE.

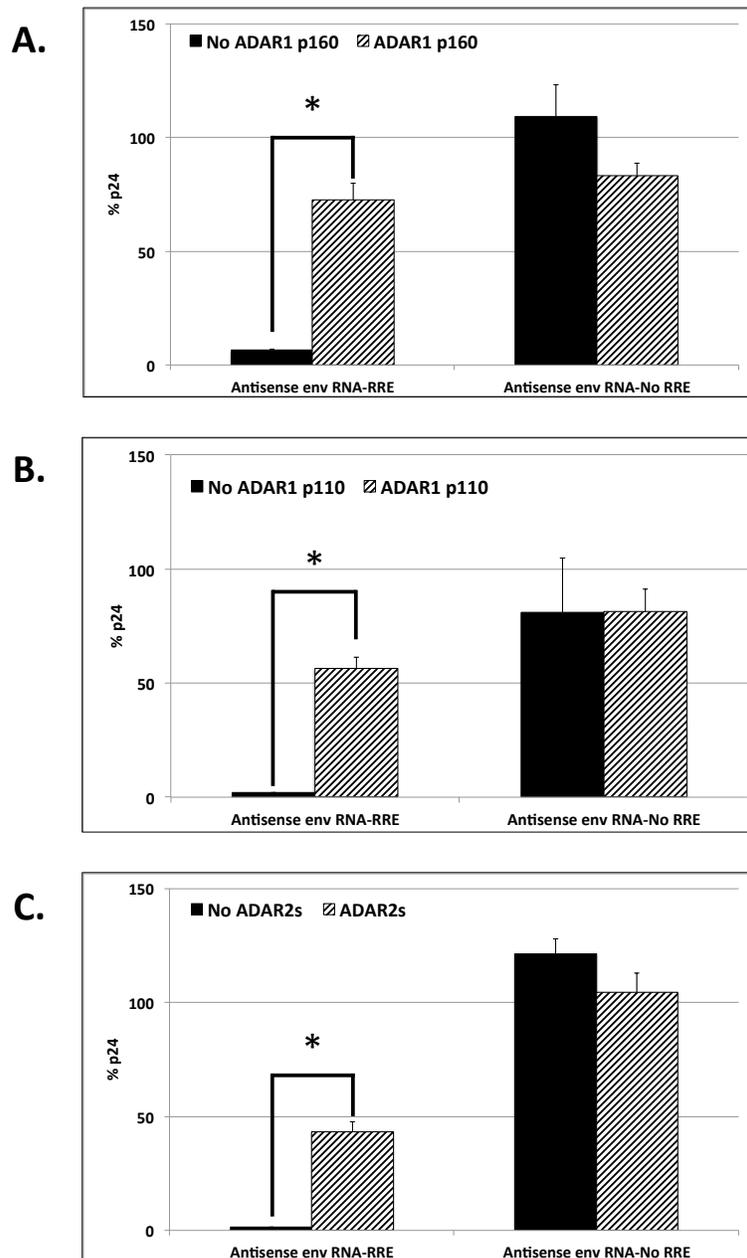


Figure 34. An overexpression of wild-type huADAR1 p160, huADAR1 p110, or huADAR2s significantly impairs RRE-driven antisense RNA-mediated inhibition of HIV replication

A total of 6 million 293T cells were co-transfected with an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense RNA vector or a vector expressing antisense RNA with no element (NO RRE) with or without either (A) huADAR1 p160, (B) huADAR1 p110 or (C) huADAR2 overexpression. Culture supernatants (viral stocks) were collected and assayed for viral production by p24 ELISA, 48 hours post-transfection. Data are presented as mean \pm standard deviation of independent experiments (A) N = 2, (B) N = 2 or (C) N = 10. *p*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances. **P* < 0.05.

We next conducted similar experiments with ADAR2s that included transfections with plasmids expressing wtADAR2s and three mutants (huADAR2s with a mutated dsRNA binding domain (EAA) X 2, huADAR2s with a deleted N-terminus, lacking the NLS ($\Delta 4-72$) and a catalytically inactive mutant of huADAR2s). Cells were transfected with an increasing amount of either of these plasmids, the AS-RRE vector or the no element (NO RRE) vector and the pNL4-3 proviral plasmid. In the case of wtADAR2s, the inhibition by AS-RRE RNA was decreased in a dose-dependent manner, confirming the results presented above. As before, little or no effect of ADAR2s were observed with the AS-RNA lacking the RRE (**Figure 35A**). When the ADAR2s ($\Delta 4-72$) mutant was expressed, inhibition was completely abolished even with very small amounts of transfected plasmid (**Figure 35C**). Since this mutant lacks the NLS and has been shown to be localized to the cytoplasm¹⁷³, this result suggests that ADAR2s can interfere with AS inhibition at the cytoplasmic level. In contrast to these results, the huADAR2s mutant with a non-functioning dsRNA binding domain (EAA X2) and the catalytically inactive huADAR2s had no effect on the inhibition (**Figure 35B** and **35D**). Thus RNA binding and catalytic activity are both important for the ability of ADAR2s to interfere with the inhibition. Overall, these findings confirm the hypothesis that huADAR enzymes and RNA editing are involved in AS- RNA-mediated inhibition, when the AS-RNA traffics on the Rev/RRE pathway.

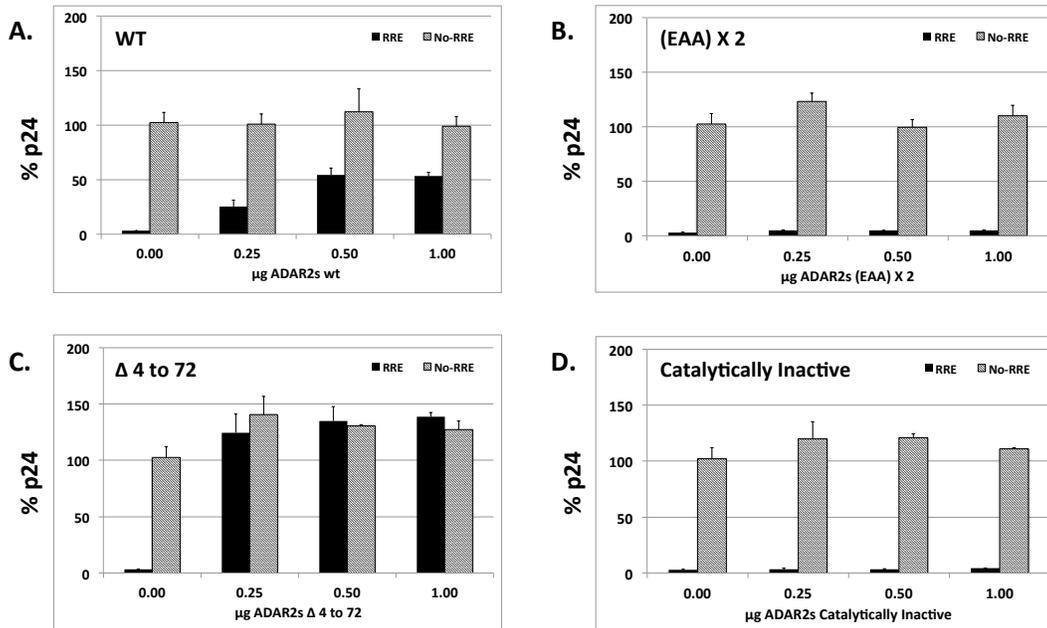


Figure 35. Overexpression of huADAR2s wt or an N terminus-deleted mutant impairs the inhibition, but this is not observed with huADAR2s lacking dsRNA binding domains or catalytic activity.

A total of 3 million 293T cells were co-transfected with an equi-molar amounts of an pNL4-3 Nef (+) proviral plasmid and the RRE-driven AS-RNA vector or AS-RNA vector with no RRE. In addition, the cells were transfected with an increasing amount of plasmid expressing **(A)** wild type (WT) huADAR2s **(B)** a huADAR2s with a mutated dsRNA binding domain (EAA) X 2, **(C)** a huADAR2s with a deleted N-terminus (Δ 4-72) or **(D)** a catalytically inactive huADAR2s. Culture supernatants (viral stocks) were collected and assayed for viral production by p24 ELISA, 48 hours post-transfection. Data are presented as mean \pm standard deviation of independent experiments (N=2).

Analysis of potential ADAR editing in AS-RNA and target HIV mRNAs with and without ADAR1L or ADAR2s overexpression

Since the experiments described above, showed that an overexpression all three ADAR enzymes inhibited AS RNA-mediated inhibition when the AS-RNA traffics on the Rev/RRE pathway (see **Figure 34**), we decided to analyze for evidence of ADAR editing in AS RNA in the presence of HIV target mRNA and determine if this changed when either ADAR1L (p160) or ADAR2s were overexpressed. ADAR1 p110 was not included in this study. To do this, we

harvested total RNAs from transfected 293T cells 48 hours post-transfection from an experiment similar to that described in **Figure 34**, and subjected specific AS and target mRNA cDNA amplicons to NGS analysis using an Illumina miSeq instrument. The procedures for cDNA synthesis, PCR amplification and library preparation are described in **Material and Methods**. Obtained sequences were analyzed using Genieous Software. As seen in **Figure 36**, the results showed that, in the absence of ADAR overexpression, high levels of A-to-G changes were observed in the AS-RRE RNA vector throughout the whole region that was complementary to the target mRNA (non highlighted region). When either huADAR1L or huADAR2s was overexpressed, the A to G changes increased in the panhandle region of the AS RNA, but decreased in the rest of the AS sequences (highlighted by yellow boxes). In the case of the AS-RNA vector lacking an RRE, there was no evidence of editing outside of the “panhandle” region whether or not ADAR was overexpressed.

We were also interested in investigating potential ADAR editing in the target env region. To this end, we used total RNA to synthesize cDNA and PCR amplify the target region in the HIV mRNAs in an experiment where the vector expressing RRE-driven AS RNA was cotransfected with pNL4-3. The results of an NGS analysis of these amplicons (**Figure 37**), showed that, whether or not huADAR proteins were overexpressed, very low levels of A-to-G changes (%G) were seen throughout the region targeted by the AS RNA and overexpression of either huADAR1L or huADAR2s did not increase the A-to-G changes in the target region.

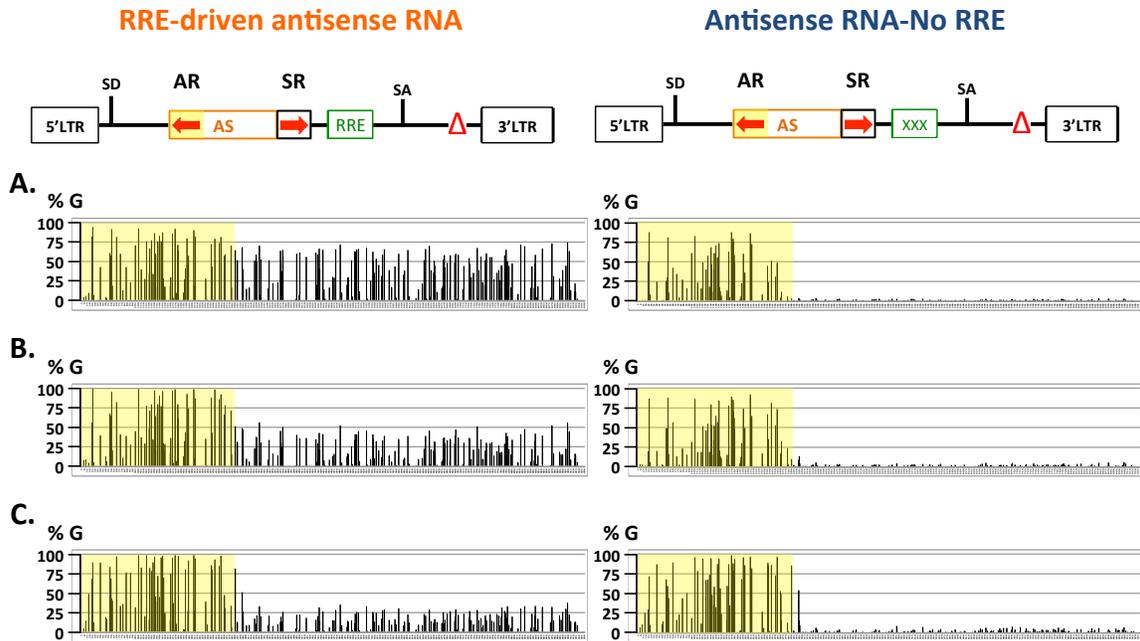


Figure 36. With a very effective inhibition, relatively high levels of A-to-G changes are observed throughout the antisense env sequence of the RRE-driven antisense RNA (AS-RRE)

NGS data of the antisense env sequence of AS-RNA with RRE isolated from 293T cells expressing an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense env RNA construct or the antisense env RNA lacking RRE construct. When compared to **(A)** a normal huADAR expression, an overexpression of either **(B)** huADAR1L or **(C)** huADAR2s results in a much lower A-to-G changes in the region outside the panhandle (yellow box). Significantly high editing level is shown in the antisense repeat region, which is a part of the panhandle of the antisense RNA vectors containing the RRE or without the RRE **(A to C)**. Diagrams of AS-RNA with or without HIV-1 RRE are shown above graphs showing % A-to-G change (%G) in the antisense env sequence. Yellow box highlights the antisense repeat (AR) region.

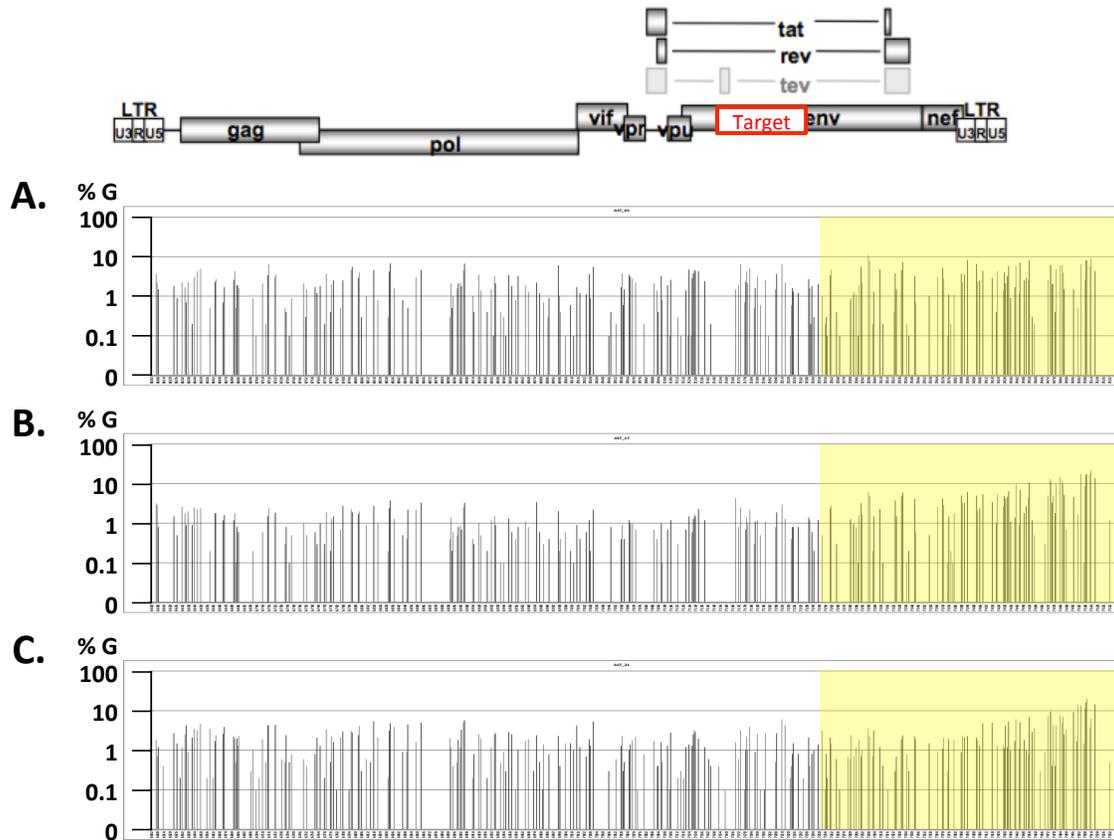


Figure 37. Low-level A-to-G changes were observed in the target region (AST) of the AS-RNA trafficking through the HIV-1 Rev/RRE pathway

NGS data of the target sequence (AST) of AS-RNA isolated from 293T cells expressing an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense env RNA construct (AS-RNA with RRE) **(A)** Under normal huADAR expression, an NGS analysis shows low-level % A-to-G changes in AST targeted by AS-RNA with RRE. An overexpression of either **(B)** huADAR1L or **(C)** huADAR2s results in a similar A-to-G changes in the AST to that targeted by the AS-RNA with RRE under normal huADAR expression. A diagram of HIV-1 (NL4-3) genome and the location of the target sequence (AST), a red text box, are shown above. The yellow highlighted region represents the region of the AST that is complementary to the antisense repeat (AR) region of the AS-RNA. % A-to-G changes (%G) are shown in the graphs using a logarithmic scale (\log_{10} scale).

We were intrigued by the fact that the AS target sequence (AST) in the HIV mRNA of the AS-RNA that traffics through the HIV-1 Rev/RRE pathway, showed only low levels A-to-G changes. This was observed whether or not huADAR was overexpressed during RRE-driven antisense RNA-mediated inhibition of HIV-1 replication. To further analyze this, we analyzed the A-to-G changes in individual cDNA clones of the 362 nucleotide target sequence (AST) of HIV-1 mRNA present in either total or cytoplasmic RNA fractions in the absence or presence of overexpressed ADAR2s (see **Figure 38**). As shown in **Figure 38A**, of the 57 clones isolated from the total RNA of the cells without ADAR2 overexpression, 42 clones showed 2 or fewer A to G changes, 8 showed between 3 to 10 A-to-G changes per clone and two sets of the 3 AST clones showed between 11 to 20 and 21 to 30 A-to-G changes per clone. Interestingly, one AST clone showed 173 A-to-G changes. When ADAR2 was overexpressed (**Figure 38B**), 77 of the 96 clones showed 2 or fewer A to G changes, 12 clones showed between 3 and 6 changes and 7 clones showed more than 50 A-to-G changes per clone. The highest number of changes in a clone was 109. The investigation of the A-to-G changes in the AST clones from the cytoplasmic RNA revealed relatively similar distributions. In the absence of ADAR2s (**Figure 38C**) forty-one of 54 clones showed 2 or fewer changes 5 showed between 3 to 10 changes and 8 showed between 11 to 50 changes per clone. In the presence of ADAR2s (**Figure 38D**), 47 of 57 clones showed 2 or fewer changes, 3 showed between 3 to 10 changes and 7 showed between 11 to 50 changes per clone.

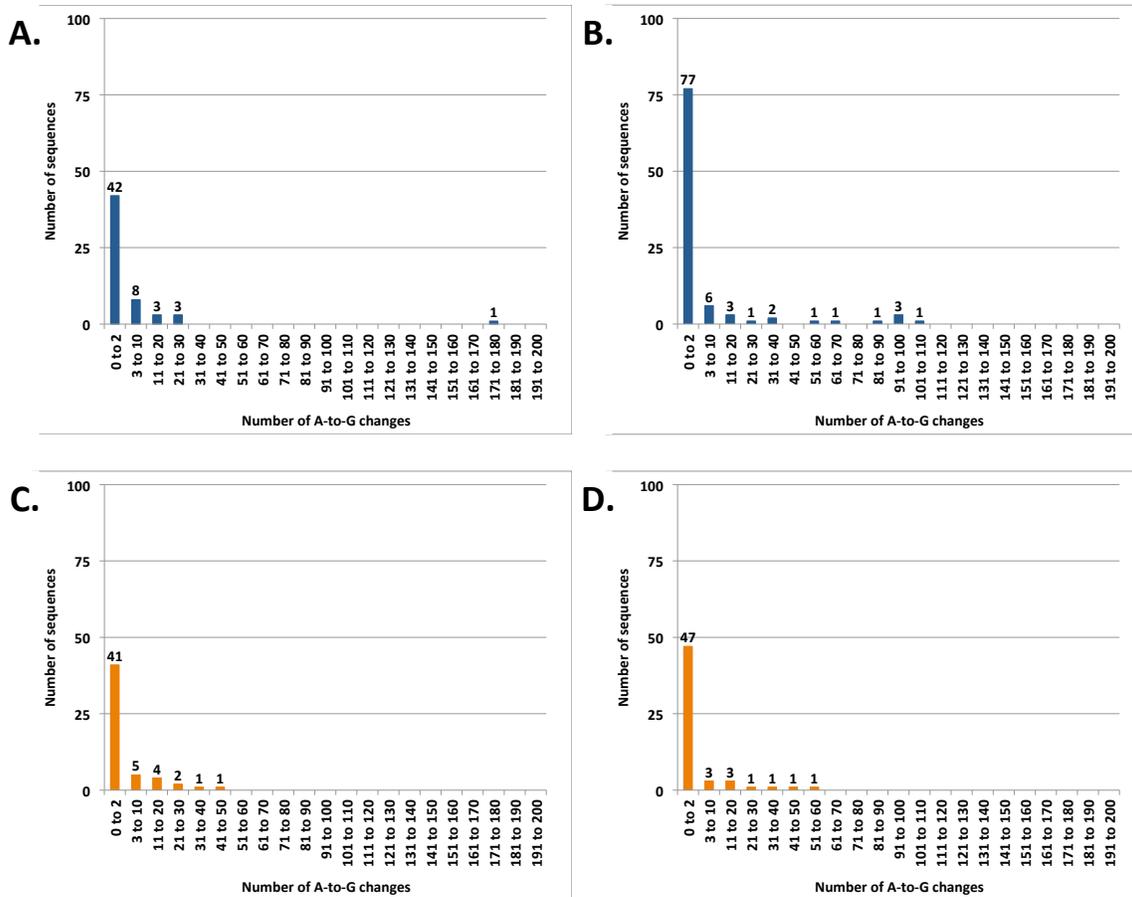


Figure 38. The distribution of individual ADAR-edited (A-to-G changes) clones of the cDNA of AS-RRE RNA target region from total and cytoplasmic HIV mRNA with and without ADAR2s overexpression

293T cells were co-transfected with an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense RNA vector (AS-RNA with RRE). In an additional transfection experiment, 293T cells were also co-transfected with a plasmid expressing a wild type huADAR2s. cDNA of AST region in HIV RNA was synthesized using a specific primer. Then, TopoTA cloning was performed on bulk PCR products of the AST cDNAs. The sequences of AST clones were analyzed for A-to-G changes. The AST clones that were edited showed more than 3 or more A-to-G changes. **(A)** and **(B)** shows the number of AST clones from total RNA of the transfected 293T cells with and without ADAR2 overexpression, respectively. **(C)** and **(D)** shows the number of AST clones from cytoplasmic RNA of the transfected 293T cells with and without ADAR2 overexpression, respectively.

As shown in **Table 2**, in cDNA obtained from total RNA, 15/57 AST clones (26%) showed A-to-G changes in the absence of ADAR overexpression. When huADAR2s was overexpressed, 19/96 clones (20%) showed A-to-G changes. In cDNA from the cytoplasmic RNA fraction, 13/54 AST clones (24%) showed A-to-G changes. When huADAR2s was overexpressed, 10/ 57 clones (18%) showed A-to-G changes. The percentage of AST clones with A-to-G changes were thus similar in cDNA from total (26% and 20% of the clones, from no ADAR2 overexpression and from ADAR2 overexpression, respectively) and the cytoplasmic fraction (24% and 18% of the clones, from no ADAR2 overexpression and from ADAR2 overexpression, respectively), supporting the notion that editing is not related to retention of edited target RNA in the nucleus.

AST clones from total RNA

Conditions	Clones showing ADAR editing	Clones not showing ADAR editing	Total number of clones
A. RRE AS RNA-mediated inhibition	15 (26%)	42 (74%)	57
B. RRE AS RNA-mediated inhibition + ADAR2 wt	19 (20%)	77 (80%)	96

AST clones from cytoplasmic RNA

Conditions	Clones showing ADAR editing	Clones not showing ADAR editing	Total number of clones
C. RRE AS RNA-mediated inhibition	13 (24%)	41 (76%)	54
D. RRE AS RNA-mediated inhibition + ADAR2 wt	10 (18%)	47 (82%)	57

Table 2. Analysis of A-to-G changes in the AS-RRE RNA target region in cDNAs from total and cytoplasmic HIV mRNA with and without ADAR2s overexpression as shown in Figure 38

The 5' and 3' nearest-neighbor preferences pattern of edited As in the AS RNA compared to the reported patterns associated with RNA editing by a particular ADAR enzyme.

5' and the 3' nearest-neighbor preferences of the edited adenosines (A) can be used to potentially predict which ADAR enzymes are responsible for the RNA editing, as shown in **Table 3** and **4**. These preferences were originally described in 1994, but have recently been updated as more data has become available^{174–176}.

Enzymes	5' nearest-neighbor preferences	3' nearest-neighbor preferences	References
ADAR1	U = A > C > G	Not determined	Polson et al. EMBO 1994 ¹⁷⁴
ADAR2	U = A > C = G	U = G > C = A	Polson et al. EMBO 1994 ¹⁷⁴

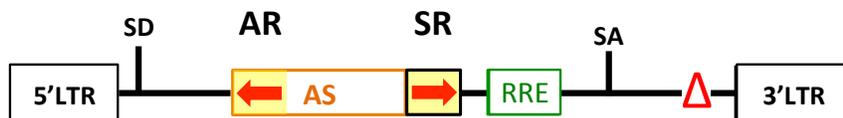
Table 3. A table showing the “old” 5' and 3' nearest-neighbor preferences of edited Adenosine (A)

Enzymes	5' nearest-neighbor preferences	3' nearest-neighbor preferences	References
ADAR1	U > A > C > G	G > C = A > U	Eggington et al. Nat. Commun. 2011 ¹⁷⁵
ADAR2	U > A > C > G	G > C > U = A	Eggington et al. Nat. Commun. 2011 ¹⁷⁵
ADAR1	U > A > C > G	G > C ~ A > U	Kuttan et al. PNAS 2012 ¹⁷⁶
ADAR2	U > A > C > G	G > C > U ~ A	Kuttan et al. PNAS 2012 ¹⁷⁶

Table 4. A table showing the “new” 5' and 3' nearest-neighbor preferences of edited Adenosine (A)

In the antisense repeat (AR), of the AS RNA Vector shown in **Figure 30**, neither the 5' nor the 3' nearest-neighbor preferences pattern of edited As in the AS RNA match any of the reported “old patterns” associated with RNA editing by a particular ADAR enzyme (compare **Table 3** to **Figure 39**). Although the 5' nearest-neighbor preferences pattern of all conditions, except for that of **B**, match

those of the “new patterns”, the 3’ nearest-neighbor preferences patterns could not determine the specific ADAR enzymes (compare **Table 4** to **Figure 39**).



Antisense Repeat: 5’ nearest-neighbor preferences

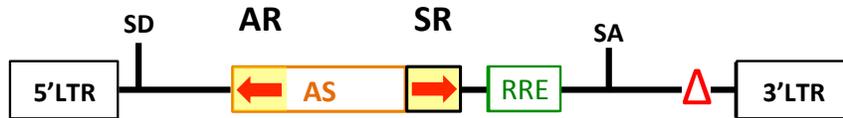
Conditions	A	C	G	U	Patterns
A. RRE-driven AS RNA Rev wt	10	4	1	21	U > A > C > G
B. RRE-driven AS RNA RevM10-Nxf1 & Nxt1	6	2	2	13	U > A > C = G
C. RRE-driven AS RNA	6	4	0	10	U > A > C > G
D. AS RNA with no export element	11	5	1	16	U > A > C > G

Antisense Repeat: 3’ nearest-neighbor preferences

Conditions	A	C	G	U	Patterns
A. RRE-driven AS RNA Rev wt	10	8	4	14	U > A > C > G
B. RRE-driven AS RNA RevM10-Nxf1 & Nxt1	10	5	0	8	A > U > C > G
C. RRE-driven AS RNA	7	7	1	5	A = C > U > G
D. AS RNA with no export element	11	7	4	11	A = U > C > G

Figure 39. The 5’ and 3’ nearest-neighbor preferences pattern of ADAR editing in the antisense repeat (AR) of the AS RNA vector (Figure 30)

As shown in **Figure 40**, neither the 5’ nor the 3’ nearest-neighbor preferences pattern of edited As in the AS RNA match any of the reported patterns associated with RNA editing by a particular ADAR enzyme in the sense repeat (SR) of the AS RNA Vector shown in **Figure 30** (compare tables 3 and 4 to **Figure 40**).



Sense Repeat: 5' nearest-neighbor preferences

Conditions	A	C	G	U	Patterns
A. RRE-driven AS RNA Rev wt	37	15	6	23	A > U > C > G
B. RRE-driven AS RNA RevM10-Nxf1 & Nxt1	16	7	12	6	A > G > C > U
C. RRE-driven AS RNA	3	1	2	3	A = U > G > C
D. AS RNA with no export element	18	1	8	8	A > G = U > C

Sense Repeat: 3' nearest-neighbor preferences

Conditions	A	C	G	U	Patterns
A. RRE-driven AS RNA Rev wt	33	13	17	18	A > U > G > C
B. RRE-driven AS RNA RevM10-Nxf1 & Nxt1	16	4	16	5	A = G > U > C
C. RRE-driven AS RNA	5	1	3	0	A > G > C > U
D. AS RNA with no export element	17	5	4	9	A > U > C > G

Figure 40. The 5' and 3' nearest-neighbor preferences pattern of ADAR editing in the sense repeat (SR) of the AS RNA vector (Figure 30)

RRE-driven AS RNA-mediated inhibition is improved when ADAR1 expression is reduced.

Our experiments demonstrated that overexpression of either ADAR1 (ADAR1 p160 and p110) or ADAR2 significantly impaired RRE-driven antisense RNA-mediated inhibition of HIV-1 viral particle production. To further investigate the potential role of human ADAR enzymes in the antisense RNA-mediated inhibition, we used shRNAs to deplete either the human ADAR1 or human ADAR2 enzymes and analyzed the effect on the AS inhibition. We performed experiments in which either shRNA against huADAR1 (**Figure 41A**) or huADAR2 (**Figure 41B**) mRNAs was cotransfected with plasmids expressing the HIV-1 env panhandle-containing AS RNAs with an RRE (pHR3476) or without an RRE (pHR3478) and pNL4-3 (pHR1145) (see **Figure 41**). Both shRNAs against human ADAR1 and human ADAR2 were also tested for their ability to reduce protein expression in transfected 293T cells (see **Figure 42** and **Figure 43**, respectively). These experiments (**Figure 41A**) showed that expression of huADAR1 shRNA resulted in a significant improvement of the inhibition, when the AS RNA trafficked on the Rev/RRE pathway. In contrast, huADAR1 shRNA had no significant effect when the AS RNA lacked the RRE. **Figure 42** shows that both ADAR1p160 and ADAR1p110 were significantly reduced in the presence of the shRNA. Surprisingly, this finding suggests that reduction of huADAR1 improves the RRE-driven antisense RNA-mediated inhibition. In contrast, shRNA against ADAR2 had no significant effect on the inhibition

(Figure 41B), in spite of the fact that the shRNA was able to significantly reduce levels of ADAR2 in cells expressing this protein (Figure 43).

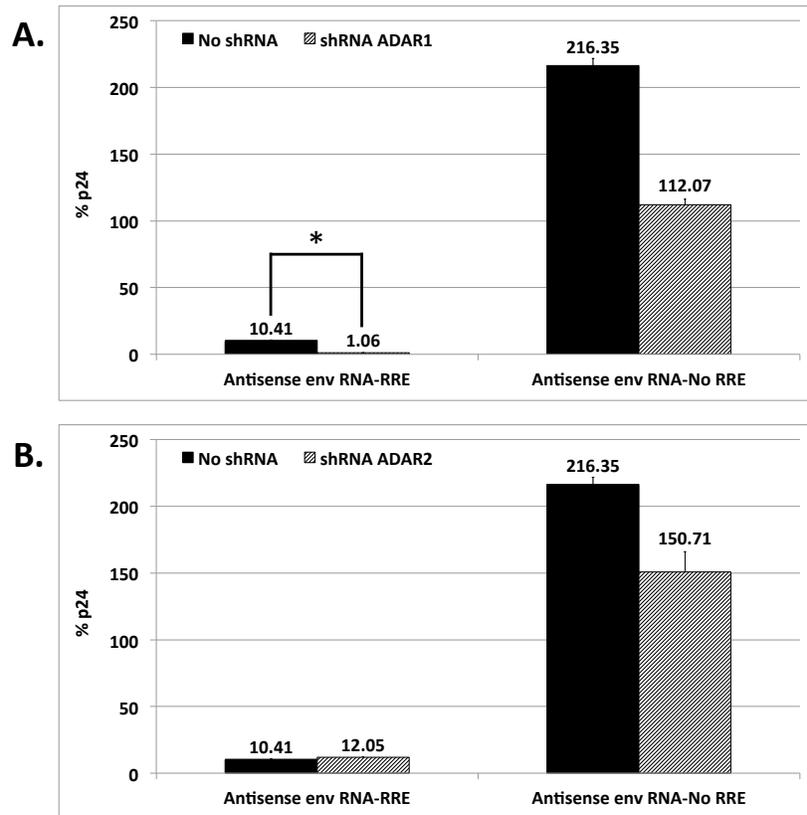


Figure 41. RRE-driven antisense env RNA-mediated inhibition is improved when the producer cells are treated with shRNA against huADAR1

A total of 6 million shRNA-treated 293T cells were co-transfected with an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense RNA vector or a vector expressing antisense RNA with no element (NO RRE) with or without either (A) 1.5 μ g of shRNA against all human ADAR1 variants or (B) 1.5 μ g of shRNA against all human ADAR2 variants. Culture supernatants (viral stocks) were collected and assayed for viral production by p24 ELISA, 60 hours post-transfection. Data are presented as mean \pm standard deviation of independent experiments (A) N = 3 and (B) N = 3. *p*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances. **P* < 0.05.

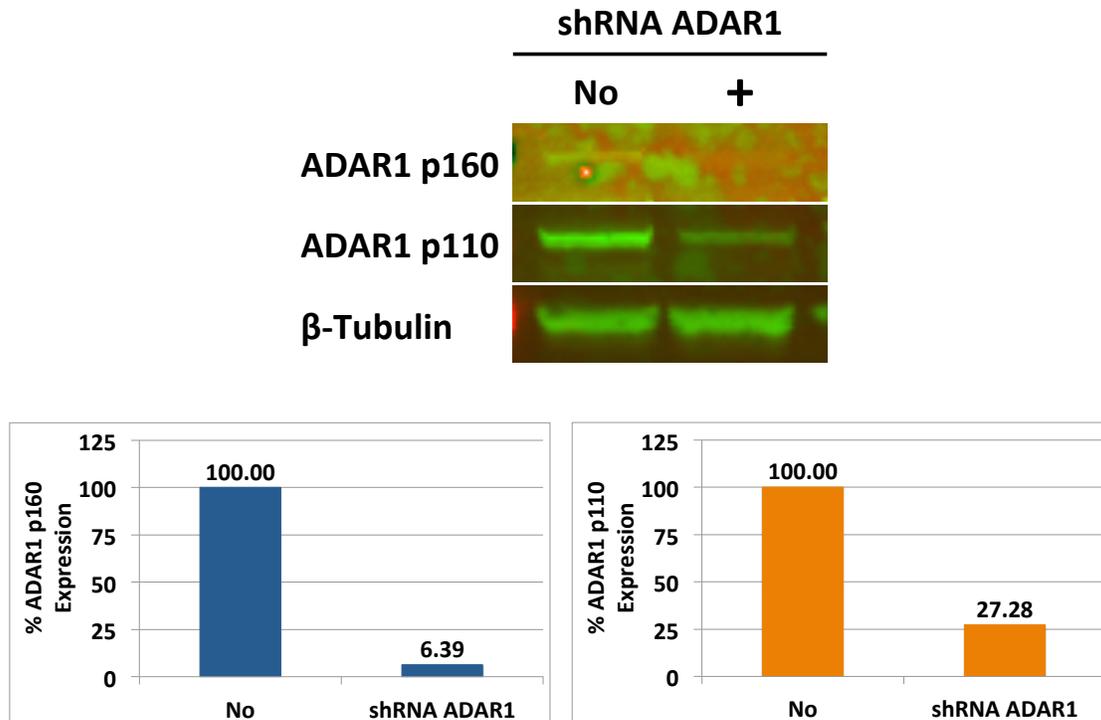


Figure 42. shRNA against ADAR1 significantly reduces ADAR1 p160 expression and moderately reduce ADAR1 p110 expression.

Western Blot analysis of transfected cells from the experiments, shown in **Figure 1A**, shows significantly reduced levels of both huADAR1 p160 and p110 expressions after shRNA treatment. The quantification of protein expression is shown in the histogram. No error bar shown here since all triplicated lysates were pooled together.

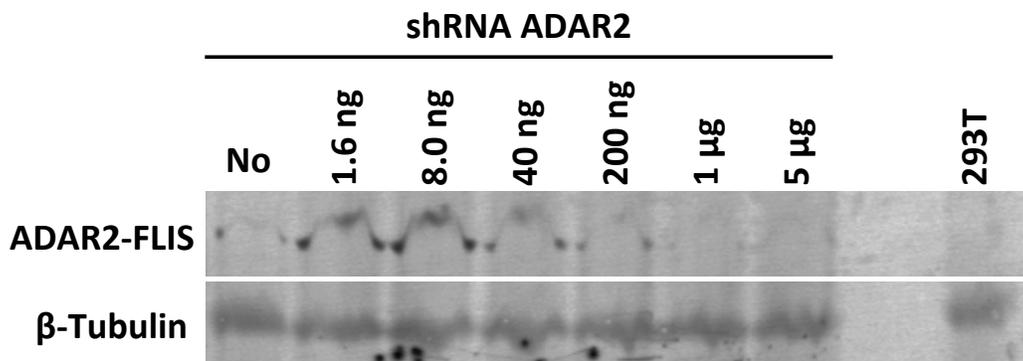


Figure 43. shRNA against huADAR2 in 293T cells overexpressing FLAG-tagged and HIS-tagged huADAR2 (huADAR2-FLIS) are able to significantly reduce ADAR2 expression.

3 million 293T cells were co-transfected with 1 μg of pHR4921 ADAR2s wt and an increasing amount of shRNA against ADAR2. Cell pellets were harvested 48 hrs post-transfection. 30 μl of each cell lysate was resolved on 12% PAGE. A mouse anti-FLAG primary antibody and a rabbit anti-β-tubulin polyclonal antibody were used. Western Blot analysis of transfected cells shows significantly reduced levels of huADAR2 expressions after shRNA treatment.

Analysis of evidence of ADAR editing in AS-RNA and target HIV mRNAs with and without ADAR1 or ADAR2 shRNA

Since the experiments described above showed that a depletion of huADAR1 enhanced AS RNA-mediated inhibition when the AS-RNA traffics on the Rev/RRE pathway (see **Figure 41B**), we decided to analyze for evidence of ADAR editing in AS RNA in the presence of HIV target mRNA (**Figure 44A**) and determine if this changed when either ADAR1 (**Figure 44B**) or ADAR2 (**Figure 44C**) were depleted. We harvested total RNAs from transfected 293T cells 60 hours post-transfection from an experiment described in **Figure 41**, and subjected specific AS and target mRNA cDNA amplicons to NGS analysis using an Illumina miSeq instrument. The procedures for cDNA synthesis, PCR amplification and library preparation are described in **Material and Methods**. Obtained sequences were analyzed using Geneious Software. As seen in **Figure 44B**, the results showed that, in the presence of shRNA against ADAR1, lower levels of A-to-G changes were observed in the AS-RRE RNA vector throughout the whole region that was complementary to the target mRNA, including the highlighted panhandle region. In the presence of shRNA against ADAR2 (see **Figure 44C**), the A-to-G changes in the entire region of the AS RNA were similar to that observed in the absence of shRNA expression (see **Figure 44A**). In the case of the AS-RNA vector lacking an RRE, there was no evidence of editing outside of the “panhandle” region, whether or not the shRNAs were present (**Figure 44C**). However, huADAR1 reduction led to decreased levels of

A-to-G changes in the panhandle, indicating that the panhandle is edited by ADAR1 even when the vector lacks the RRE.

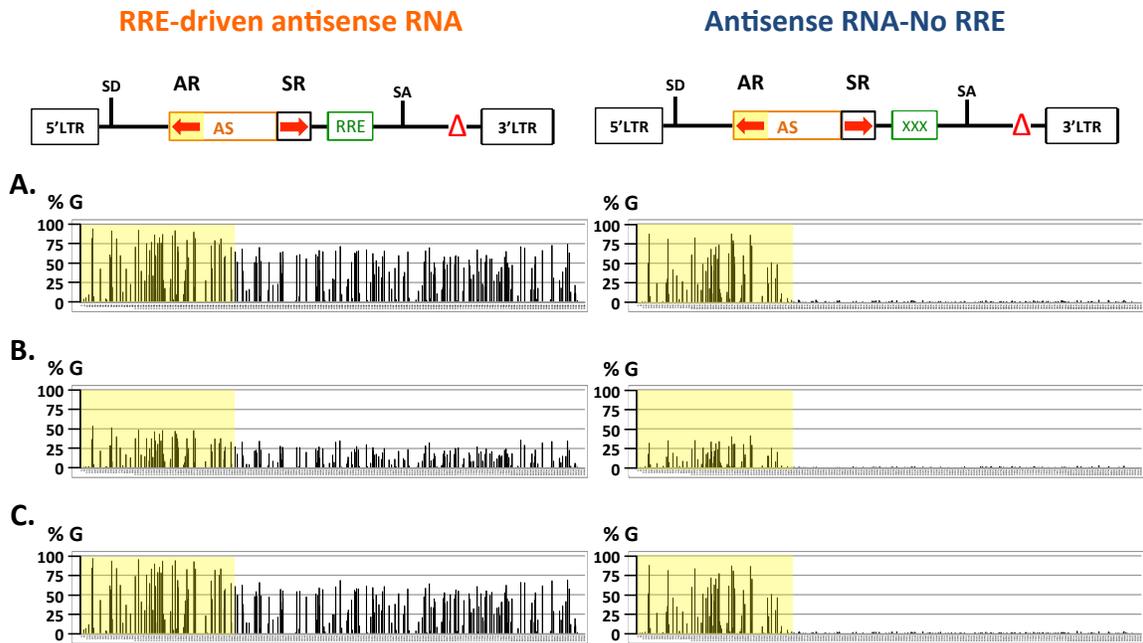


Figure 44. A-to-G changes in the antisense env sequence of the RRE-driven or No RRE antisense RNA (AS-RRE) in the absence or presence of shRNA against ADAR1 or ADAR2

NGS data of the antisense env sequence of AS-RNA isolated from 293T cells expressing an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense env RNA construct or the antisense env RNA-No RRE construct. **(A)** no shRNA treatment **(B)** ADAR1 shRNA **(C)** ADAR2 shRNA. Diagrams of the AS-RNA constructs are shown at the top. The graphs show % A-to-G change (%G) in the antisense env sequence. The yellow box highlights the antisense repeat (AR) region.

We were also interested in investigating potential ADAR editing in the target env region. To this end, we used total RNA to synthesize cDNA and PCR amplify the target region in the HIV mRNAs in an experiment where the vector expressing RRE-driven AS RNA was cotransfected with pNL4-3. The results of an NGS analysis of these amplicons, showed that in RNA from cells lacking shRNA **(Figure 45A)** or expressing shRNA against huADAR2 **(Figure 45C)**, relatively low levels of A-to-G changes (%G) were seen throughout the region

targeted by the AS RNA. Interestingly, the level of A-to-G changes (%G) in the target region was even lower (~1%) when shRNA against huADAR1 was expressed (Figure 45B). Thus very low levels of A-to-G changes (%G) are observed in both the AS RNA and target RNA when ADAR1 reduced and this is associated with an improvement of the inhibition of viral production by the RRE-driven antisense RNA (see Figure 41A).



Figure 45. A-to-G changes in the target env sequence in the absence or presence of shRNA against ADAR1 or ADAR2

NGS data of the target sequence (AST) of AS-RNA isolated from 293T cells expressing an equal molar amount of a pNL4-3 Nef (+) proviral plasmid and the RRE-driven antisense env RNA construct (AS-RNA with RRE). (A) no shRNA treatment (B) ADAR1 shRNA (C) ADAR2 shRNA A diagram of the HIV-1 (NL4-3) genome with the location of the AST sequence (red text box), is shown above. The yellow highlighted region represents the region of the AST that is complementary to the antisense repeat (AR) region of the AS-RNA. % A-to-G changes (%G) are shown in the graphs using a logarithmic scale (\log_{10} scale).

RRE-driven antisense env RNA-mediated inhibition is less effective when cellular huTudor-SN (SND1) is depleted.

Above, we demonstrated that multiple ADAR signature changes can be detected in the panhandle and other regions of the RRE-driven AS RNA vector (see **Figures 30 and 36**). ADAR-edited, inosine-containing double-stranded RNA could potentially be a source of small RNAs (siRNAs or miRNAs) that could mediate the inhibition of viral particle production via an RNAi mechanism. Human Tudor staphylococcal nuclease (huTudor-SN nuclease or SND1) has been shown to cleave ADAR-edited, inosine-containing dsRNAs¹⁵¹. It is possible that huTudor-SN could generate miRNA-sized small RNAs, which could be used by the RNAi machinery. To analyze for the potential involvement of Tudor-SN in the inhibition, we used shRNAs to deplete huTudor-SN in 293T cells (see **Figure 46B**). The cells were then transfected with pNL4-3 and vectors expressing AS RNA (RRE, CTE or with no export element, NoE). As can be seen in **Figure 46A** shRNA-mediated depletion of huTudor-SN (SND1) led to a less effective inhibition in the case of the RRE-driven antisense RNA. In contrast, there were no significant effects of huTudor-SN depletion with the other two AS vectors, where no inhibition was observed whether hu-Tudor-SN was depleted or not. The results thus suggest that huTudor-SN may play a role in AS inhibition, when the AS RNA traffics on the Rev/RRE pathway.

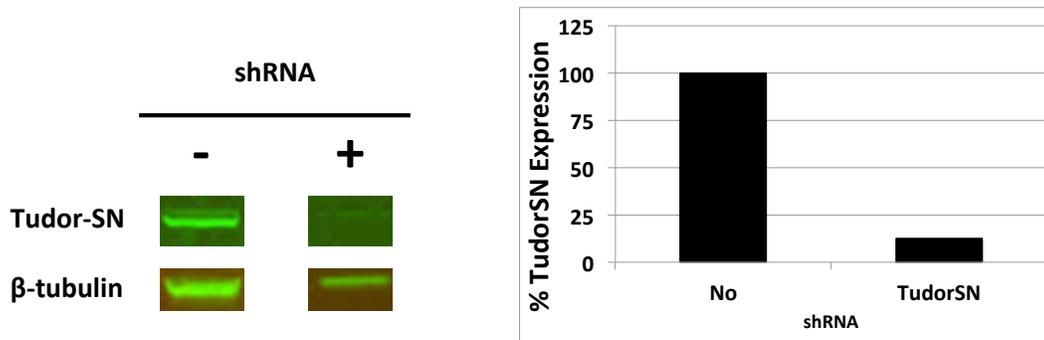
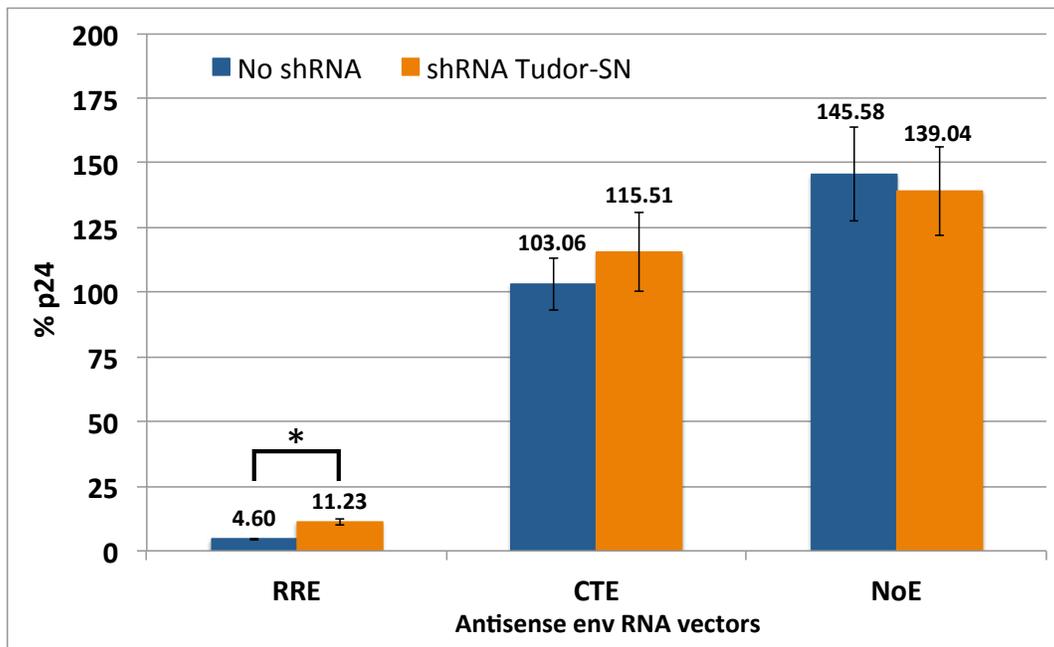


Figure 46. RRE-driven antisense env RNA-mediated inhibition is less effective when the producer cells are treated with shRNA against huTudor-SN

(A) shRNA-treated 293T cells were co-transfected with an equal molar amount of the target NL4-3 proviral plasmid and one of the 3 AS RNA constructs. As a positive control, shRNA-treated 293T cells were transfected with pNL4-3 Nef (+) alone. Similar transfection experiments were also performed in untreated 293T cells for a comparison. Cultured supernatants were collected 48 hours post-transfection and assayed for viral production by p24 ELISA. Cell pellets were also harvested for Western blot analysis. Data are presented as mean \pm standard deviation of seven independent experiments. The % p24 is calculated from a ratio of p24 values in culture supernatant of the 293T cells expressing NL4-3 and the antisense vector to that of a positive control. Statistical significance was assessed using a two-tailed unpaired Student's *t*-test. Error bars indicate Standard deviation means ($n=2$). *P*-values were calculated using the two-tailed unpaired Student's *t*-test with equal variances. **P* < 0.05. **(B)** Western Blot analysis of cells from one of the experiments diminished levels of huTudor-SN expression after shRNA treatment. The quantification of protein expression is shown in the histogram. No error bar shown here since the two duplicated lysates were pooled together.

AS RNA-derived miRNAs are specifically detected in total RNA isolated from cells expressing RRE-driven AS RNA

The experiments described above indicated that the RRE-driven AS RNA was subject to ADAR editing, both in the intra-molecular panhandle and other regions that are, complementary to the HIV target RNA. We also showed that potential involvement of Tudor-SN, which is known to cleave ADAR-edited, inosine containing dsRNAs¹⁵¹, during the inhibition. Processed ADAR-edited dsRNA regions could potentially generate small RNAs that could enter the RNAi machinery as either the miRNAs or the siRNAs^{177,142}. Subsequently, the RNAi machinery could target HIV mRNA species that contain the env target region (9-kb 9 kb and 4 kb HIV-1 mRNAs, see diagram in **Figure 47A**). This notion is supported by the previous study by Ward et. al. that showed that expression of less than molar amounts of AS-RRE RNA lead to efficient reduction of the HIV-1 GagPol production, but does not affect ribosome association of the targeted RNA¹²⁷. These results are consistent with RNAi mediated inhibition. To directly explore this hypothesis, we proceeded to investigate whether the small AS RNA-derived RNAs could be detected during the inhibition of HIV-1 replication, using the Next Generation Sequencing method.

To do this, 293T cells were co-transfected with an equal molar amount of the target NL4-3 Nef (+) proviral plasmid and a vector expressing either AS RNA with the RRE or AS RNA without an export element (NoE). Both of the AS RNA vectors used contained the 287 bp panhandle. As a control, 293T cells were transfected with pNL4-3 Nef (+) alone. Total RNAs were extracted from the

transfected cells 48 hours post transfection. Small RNA libraries were synthesized from total RNA of the transfected cells (see Material and Methods). NGS was then performed on the small RNA libraries using an Illumina miSeq instrument. The NGS data of small RNA libraries were then mapped against the human reference genome (NCBI GRCh38) using the Bowtie algorithm version 1.1.2 developed by Ben Langmead and Cole Trapnell, which is available in GENEIOUS bioinformatics software. The unmapped sequences were subsequently mapped against the HIV-1 reference genome (M19921.2 HIV NL4-3 Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant clone pNL4-3), using the Bowtie algorithm version 1.1.2, in order to detect the small sequences that aligned to the antisense target region. The NGS analysis showed that HIV-1 derived small RNAs (~150-nt or less) constituted 1% of the total small RNA reads when the RRE-AS RNA was present (**Table 5**). Only a small portion of these were HIV-1-derived 19-nt to 23-nt small RNAs. However, significantly more HIV-1 19-nt to 23-nt small RNAs (including AS RNA-derived miRNAs) were detected in the “**RRE**” library, (see **Table 5**). No small RNA aligned to the AS sequences present in the vectors in the case of the pNL4-3 control.

AS-RNAs	All sequences	Aligned to HIV-1 NL4-3 (sequences)	HIV-1 derived miRNAs (sequences)	HIV-1 derived miRNA aligned to AST (sequences)
RRE	2,658,754	29,547 (1.111%)	171 (0.0064%)	73 (0.0027%)
NoE	3,013,411	4,377 (0.145%)	7 (0.0002%)	2 (0.0001%)
NL4-3 only	2,158,583	18,727 (0.868%)	27 (0.001%)	0

Table 5. NGS analysis of small RNAs present in total RNA from cells transfected with pNL4-3 in the presence or absence of AS vectors

The table shows the total number of sma sequences in small RNAs isolated from total RNA from the following transfection conditions: RRE-driven AS env RNA+ pNL4-3 (**RRE**), iAS env RNA with no RRE (**NoE**) +pNL4-3, and pNL4-3 only (NL4-3). After removing the small RNA sequences that mapped to the human genome, the remaining sequences were mapped against the HIV-1 reference genome. The sequences of the 19-23 nt small RNA size class is referred in this table as HIV-1 miRNAs (the AS RNA-derived miRNAs are also included in this population). The total percentages of the different small RNAs are shown in parentheses.

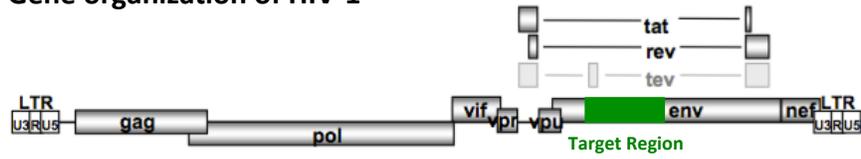
In a separate analysis (**Figure 47**), the HIV-1 small RNAs (**Table 5, third column**) were mapped against the HIV-1 reference genome (M19921.2 HIV NL4-3 Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant clone pNL4-3) in order to detect viral 19-23 nt small RNAs mapping to the target region of the AS RNAs, as shown in **Figure 47A**, highlighted in green. As shown in **Figure 47B** (RNA from an inhibition experiment in the presence of RRE-driven AS RNA), 1,777 of the 29,547 HIV-1 small RNA sequences aligned to the target region. The size distribution of those 1,777 HIV-1 small RNAs showed that only 73 sequences were of the 19-23 nt small RNA size class (highlighted by the yellow box and see size distribution in an inlet). On the other hand, as shown in **Figure 47C** (RNA from an inhibition experiment in the presence of AS RNA lacking the RRE), only 2 out of the 276 sequences represented 19-23 nt small RNAs. As shown in both **Figure 47B** and **Figure 47C**, a small number of the

larger HIV-1 small RNA sequences also aligned to the target region of the AS RNA. As for the control group, as shown in **Figure 47D** (NL4-3 only), there was no viral 19-23 nt small RNA detected in the target region. The smaller sequences detected, which were less than 14-nt, could represent small fragments of the HIV-1 RNAs, resulting from RNA degradation.

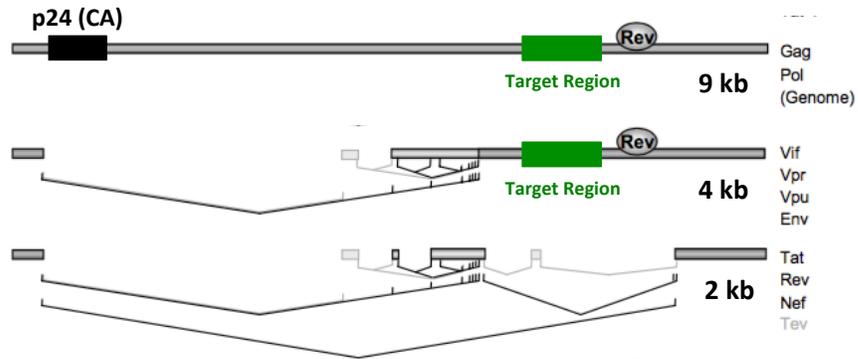
The 73 small RNAs (**Figure 47B**) aligned across the antisense target region. Of the 73 sequences (**Table 5, RRE** and **Figure 47B**), 39 sequences were AS RNA-derived small RNAs, whereas 34 sequences were derived from the sense RNA. Furthermore, 12 of the 39 AS RNA-derived miRNAs showed ADAR-editing, whereas 4 of the 34 sense RNA-derived miRNAs showed ADAR-editing. Interestingly, the remaining of the HIV-1-derived 19-23 nt small RNAs generated during RRE-driven AS RNA-mediated inhibition (**Table 5, RRE**) aligned primarily to the HIV-1 5' and 3'LTRs: 32 sequences aligned to the 5'LTR and the upstream region to the HIV-1 GagPol gene and 66 sequences aligned to the 3'LTR. The detection of 19-23 nt AS RNA-derived miRNAs, generated during the inhibition of HIV-1 replication (**Table 5, RRE** and **Figure 47B**), suggests the possibility that the RNAi machinery acts to inhibit viral protein production as previously suggested^{126,127}.

A.

Gene organization of HIV-1

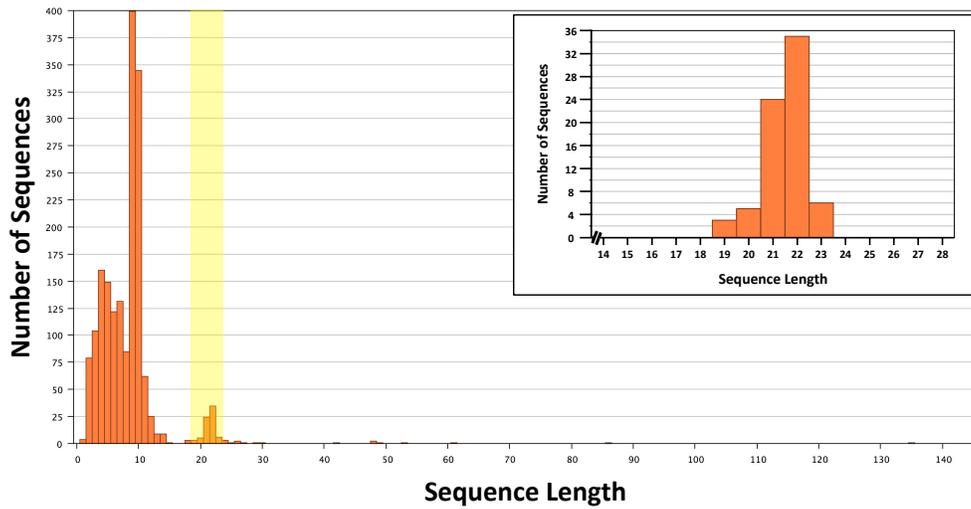


HIV-1 transcripts



B.

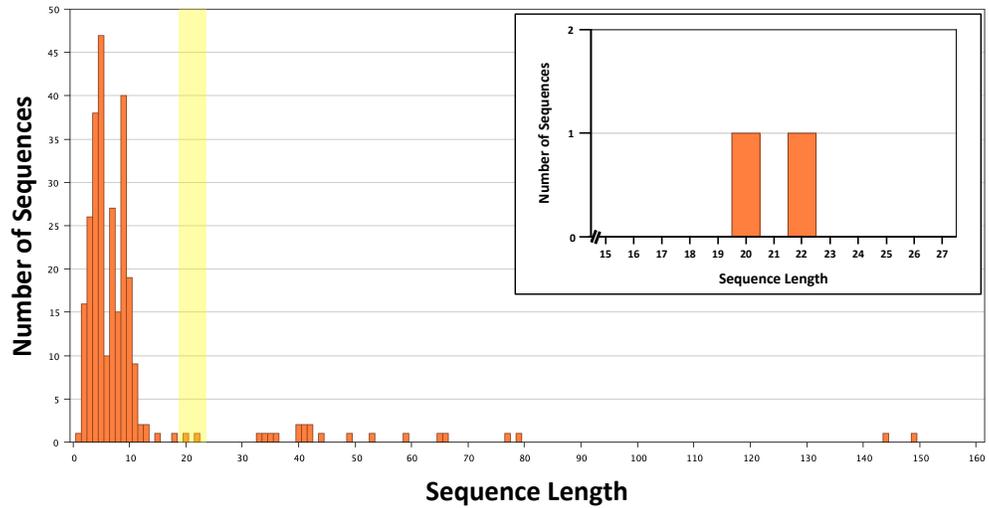
1:1 NL4-3 to RRE-driven antisense env RNA vector



Total number of sequence shown: 1,777

C.

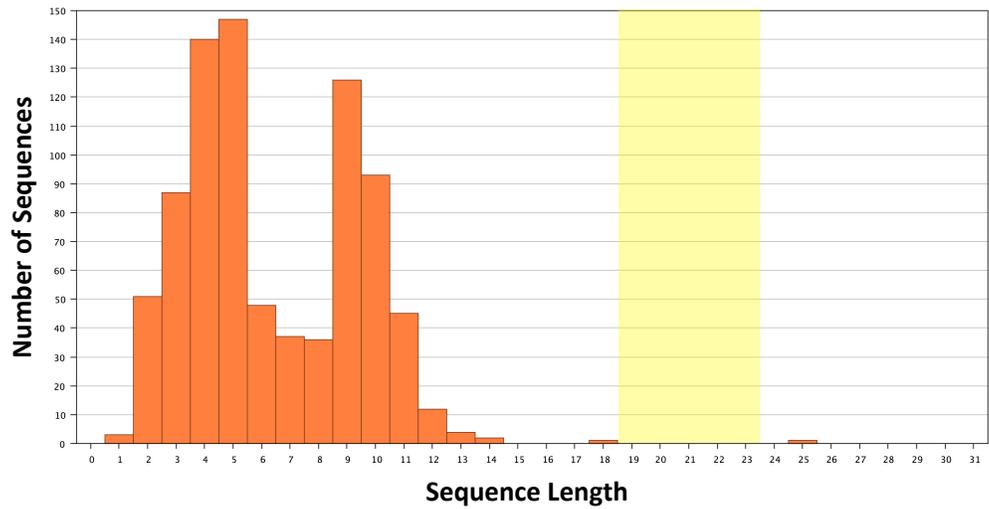
1:1 NL4-3 to antisense env RNA vector with no RRE



Total number of sequence shown: 276

D.

NL4-3 only



Total number of sequence shown: 833

Figure 47. Alignment of small RNAs present in transfected cells to the HIV AS target region

(A) The diagrams show the gene organization of HIV-1 virus and the 3 different size classes of viral mRNAs. The green box highlights the target of the antisense RNA. The black box indicates the ORF of the HIV-1 p24 capsid protein in the 9-kb mRNA. Figures adapted from Kammler et al. *Retrovirology* 2006²³. The histograms show size distribution of the HIV-1 derived small RNAs, aligned to the antisense RNA's target region, in total RNA extracted from 293T cells transfected with: (B) pNL4-3 and the vector expressing AS-RNA containing an RRE (C) pNL4-3 and the vector expressing AS-RNA with no export element, and (D) pNL4-3 only. 293T cells were co-transfected with an equal molar amount of the target NL4-3 Nef (+) proviral plasmid and the AS vectors. As a control, 293T cells were transfected with pNL4-3 Nef (+) alone. The yellow box highlights miRNAs (19-23 nt). The size distributions of the HIV-1 derived miRNAs are shown in inlets. Total RNAs were extracted 48 hours post transfection. Culture supernatants were also collected 48 hours post-transfection and assayed for viral production by p24 ELISA. The p24 ELISA result was consistent with those from previous experiments. The small RNA libraries were prepared from 1 µg of the total RNAs, using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (Set 1 index 1 to 12: NEB E7300S). Prior to the NGS, the libraries were size selected for the small RNA species (10 – 150 nt) with the 6% PAGE gel. The NGS was performed on the 3 small RNA libraries using the Illumina MiSeq system.

Discussion

The analysis of AS RNA-mediated inhibition of HIV points to specific interactions between the ADAR RNA editing machinery and the HIV-1 Rev/RRE pathway.

Our results clearly suggest that there is a specific connection between the HIV-1 Rev/RRE nuclear export pathway and ADAR-mediated RNA editing. This may be related to the fact that HIV Rev proteins have been reported to direct HIV-1 RNAs trafficking through nucleoli¹⁷⁸, where ADAR1 p110 and ADAR2 proteins also localize¹⁷⁹. In this and the previous study by Ward et al. it has also been clearly demonstrated that efficient AS RNA-mediated inhibition of HIV-1 replication is Rev/RRE pathway dependent¹²⁷. In addition, our data have demonstrated that presence of a panhandle in RRE-driven AS env RNA significantly promotes AS inhibition. We have shown that these panhandles are

subject to high levels ADAR editing in the presence of Rev, which increases when ADAR enzymes are over-expressed. This suggests that the RRE-AS RNA may be a target for ADAR editing in the context of the nucleolus, dependent on Rev expression that directs these RNAs to this location.

From the results shown in **Figure 30A** and **Table 1**, multiple ADAR-mediated A-to-G changes are seen in both the antisense and sense repeats of the double-stranded RNA region (the panhandle) present in the AS env RNA vector that traffics through the HIV Rev/RRE-pathway. The level of the A-to-G changes at each edited adenosine (A) position is high. This finding indicates that AS RNA that traffics via the HIV-1 Rev/RRE pathway could encounter the ADAR enzymes in a different way that leads to the extensive editing. Furthermore, RNA editing in the panhandle still occurs without the AS RNA's target.

When the RRE AS-RNA was directed to the NXF1 pathway, using RevM10Nxf1 that binds like Rev, but exports the RNA on the Nxf1 pathway, the number of A-to-G changes in the panhandle region were significantly decreased. (**Figure 30B** and **Table 1**). The observed RNA editing of AS-RNA that traffics on the Nxf1 pathway (**Figure 30B**) could be a result of co-transcriptional editing by ADAR.

Interestingly, the level of editing of the panhandles in cells expressing the AS-RRE RNA alone (shown in **Table 1** and **Figure 30C**) was somewhat lower than that observed the AS RNA vector with no RRE (**Figure 30D**). This vector contains a mutated HIV-1 RRE element with 37 mutations that makes it incapable of interacting with Rev^{49,180}. Since the RNAs expressed from both of

these vectors would be expected to be retained in the nucleus of transfected cells, we would have expected similar ADAR editing levels in the panhandle of both AS RNA vectors. It is thus possible that the AS-RNA with the mutated RRE leads to a different localization in transfected cells, which could affect the editing pattern. However, this would have to be further explored using in situ hybridization.

In the absence of HIV target RNA, only the ds panhandle region would be expected to be subject to ADAR editing. Next generation Sequencing (NGS) analysis confirmed that the ADAR editing takes place only in the panhandle, not in the “single” stranded regions of the AS RNA (**Figure 32A**). However, in the presence of overexpressed wt ADAR1L (p160) the level of ADAR editing increased in the antisense repeat (AR) of the panhandle and extended to certain position along the loop region (**Figure 32B**). It is possible that sequences in the loop region could form short dsRNA regions in the RNA that could be edited by overexpressed ADAR1L p160 in the cytoplasm with time. However, when ADAR2s was overexpressed (**Figure 32C**), only the panhandle was edited, possible reflecting the nuclear localization of this enzyme

Natural levels of ADAR editing benefits Rev/RRE mediated AS RNA inhibition of HIV-1 replication, whereas too much editing impairs it.

Our results demonstrate high levels of ADAR editing in the panhandle regions of the AS RNA that traffic through the HIV-1 Rev/RRE pathway in the absence of the target (pNL4-3). During the inhibition of HIV-1 replication, dsRNA

substrates of the ADAR enzymes could be further generated through base pairing between the sense target region in the HIV-1 mRNAs and the AS RNAs. We have proposed that ADAR-edited dsRNA, for example the edited panhandle, could be cleaved by the nuclease(s) to generate miRNAs, resulting in the inhibition of HIV-1 replication^{141,142,177}. This is why we decided to try to investigate the involvement of ADAR enzymes in the inhibition, by examining whether the inhibition was affected by the overexpression of ADAR enzymes.

There are two functioning isoforms of ADAR1 enzymes in mammalian cells, i.e. ADAR1 p160 and ADAR1 p110. We decided to focus on ADAR1 p160, since this isoform is interferon-inducible and is localized in cytoplasm in response to viral infection^{166–168}. ADAR1 p160 has been shown to suppress PKR, which promotes an interferon induction following viral infection^{181–183}. In the case of ADAR 2, we decided to use one of the isoforms, ADAR2s. This isoform lacks an exon created by an in frame Alu element. Although there are few studies of human ADAR2, there are no clear functional differences between the two isoforms.

Interestingly, an overexpression of all three ADAR enzymes, ADAR1 p160, ADAR2s or ADAR1 p110 (**Figure 34**) specifically led to a very significant reduction in the RRE-driven AS-RNA-mediated inhibition. When three ADAR2s mutants were analyzed (**Figure 35**) the results demonstrated that neither a catalytically inactive, nor a dsRNA binding domain-defective huADAR2s (huADAR2s EAA X 2) had any effect on the inhibition (as shown in **Figure 35B** and **Figure 35D**, respectively). However, as shown in **Figure 35C**, the RRE-

driven AS RNA-mediated inhibition was severely impaired by overexpression of a huADAR2s mutant that lacked the amino terminal domain containing the NLS (Δ 4-72). These findings confirm the involvement of ADAR enzymes during the inhibition and suggest that too much RNA editing activity serves to negatively affect the Rev/RRE AS-RNA mediated inhibition. This could potentially be the result of too much editing, especially in the panhandle region, that makes it difficult to generate functional miRNAs.

In the presence of HIV 1 target mRNAs, ADAR RNA editing is seen throughout the 937-nt antisense AS region, whereas only a few of the target mRNAs are edited.

The Next Generation Sequencing (NGS) data showed that high-level ADAR-editing was observed throughout the antisense region of the RRE-driven AS-RNA in the presence of target HIV mRNAs (**Figure 36A**). This result suggests editing of the AS RNA and the target mRNA can only occur when both AS and target RNAs traffic through the HIV-1 Rev/RRE pathway. Lower level of ADAR-editing were observed in the loop region of the RRE-driven AS-RNA when either huADAR1 p160 or huADAR2s was overexpressed, as shown in **Figure 36B** and **Figure 36C**, respectively. In contrast, ADAR editing in the panhandle region increased. This may indicate less basepairing between AS-RRE RNA and the target mRNA when the ADAR enzymes are overexpressed. It is possible that this reflects nuclear matrix retention of the AS RNA, as has been previously been

shown with heavily edited RNAs^{150,184}. This would make it unavailable for base pairing with the target HIV mRNA.

The results from the NGS analysis of ADAR editing in the AS-RRE RNA (**Figure 36**) and the AST region in the HIV mRNAs (**Figure 37**) and the fact that the a 287 panhandle in the AS RNA is sufficient for efficient inhibition, support the notion that ADAR editing of the panhandle region in the RRE-driven AS RNA serves to generate small RNAs that mediate most of the inhibition. The editing observed in the rest of the AS-RNA that appears to be the result of base pairing between the AS RNA and the AST region in the the HIV mRNAs (**Figure 36**), could lead to generation of additional small RNAs that could contribute to efficient inhibition. The fact that high levels of editing are observed throughout the AS RNA (**Figure 36**), whereas very low levels of editing are observed in the AST region of the HIV mRNA speaks against inhibition being a direct consequence of basepairing and editing of AS-RNA and target mRNAs hybrids. These results suggest that, whereas a relative large fraction of the AS RNA hybridizes with the target mRNA, a very small fraction of the HIV target mRNA ever base pairs with the AS RNA, possibly because of lower amounts of AS RNA compared to target mRNAs.

Lower expression level of huADAR1 enzymes improves the RRE-driven AS RNA-mediated inhibition of HIV-1 replication. The improved inhibition is associated with lower level of A-to-G changes (%G) in both the AS RNA and the target RNA.

Our results demonstrated that an shRNA-mediated reduction of huADAR1 enzymes led to a better RRE-driven AS RNA-mediated inhibition of HIV-1 replication (see **Figure 41A**). Interestingly, Next Generation Sequencing (NGS) data showed that a low level of huADAR1 enzymes decreased the ADAR editing level (%G or % A-to-G changes) in both the RRE-driven AS RNA and its target RNA (see **Figure 44B** and **Figure 45B**). We hypothesize that moderately edited dsRNAs (i.e. a panhandle or dsRNA formed through basepairing between AS RNA and its target RNA) may be processed by endonucleases to form miRNAs or siRNAs^{151,185–189}, which would then mediate efficient inhibition via the RNAi machinery¹⁴². When huADAR enzymes are overexpressed, hyper-editing may dissociate the dsRNA and thus limit the availability of ADAR-edited dsRNA endonuclease substrates.

Human Tudor-SN (SND1) may be involved in the inhibition of HIV-1 mediated by the RRE-driven AS-RNA vector.

Based on the results of our experiments, we hypothesize that the ADAR-edited dsRNA panhandle could give rise to miRNAs, which would inhibit the viral protein production at the level of translation as previously proposed by Ward et al¹²⁷. The ADAR-edited dsRNAs could potentially be cleaved by human Tudor

staphylococcal nuclease (huTudor-SN or SND1), resulting in the generation of small RNAs, miRNAs or siRNAs that could potentially inhibit viral particles production via a RNAi machinery^{142,190}. Proteome analysis using SILAC-based quantitative mass spectrometry has identified TudorSN (SND1) among the 51 candidate proteins that can establish biological interactions with HIV gag. HIV gag interacting RNAi components were TudorSN (SND1), Lin28B, DICER and MOV10¹⁹¹. A study of a crystal structure of Tudor-SN shows that it contains four tandem repeats of staphylococcal nuclease-like domains (SN1–SN4) followed by a Tudor and C-terminal SN domain (SN5)¹⁹². Tudor-SN depends on tandem repeats of SN domains for both RNA binding and cleavage activity. SN3, SN4, Tudor and SN5 domains are shown to form a crescent-shaped structure. A concave basic surface, which is formed by SN3 and SN4 domains, is suggested to be involved in RNA binding, whereas citrate ions are bound at the putative RNase active sites¹⁹². Computational modeling also proposes that Tudor-SN preferentially cleaves RNA containing multiple I:U wobble-paired sequences¹⁹².

Our investigation showed that the depletion of huTudor-SN leads to a less effective AS RRE_RNA-mediated inhibition (**Figure 46**). Tudor-SN has been implicated in processing/degradation of inosine (I)-containing dsRNAs^{151,185,186}. It is therefore possible that ADAR-edited, inosine-containing, double-stranded AS RNA expressed from the RRE-driven antisense RNA vectors can be processed by huTudor-SN to generate AS RNA-derived miRNAs that could cause the inhibition of viral production via the RNAi machinery. However, when either huADAR1 or huADAR2 is overexpressed, extensive editing may lead to the

dissociation of the panhandle and a subsequent reduction of substrate for generating miRNA-like small RNAs by Tudor-SN. Other studies have suggested that human Endonuclease V (huENDO V), a ribonuclease specific for inosine-containing RNA may also be involved in processing of ADAR-edited RNAs^{187–189}. However, shRNA-knockdown of huENDO V did not show any effect on the Rev/RRE-mediated inhibition of HIV replication (data not shown).

19-23 nt AS RNA-derived miRNAs/siRNAs can be specifically detected in RNA from cells where HIV replication is inhibited by RRE-driven AS RNA.

Our experiments have shown that a panhandle in RRE-driven AS RNA promotes the inhibition of HIV-1 replication. ADAR editing of this and other regions in the AS RNA were also demonstrated. Furthermore, shRNA mediated reduction human Tudor staphylococcal nuclease (huTudor-SN or SND1), which is capable of cleaving ADAR-edited dsRNAs to generate miRNAs, in transfected cells were shown to significantly reduce the inhibition. These findings strongly support the involvement of the RNAi mechanisms the inhibition. To further explore these hypothesis, we investigated whether 19-23 nt AS RNA-derived miRNAs were generated during the antisense RNA-mediated inhibition of HIV-1 replication.

Next Generation Sequencing (NGS) data of the small RNA libraries (**Table 5**) revealed that the number of HIV-1-derived small RNAs, were very low. This may be due to the fact that the NGS libraries were prepared from total RNAs that

were not subjected to miRNA enrichment. However, the NGS data demonstrated that significantly more 19-23 nt small RNAs were generated in cells expressing RRE-driven AS env RNA, compared to when cells were expressing an AS RNA lacking the RRE (**Table 5**). There are also more 19-23 nt small RNAs generated during the inhibition mediated by the RRE-driven AS RNA (**Table 5, RRE**) than those generated when there is no inhibition (**Table 5, NL4-3 only**).

When the 19-23 nt viral small RNAs detected in cells expressing RRE-AS RNA were mapped to the HIV genome, 73 of the 171 sequences aligned to the target region of the HIV-1 reference genome (**Table 5, RRE** and **Figure 47B**). In addition, some of those sequences showed ADAR editing. In cells expressing the AS RNA with no RRE, only two sequences aligned to the target region (**Table 5, NoE** and **Figure 47C**). No 19-23 nt small RNA species aligned to the target region in cells expressing pNL4-3 alone (**Table 5, NL4-3 only** and **Figure 47D**). Slightly more than half of the 19-23 nt small RNA species detected in cells expressing RRE-AS RNA (39 sequences) could be identified as AS RNA-derived miRNAs. The remaining 34 19-23 nt sequences that were detected were sense sequences. These could have been derived from Tudor-SN mediated degradation of AS dsRNA duplexes or the RNAi-mediated RNA degradation of targeted HIV mRNAs. The NGS analysis of a small RNA library prepared from miRNAs associated with polyribosomes in transfected cells may improve the detection of AS RNA-derived miRNAs, to further strengthen our hypothesis. Nevertheless, these findings strengthen the hypothesis that ADAR editing and RNAi mechanisms play a significant role in the efficient inhibition that is observed

with AS RNA that traffics on the Rev/RRE pathway.

Interestingly, we also detect several 19-23 nt small RNAs (miRNAs) derived from the HIV-1 5'LTR and 3'LTR of the HIV-1 genome. Several research groups have reported the presence of HIV-1 LTR-derived miRNAs in chronically infected T lymphocytes and their biological function during the infection. For example, in one study approximately 60% of the 47,773 small RNA sequences analyzed were miRNAs, and 125 sequences were those of HIV-1 origin¹⁹³. The HIV-1 TAR-derived miRNA, the earliest known HIV-1 miRNA, has been shown to be the most abundant¹⁹³⁻¹⁹⁶. Others are those derived from the HIV-1 RRE, Nef gene, and LTR^{193,197,198}. The HIV-1-derived miRNAs are capable of modulating cellular and/or viral gene expression, as well as, promoting the viral replication^{199,200}. The HIV-1 TAR-derived miRNA has been shown to suppress the apoptosis in the infected cells²⁰¹. Interestingly, HIV-1 TAR-derived miRNA could be incorporated in to the exosomes of the infected cells²⁰². Recent studies have reported pro-viral effect of novel HIV-1 miRNAs, for instance enhancing viral replication by targeting TATA box region²⁰³ or stimulating TNF α release in human macrophages via the TLR8 signaling pathway contributing to chronic immune activation²⁰⁴. Since there are more LTR-derived miRNAs generated during the RRE-driven antisense RNA-mediated inhibition, it remains to be investigated whether these miRNAs contribute to the inhibition of HIV-1 replication.

Chapter 4 – Concluding remarks and future directions

In this study, we have demonstrated that the presence of a panhandle in the RRE-driven antisense RNA (AS RNA) vectors leads to the most efficient AS RNA-mediated inhibition of HIV viral production and is HIV-1 Rev/RRE pathway-dependent. Removal of the panhandle in the RRE-driven AS env RNA reduces the efficiency of the inhibition, but compared to the AS RNA lacking the RRE, there is still significant inhibition. Surprisingly we also found that the RRE-driven AS RNA vector, which contains only 287 bp panhandle panhandle sequences derived from HIV-1 env, still gives efficient inhibition of viral production. Furthermore, we also showed that a construct with an EGFP panhandle caused inhibition of viral production and EGFP expression from the NL4-3-derived proviral clone, pHIG, which contains the EGFP sequence in viral mRNA. This inhibition was also dependent on trafficking through the HIV Rev/RRE pathway

A previous study by Lu et al. reported that cDNA sequencing analysis of episomal HIV DNAs, isolated from infected SupT1 cells stably expressing AS env RNA, revealed multiple Adenosine-to-Guanosine (A-to-G) changes in the targeted region of the antisense *env* RNA in proviral HIV DNA¹²⁶. They suggested that RNA editing, mediated by ADARs (adenosine deaminases that acts on RNA), in the HIV genome mRNA of cells expressing AS RNA may have resulted in A to G changes during reverse transcription in a new round of infection¹²⁶. Our results provide further evidence for ADAR editing during antisense inhibition.

Our results also show that the “panhandle” regions, which were lacking in the vector employed by Lu et al., showed evidence of high levels of ADAR editing. This was observed both in the presence and absence of target mRNA, and was likely a consequence of ADAR acting on the long double stranded stem that can form intramolecularly within the AS RNA. In contrast, relatively low levels of editing were seen in the double stranded RNA that would be formed by base pairing of the AS RNA with HIV target mRNA. Taken together, our results do not support the hypothesis by Lu et. al. that heavy editing of AS and target RNA hybrids in the nucleus leads to retention by the nuclear matrix. This hypothesis is further contradicted by the fact that edited RNA is found in the cytoplasm and that the target mRNA still associates with ribosomes²⁰⁵. Also, only a small fraction of the target mRNA was edited and overexpression of wild-type huADAR1 and huADAR2s led to relief of the RRE-driven AS RNA-mediated inhibition, rather than the enhancement that might have been expected if the Lu et al. hypothesis was true. Thus, we conclude that another mechanism must be responsible for the efficient inhibition that is seen when the AS RNA trafficked on the Rev/RRE pathway.

The previous AS RNA vector from our lab study¹²⁷ presented some evidence that suggested that RNAi mechanisms may be involved in the inhibition. Although the targeted GagPol mRNAs were exported to the cytoplasm and associated with polyribosome complexes, the levels of HIV-1 capsid protein, p24 made from this mRNA were greatly reduced. We have hypothesized that ADAR-edited dsRNAs could be cleaved by the human Tudor staphylococcal

nuclease (huTudor-SN or SND1), resulting in the generation of either miRNAs or siRNAs that could inhibit viral particle production via a RNAi machinery¹⁴². Indeed, our experiments have shown that the depletion of huTudor-SN led to significantly less effective AS RNA-mediated inhibition. Additionally, AS RNA-derived miRNAs and siRNAs were detected in cells expressing the AS-RNA. Although only a small number of such RNAs were detected, the numbers were significantly increased when the AS-RRE RNA was expressed. To date, all of the observations from our study and those from the previous ones are consistent with the involvement of both the ADAR RNA editing and the RNAi machinery in AS RNA-mediated inhibition of HIV-1 replication when the RNA traffics on the HIV-1 Rev/RRE pathway. A model for the proposed mechanism is shown in **Figure 48**.

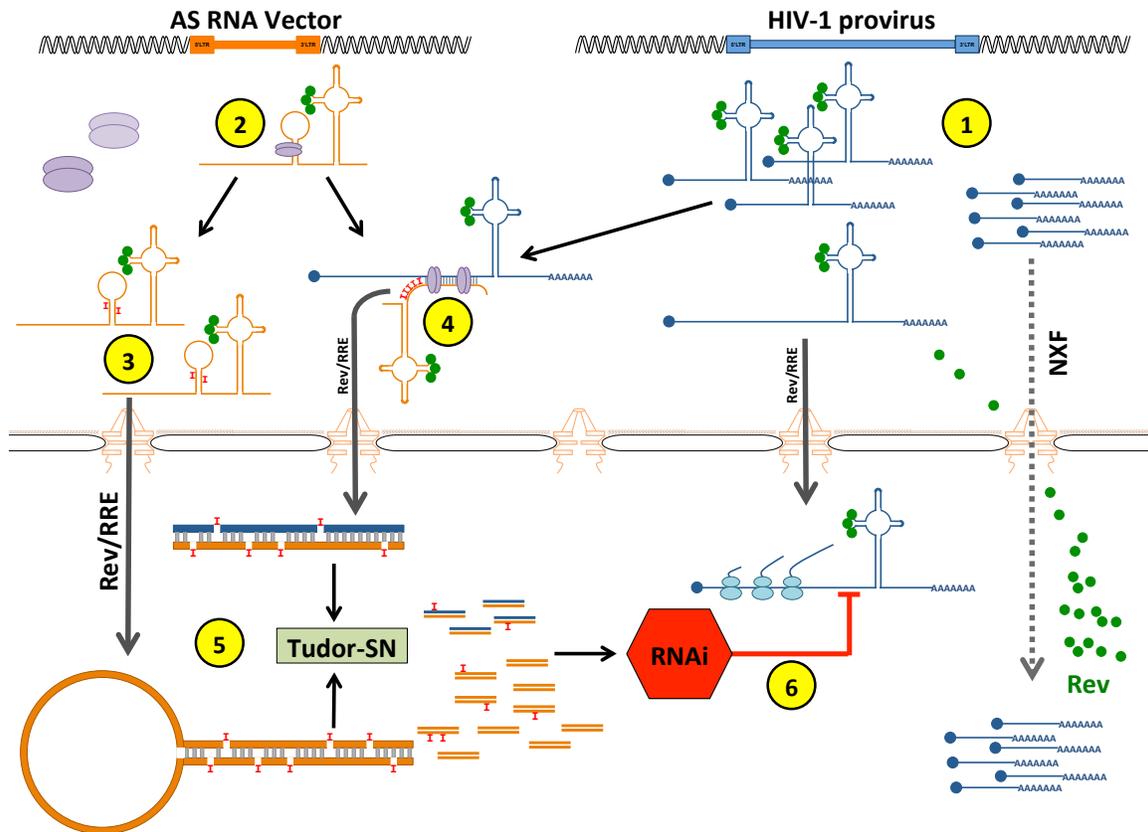


Figure 48: The diagram demonstrates the hypothetical mechanism of the Rev/RRE-pathway dependent antisense RNA (AS RNA)-mediated inhibition of HIV-1 replication

(1) The transcription of HIV-1 provirus yields the 9-kb intron-retaining unspliced mRNAs, 4-kb intron-retaining singly spliced, and the 2-kb multiply spliced mRNAs. The 2-kb multiply spliced HIV-1 transcripts are exported to the cytoplasm via the NXF1 pathway. The HIV-1 Rev proteins are translated from the 2-kb HIV-1 Rev transcripts. The HIV-1 Rev proteins are imported into the nucleus, where they bind to the RRE element of the intron-retaining mRNAs in order to facilitate the nuclear export. (2) The nuclear localized HIV-1 Rev protein (green circles) bind to the RRE element of the AS RNA co-transcriptionally to facilitate the nuclear export of the RNA³⁶⁻³⁹. The HIV-1 Rev/RRE pathway would also allow AS RNA to be edited by ADAR1 p110 and ADAR2 (purple ovals) localized in the nucleoli^{178,179}. (3) Since the HIV-1 Rev/RRE pathway can overcome nuclear retention, the AS RNA, containing the edited panhandle, is exported to the cytoplasm¹⁵⁰. Only the edited panhandle and the loop region of the AS RNA are shown in the cytoplasm. (4) AS RNA (orange) vector can also hybridize with the target region (AST) in HIV-1 transcripts (blue) to create long dsRNA. ADAR enzymes can also edit the dsRNA. Ultimately, the edited dsRNA is exported to the cytoplasm through the HIV-1 Rev/RRE pathway. Only the edited (orange and blue) dsRNA portion is shown in the cytoplasm. (5) In the cytoplasm, the human Tudor staphylococcal nuclease (Tudor-SN) cleaves the ADAR-edited dsRNAs and generates AS RNA-derived miRNAs, which could be incorporated into the RNAi machinery. (6) The RNAi machinery inhibits translation of HIV-1 mRNAs containing sequences complementary to the AS RNA

RRE-driven antisense RNA (AS RNA) Lentivirus vectors have been shown to be the effective tool for inhibiting HIV-1 replication. Therapeutic use of an RRE-driven AS RNA vector (VRX496), has been tested in HIV-1 infected individuals in clinical trials (see www.clinicaltrials.gov, number NCT00295477¹²⁸) and has shown some promising results. The modified CD4 T cells transduced with the VRX496 have a half-life of approximately 5 weeks in the blood in the HIV-1 infected individuals. This long-term prevention of viral replication provided by the VRX496-transduced CD4 T cells may enhance the effect of the anti-retroviral regimens. Interestingly, passaging of HIV-1 in the SupT1 cells transduced with the RRE-driven AS RNA vector was shown to result in non-infectious viruses, demonstrating deletions or mutations in the sequence targeted by the AS RNA¹²⁶. It would be interesting to perform similar passaging experiment with our RRE-driven AS RNA vector that contains the panhandle, to analyze if similarly leads to the generation of defective viruses and to compare the inhibition in a situation where the AS RNA is stably expressed. Such experiments should include studies to determine if HIV-1 can develop resistance to the AS RNA, as was previously shown with RevM10²⁰⁶.

However, the use of an RRE-driven AS RNA vector as a potential prevention measure against the HIV-1 transmission in healthy individuals remains to be explored and may be problematic. Additional studies have to be conducted in order to establish the safety of using AS RNA construct in healthy individuals, since several adverse effects of retrovirus vectors have been reported in the past. Although a much safer the third generation lentiviral vector

is used in the VRX496 clinical trial¹²⁸, it is not entirely exempt of risks, since the viral proviral genome still integrates rather randomly into the host genome. The integrated vector could disturb the function of host cells by insertional mutagenesis. Alternatively, the lentiviral vector could be replaced by an adeno-associated virus (AAV) vector, which can infect both dividing and quiescent cells and persist episomally, largely avoiding insertional mutagenesis²⁰⁷⁻²¹⁰. The successful use of an AAV vector in a clinical trial designed to treat Leber's Congenital Amaurosis has been recently reported²¹¹. It is possible that AAV vectors could be engineered to express AS RNA in a Rev/RRE dependent manner.

Our experiments clearly demonstrated that the dsRNA panhandle, which contained the 287-nt inverted repeat derived from the antisense sequence of the HIV-1 env sequence, significantly enhanced the AS RNA-mediated inhibition of the HIV-1 replication. The panhandle may stabilize the AS RNA vector in the transduced cells, leading to increased levels of AS-RNA, which could enhance the effects. This could be further analyzed by comparing the amounts of AS RNA in transfected cells with and without the panhandle. ADAR editing creates moderately edited panhandles that could be a major source for AS RNA-derived miRNA/siRNA production (**Figure 48: step 1, 2 and 4**). However, miRNA/siRNA could also be derived by hybridization of the AS RNA to the mRNA target. Hyper-edited panhandles would be expected to dissociate, allowing the 650-nt single stranded loop region to hybridize with the target region. The resulted dsRNA could then be edited and turned into the AS RNA-derived

miRNAs/siRNAs (**Figure 48: step 3 and 4**). Therefore, it would be interesting to further investigate how each of these mechanisms contribute to the production of the AS RNA-derived small RNAs by further enriching for small RNAs and mapping them to the HIV genome.

As shown in **Figure 48**, ADAR editing and efficient inhibition is related to trafficking of the AS RNA on the Rev/RRE pathway. However, this RNA is still able to rather efficiently inhibit HIV mRNAs redirected to the CTE pathway¹²⁷. This suggests that inhibition may not be dependent on the AS RNA and the target mRNA trafficking on the same pathway when the AS RNA contains a panhandle. In this regard, we have also demonstrated that an AS RNA vector with an EGFP sequence panhandle effectively inhibits p24 production from pHIG, in an HIV Rev/RRE pathway dependent manner. pHIG is a laboratory clone of HIV-1 containing an EGFP sequence that is expressed in both in a Rev/RRE dependent and independent way. It would be interesting to see whether the panhandle could be modified to target the transcripts of other cellular proteins expressed from mRNAs not expressed as part of HIV and that exclusively traffic through the Nxf1 pathway. Additional experiments may also be necessary to establish the optimized number of ADAR targets (A:U pairs) and the length of the dsRNA panhandle required for effective inhibition. For example, a dsRNA panhandle, containing an A:U pair for every 20-21 basepair along the stem, could theoretically be used to deliver multiple siRNAs.

Our results suggest that ADAR-edited inosine (I) containing dsRNA could be turned into miRNAs/siRNAs by the Tudor-SN in cytoplasm, as shown in

Figure 48. Only moderately edited dsRNAs could be turned into miRNAs/siRNAs, if this occurs at edited inosines. In this case, it would be expected that hyper-editing by ADAR enzymes would diminish the population of moderately edited dsRNAs, as shown in **Figure 49**. This could be one explanation for the observed reduction in inhibition when ADAR enzymes are over-expressed. Another potential explanation is that over-expression of ADAR proteins leads to the squelching of factors needed for processing of edited RNAs. We have tried to use CRIPR/Cas9 to develop the 293T cell lines lacking catalytically active ADAR enzymes: ADAR1 p160, ADAR1 p110, and ADAR2. This would serve to demonstrate that the AS RNA effects that we see are dependent on editing. However, we failed to generate 293T clones that lacked ADAR expression. The genomic DNA sequencing of generated clones showed that at least one copy of the ADAR1 gene remained intact. It has been shown that the ADAR1-deletion in mice is embryonically lethal. Furthermore, ADAR2-deficiency in mice results in early postnatal death²¹². Since both ADAR genes appear to be essential to the normal cellular function, this could explain the unsuccessful development of the ADAR-deficient 293T clones. Alternatively, conditional ADAR knockdown 293T cell lines could be developed for the future investigations.

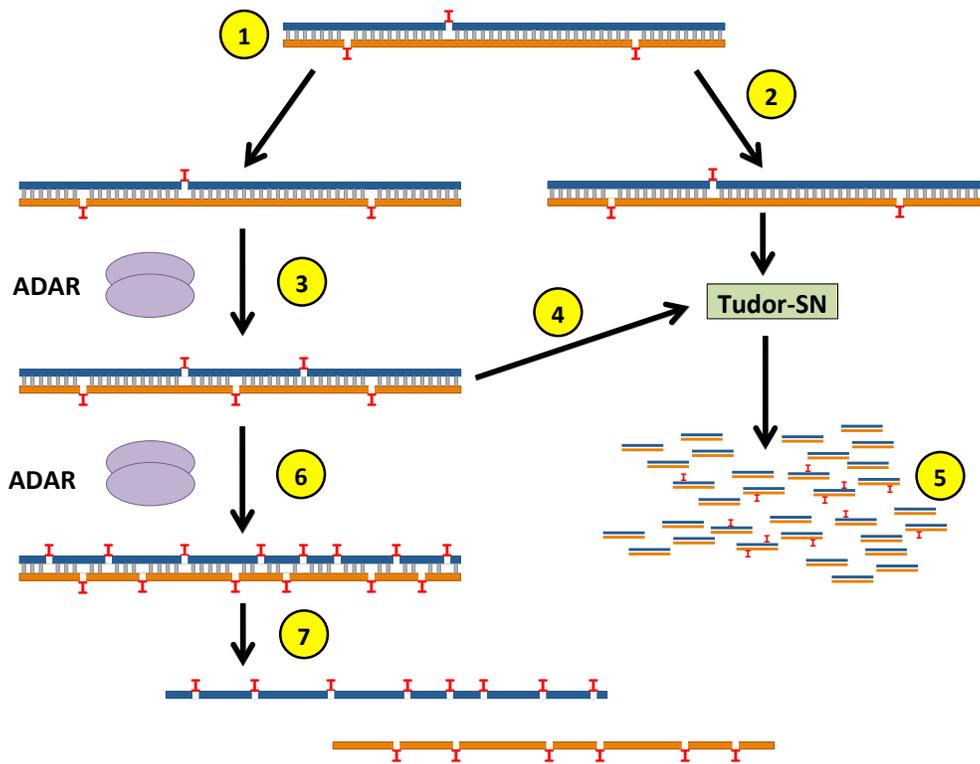


Figure 49: The diagram demonstrates the hypothetical mechanism explaining that ADAR overexpression impairs the RRE-driven antisense RNA-mediated inhibition of HIV-1 replication

During the inhibition, the double stranded RNA (dsRNA) comes from the panhandle of the AS RNAs or the hybridization between the antisense RNA and the target sense RNA (AsT), as shown in **Figure 48B**. **(1)** The moderately ADAR edited dsRNA, once exported to the cytoplasm, could be cleaved by **(2)** Tudor staphylococcal nuclease (Tudor-SN) to generate the **(5)** the AS RNA-derived miRNAs that mediated the inhibition of viral protein production. **(3)** The ADAR enzymes (i.e. ADAR1 and/or ADAR2) could further edit the dsRNA containing I:U mismatches. **(4)** The edited dsRNA, containing more I:U mismatches, could be exported to the cytoplasm to be processed by Tudor-SN. However, **(6)** hyper-editing to the dsRNA leads to the presence of more I:U mismatches, which would ultimately disrupt the dsRNA **(7)**. Therefore, the overexpression of either ADAR1 or ADAR2 could result in the depletion of the moderately edited dsRNA substrate for generating the AS RNA-derived miRNAs.

Our experiments have suggested that both ADAR-edited panhandles and dsRNA duplexes of the target and the AS RNAs can be processed into miRNAs/siRNAs. It will be interesting to investigate whether other nucleases could also be involved in the inhibition. For instance, in addition to Tudor-SN, human Endonuclease V (huENDO V) has been shown to be involved in

processing of ADAR-edited RNAs^{187–189}. huENDO V has been shown to cleave the hyper-edited inosine-containing ssRNAs, whereas the Tudor-SN process the ADAR-edited dsRNAs²¹³. Additionally, although only a DICER protein lacking the amino-terminal helicase domain has the ability to process long dsRNA substrates²¹⁴, the normal isoform of DICER may assist Tudor-SN by processing shorter dsRNA substrates generated by this enzyme. The combination of these three nucleases could thus potentially efficiently generate AS RNA-derived miRNAs/siRNAs. Since our investigation has detected only a few miRNAs/siRNAs, we wish to improve the detection in the future by using the PhotoActivatable Ribonucleoside-enhanced Crosslinking and ImmunoPrecipitation (PAR-CLIP), which helps detecting the AS RNA-derived miRNAs associated with the RISC complex^{215–218}.

Future experiments should expand the knowledge of the mechanism of the antisense RNA-mediated inhibition of HIV-1 replication. We can also gain more understanding of the complicated cellular mechanisms behind the inhibition, for example the potential relationships between ADAR editing and RNA interference during the antisense RNA-mediated inhibition. Furthermore, the more information that we gain additional investigations would improve the strategies for using antisense RNA gene therapy against either HIV-1 or other viral infections and diseases. Finally, the new knowledge gained from the proposed experiments, would likely lead to greater understanding of the biology of HIV-1 virus and the regulation that may be achieved through expression of natural HIV AS RNA.

Antisense RNA-based therapy for non-infectious diseases

Antisense (AS) RNAs have also been used in the development of treatments for non-infectious diseases. Weiss et al have in a review described several potential therapeutic applications of AS RNAs in cancer therapy, for instance AS RNAs targeting Insulin-like growth factor receptor type I (IGF-IR), Calmodulin, Gastrin, Vascular Endothelial Growth Factor (VEGF), and the BCR-ABL gene product of the Philadelphia chromosome, p210²¹⁹. AS RNA approaches have also been explored for treatment of cardiovascular diseases and neurologic disorders, for example AS RNAs targeting angiotensin and dopamine receptors²¹⁹. AS RNAs targeting mRNAs of these proteins have been reported to have some effect in tissue culture and animal models of the diseases²¹⁹.

Though AS RNA-based therapy has been shown to be a potentially promising therapeutic approach, the mechanism of AS RNA action is not well understood²¹⁹. Recently, several oligonucleotide-based drugs regulating disease-causing small RNAs, mRNAs or non-coding RNAs have been tested in clinical trials²²⁰. However, these are not long antisense RNAs, but short Antisense Oligonucleotides (ASOs) or other modified forms of ASOs, for example a Locked Nucleic Acid (LNA) inhibiting miR-122 in HCV-infected patients, a 2'-O-methyl phosphorothioate or Phosphorodiamidate Morpholino Oligomer (PMO) treating Duchenne Muscular Dystrophy (DMD), and N-acetylgalactosamine (GalNAc)-conjugated ASOs inhibiting Nonalcoholic Steatohepatitis-causing miR-103/107²²⁰.

To further complicate matters, ectopic expression of AS RNAs could potentially interfere with naturally occurring antisense RNAs, which may be involved in the regulation of expression of the target genes²²¹. Interestingly, some naturally occurring antisense RNAs could also be the cause of certain disorders. In the case of neuromuscular disorder spinal muscular atrophy (SMA), SMN-AS1 is a long non-coding RNA (lncRNA) transcribed from the antisense strand of the survival motor neuron (SMN) gene. This RNA has been shown to transcriptionally repress SMN expression by recruiting the epigenetic Polycomb Repressive Complex-2²²². Antisense Oligonucleotides (ASOs) have been used to degrade the SMN-AS1 lncRNA in order to increase the expression of the SMN protein²²². This may potentially be developed to treat SMA in the future. The understanding of emerging lncRNAs and their biological functions is likely essential for future use of AS RNA-based therapy for non-infectious diseases in many cases. Finally, other problems associated with AS RNA-based therapy are how to limit the uptake and delivery to the proper target tissues and how to achieve efficient expression in the intended target cells²¹⁹. When such restrictions have been addressed, many more diseases or disorders could potentially be treated by AS RNA-based therapy.

Chapter 5 – MATERIALS AND METHODS

A. Construction of antisense plasmids and nomenclatures:

All plasmids used in the study were indexed and archived as numbers in the form of pHRXXXX. See **Table 6** for a complete list of the plasmids used in the antisense experiments. The design of an RRE-driven antisense env RNA vector (pHR3476), a CTE--driven antisense env RNA vector (pHR3477), an antisense env RNA with no exporting element (NoE) (pHR3478), a CTE-driven antisense EGFP RNA vector (pHR3756), and an antisense EGFP RNA with no exporting element (NoE) (pHR3757), as well as the six first plasmids in Table 1 have been previously described¹²⁷. The RRE-driven antisense EGFP RNA vector (pHR3755), which was used in the previous study¹²⁷, was corrected for a missing Nsil site and a stop codon immediately upstream of the antisense EGFP sequence and was indexed as pHR5284. In order to create “no panhandle” antisense env RNA vectors, a 287-bp inverted repeat in pHR3476, pHR3477, and pHR3478 was removed by digesting the antisense vectors with BsaBI and SpeI. The digested vectors were, then repaired with the “Klenow fragment” polymerase and re-ligated to create the RRE-driven antisense env RNA no panhandle (pHR 5151), the CTE-driven antisense env RNA no panhandle (pHR 5153), and the antisense env RNA with no exporting element (NoE) no panhandle (pHR 5155).

pHRXXXX	Plasmids used in this study
pHR30	pCMV HIV-1 Rev wt
pHR136	pCMV HIV-1 Tat wt
pHR1145	pNL4-3 Nef (+)
pHR1831	pCMV SEAP
pHR2155	pCMV RevM10-NXF (TAP)
pHR2416	pCMV NXT1
pHR3476	RRE-driven antisense env RNA vector with an env panhandle
pHR3477	CTE-driven antisense env RNA vector with an env panhandle
pHR3478	Antisense env RNA with no exporting element vector (NoE) with an env panhandle
pHR5284	RRE-driven antisense EGFP RNA vector
pHR3756	CTE-driven antisense EGFP RNA vector
pHR3757	Antisense EGFP RNA vector with no exporting element (NoE)
pHR4039	pCS ADAR1 full length (ADAR1L) FLIS (N-terminal Flag-tagged and C-terminal HIS tagged)
pHR4044	pCS ADAR2s FLIS (N-terminal Flag-tagged and C-terminal HIS tagged)
pHR4138	pHIG or pNL4-3.HSA.R+.E- + HSA-IRES-EGFP
pHR4921	pCDNA3.1 ADAR2s wild-type
pHR4922	pCDNA3.1 ADAR1 M296 to End p110
pHR4923	pCDNA3.1 ADAR2s (EAA) x 2
pHR4924	pCDNA3.1 ADAR2s Δ 4 to 72
pHR4925	pCDNA3.1 ADAR2s Catalytically Inactive
pHR5151	RRE-driven antisense env RNA vector with no env panhandle
pHR5153	CTE-driven antisense env RNA vector with no env panhandle
pHR5155	Antisense env RNA vector with no exporting element (NoE) with no env panhandle
pHR5247	RRE-driven antisense EGFP RNA vector with an env panhandle
pHR5249	RRE-driven antisense EGFP RNA vector with the EGFP panhandle
pHR5253	CTE-driven antisense EGFP RNA vector with an env panhandle
pHR5255	CTE-driven antisense EGFP RNA vector with the EGFP panhandle
pHR5259	Antisense EGFP RNA vector with no exporting element (NoE) with an env panhandle
pHR5261	Antisense EGFP RNA vector with no exporting element (NoE) with the EGFP panhandle

Table 6: A list of proviral clones, antisense RNA vectors, and expression plasmids used in this study.

In order to construct the antisense EGFP RNA vectors with an env panhandle, an NsiI/NotI double digested gene block (gBlock) env panhandle (see sequence below), containing a 287-bp DNA sequence complementary to a region in the antisense region, was ligated into pHR5284, pHR3756, and pHR3757, which were also digested with both NsiI and NotI. The resulting antisense vectors were an RRE-driven antisense EGFP RNA vector with an env panhandle (pHR5247), a CTE- driven antisense EGFP RNA vector with an env panhandle (pHR5253), and an antisense EGFP RNA vector with no exporting element (NoE) and an env panhandle (pHR5259).

To create the antisense EGFP RNA vector with the EGFP panhandle, a SpeI/BsaBI double digested gene block (gBlock) EGFP panhandle (see the sequence below) was ligated into pHR5284, pHR3756, and pHR3757, which were also digested with both the SpeI and the BsaBI. The resulted antisense vectors were an RRE-driven antisense EGFP RNA vector with the EGFP panhandle (pHR5249), a CTE-driven antisense EGFP RNA vector with the EGFP panhandle (pHR5255), and an antisense EGFP RNA with no exporting element (NoE) with the EGFP panhandle (pHR5261).

gBlock env panhandle: NsiI- stop codon TGA-NcoI-[287-nt env antisense repeat]-NotI

```
ACGTACGTATGCATTGACCATGGGTCCACTGATGGGAGGGGCATACATTGCTTTTCTACTTCCTGCCACA
TGTTTATAAATTGTTTTATTCTGCATGGGAGTGTGATTGTGTCACTTCCTTCAGTGTTATTTGACCCTTCA
GTACTCCAAGTACTATTAAACCAAGTACTATTAAACAGTTGTGTTGAATTACAGTAGAAAAATTCCCCTCC
ACAATTAATACTGTGCGTTACAATTTCTGGGTCCCCTCTGAGGATTGCTTAAAGATTATTGTTTTATTAT
TTCCAAATTGTTCTCTTAATTTGCTAGCGGCCGCACGTACGT
```

gBlock EGFP panhandle: SpeI-TC-[Last 287-nt Sense EGFP]-AAATTA-BsaBI

```
ACGTACGTACTAGTTCACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAA
GGTGAACCTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCATCGGCGACGGCCCCGTGCTGCTGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAA
GACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAAAAATTAGATGTTTCATCACGTACGT
```

B. Proviral clones used in the study:

Two different proviral clones were used in the study as targets for the antisense vector (pNL4-3 Nef (+) (pHR1145) and pHIG (pHR4138). The former was obtained from Dr.Teh Jeang NIAID²²³. pHIG was derived from pNL4-3.HSA.R+.E-²²⁴. It is an HIV-1 proviral clone which has the murine heat stable antigen followed by an internal ribosome entry site (IRES)-enhanced green fluorescent protein (eGFP) sequence inserted into the XhoI restriction site in the Nef region of NL4-3¹⁶⁴. pHIG expresses EGFP (the enhanced green fluorescent protein), and HSA (the murine heat stable antigen). Thus, EGFP and HSA protein expressions can be used, in addition to HIV-1 p24 protein, to monitor HIV-1 virus production.

C. shRNA in vitro transcription and shRNA treatment of 293T cells:

The ssDNA oligonucleotides shown below were purchased from IDT. Oligos 2742 and 2743 and oligos 3039 and 3040 were annealed to each other and subsequently used as templates for transcription of shRNA against huADAR2 and huADAR1, respectively. Oligos 3231 and 3232 and oligos 3233 and 3234 were annealed to each other and subsequently used as templates for transcription of shRNA against Tudor-SN as described below:

HamRek Oligo 2742:

CATAGAACTTAATACGACTCACTATAAGCCTGGTTTGCAGTACACAAGTTCTCTGTGTACTGCAAACCAGG
CTTT

HamRek Oligo 2743:

AAAGCCTGGTTTGCAGTACACAGAGAACTTGTGTACTGCAAACCAGGCTTATAGTGAGTCGTATTAAGTTC
TATG

HamRek Oligo 3039:

CATAGAACTTAATACGACTCACTATA**GAGTTCCTCACCTGTAATATACAAGTTCTCTGTATATTACAGGTG**
AGGAACTCTT

HamRek Oligo 3040:

AAGAGTTCCTCACCTGTAATATACAGAGAACTTGTATATTACAGGTGAGGAACTCTATAGTGAGTCGTATT
AAGTTCTATG

HamRek Oligo 3231:

TAATACGACTCACTATA**AAGGCATGAGAGCTAATAATCAAGTTCTCTGATTATTAGCTCTCATGCCTTTT**

HamRek Oligo 3232:

AAAAGGCATGAGAGCTAATAATCAGAGAACTTGATTATTAGCTCTCATGCCTTTTATAGTGAGTCGTATTA

HamRek Oligo 3233:

TAATACGACTCACTATA**AAGGAGCGATCTGCTAGCTACAAGTTCTCTGTAGCTAGCAGATCGCTCCTTTT**

HamRek Oligo 3234:

AAAAGGAGCGATCTGCTAGCTACAGAGAACTTGTAGCTAGCAGATCGCTCCTTTTATAGTGAGTCGTATTA

The T7 RNA polymerase promoter is underlined. The ADAR siRNA sequences, in blue, are adapted from Jayan GC and Casey JL *J Virol.* 2002¹⁷². The Tudor-SN siRNA sequences, in blue, are from Välineva et al. *J. Biol. Chem.* 2005²²⁵. A sequence of the loop of the shRNA is in red.

Annealing of shRNA templates:

dsDNA templates for making shRNA Tudor-SN #1 and #2 were obtained by annealing an equal molar amount of oligo 3231 and an oligo 3232 and an equal amount of oligo 3233 and oligo 3234, respectively. 5 µl of 10x Annealing Buffer (1X TE 500 mM NaCl) was added into the mixture of the two oligos. The final volume was adjusted to 50 µl with 1X TE. The oligos were annealed in the following annealing cycle, using a thermocycler.

Annealing cycle:

- Step 1: 95°C for 5 min
- Step 2: Decreasing 1°C/cycle/1 minute until T_m
- Step 3: Incubate the oligos at T_m for 30 minutes
- Step 4: Decreasing 1°C/cycle/1 minute
- Step 5: Incubate the oligos at 25°C for 10 minutes
- Step 6: Hold at 4°C
- Step 7: End the cycles.

Note: T_m of shRNA huADAR1 and huADAR2 templates are 66.5°C and 68.9°C, respectively. T_m of shRNA Tudor-SN #1 and #2 templates are 67.7°C and 69.6°C, respectively.

In vitro transcription of shRNA:

For a 500- μ l reaction volume, 100 μ l of 5X reaction buffer (400 mM Tris, 0.05% Triton X-100, 10 mM spermidine, and 50 mM $MgCl_2$), 100 μ l of 200 mM DTT, 100 μ l of 10 mM rNTPs, 12.5 μ l of T7 RNA Polymerase (provided by Stuart F.J. Le Grice, National Cancer Institute), 10 μ l of 40 U/ μ l RNasin, 10 μ l of an annealed shRNA template was combined and the final volume was adjusted to 500 μ l with nuclease-free water. The reaction was then incubated at 37°C for 2-3 hours. To degrade the dsDNA template, 5 μ l 0.1 M $CaCl_2$ and 1 μ l RQ DNase I (Promega 1 U/ μ l) were added into the sample and the sample was incubated at 37°C for a further 30 minutes. The shRNA was then purified using the CHROMA SPIN-30 DEPC- H_2O column (Clontech Cat. No. 636087) following the manufacturer's instruction.

Treat 293T cells with shRNA:

Equal molar amount of two shRNAs (i.e. 5 μ g of each shRNA) targeting Tudor-SN (shRNA #1 and #2) were transfected into 10 million 293T/17 cells by using the calcium phosphate method as previously described²²⁶. 48 hours after transfection, the shRNA-treated cells were trypsinized. 6 million shRNA-treated cells were re-plated on 10-cm tissue culture plates for further transfection.

D. Cell lines and transfections:

293T/17 cells were maintained in Iscove's minimal essential medium supplement with 10% bovine calf serum and 0.1% gentamicin. Transient transfections were performed using the calcium phosphate method as previously described²²⁷.

E. p24 ELISA:

Tissue culture supernatants from transfected cells were collected at 48 or 60 hours after transfection. The supernatants were then centrifuged at 5000 rpm for 3 minutes in order to remove the transfected cells and cell debris. The supernatants were then diluted 10- to 10000-fold before application to enzyme-linked immunosorbent assay (ELISA) plates. The ELISA was performed using a p24 monoclonal antibody (NIH AIDS Reagent Repository catalog no. 1513) and pooled human anti-HIV-1 immunoglobulin G (NIH AIDS Reagent Repository catalog no. 3957) following a protocol developed by Dr. Bruce Chesebro (National Institute of Allergy and Infectious Diseases, Rocky Mountain Laboratories)²²⁸.

F. Western blot analysis:

The medium was aspirated off of the 10-cm tissue culturing plate containing approximately 20 million transfected 293T cells. The plate was then carefully washed twice with 10 ml of ice-cold 1X DPBS. The cells were scraped off using a sterile "cell scraper" and transferred into a 1.5-ml microcentrifuge

tube. Then centrifuged the cells at 4°C, 3000 rpm for 5 minutes. Remove the remaining supernatant from the microcentrifuge tube. Resuspend the cell pellets in 300 µl of Protease Inhibitor Cocktail (Sigma Aldrich Cat. No. P2714). Cell lysates were prepared by combining 5 µl of the resuspended cells (approximately 3.3×10^5 cells), 12.5 µl NuPAGE® LDS Sample Buffer (4X), 5 µl NuPAGE® Sample Reducing Agent (10X), and 27.5 µl ddH₂O. The cell lysates were then passed through a sterile plastic syringe with a 28 G needle or an insulin needle until it became less viscous. Heat the sample at 70°C for 10 minutes. 10 µl of each cell lysate was resolved on NuPAGE™ Novex™ 10% 10-well 1.0 mm thick Bis-Tris Protein Gels (Thermo Fisher Scientific Cat. No. NP0301BOX). The gel electrophoresis was run with 1X MOPS buffer (Thermo Fisher Scientific Cat. No. NP0001), at 200 V for 50 minutes. The remaining cell pellets resuspended in the Protease Inhibitor Cocktail and the cell lysate samples could be stored at -80°C.

The resolved proteins were then transferred to Immobilon-FL membrane by electro-transfer and the membranes were blocked using 5% milk in phosphate-buffered saline. A mouse monoclonal IgG anti-ADAR1 (clone 15.8.6) antibody (Santa Cruz Biotechnology SC-73408) was used to detect huADAR1. A mouse monoclonal IgG anti-SND1 antibody (Sigma Aldrich SAB4200504) was used to detect huTudor-SN. A goat polyclonal anti-FLAG antibody (Novus Biologicals NB600-344) was used to detect FLIS-ADAR2s. A polyclonal rabbit antibody to human beta-tubulin (Abcam, Cambridge, MA) was used to detect cellular beta-tubulin as a loading control. After incubation with secondary

antibodies (IRDye800 anti-mouse or Alexa Fluor 680 anti-rabbit antibodies; Rockland Immunochemicals, Gilbertsville, PA), blots were visualized using the Odyssey Infrared Imaging System (Li-Cor Biosciences, Lincoln, NE) and the protein band intensities were analyzed by using the ImageJ32 package.

G. Total RNA extraction and cytoplasmic RNA fractionation

Total RNA extraction

The following procedure was used for extracting total RNA from 293T cells cultured in a 15-cm plate. To start, 50-ml Lysis buffer was prepared by combining 1 ml of 10 mg/ml Proteinase K (Sigma Aldrich Cat. No. P2308) solution, 10 ml 10% SDS, and 40 ml of stock buffer (0.2 M Tris-HCl pH 7.5, 0.2 M NaCl, 1.5 mM MgCl₂). The medium was then aspirated off of the 15-cm plate. The plate was then carefully washed twice with 15 ml of room temperature PBS and the cells were scraped off using a sterile “cell scraper” into a 50-ml conical tube in 10 ml of lysis buffer. Cells were then pipetted up and down in the tube to assure total lysis.

The lysate was then passed through a sterile plastic syringe with a 21 G needle until the viscosity resembled that of the lysis buffer. It was then incubated at 45°C for 1 to 2 hours in a slow shaking water bath. After the incubation, 630 µl of 5M NaCl was added per 10 ml of the lysate. If the DNA precipitation appeared after adding NaCl, a 21 G needle was used to shear any precipitating DNA. At room temperature, 10 ml of phenol (pH6.8)/CHISAM (1:1) was added to the lysate. It was then vigorously shaken for 1 minute and centrifuged at room

temperature, 3000 rpm for 5 minutes. The extraction was repeated twice with 10 ml CHISAM (chloroform:isoamyl alcohol = 49:1). After the second CHISAM extraction the aqueous phase was transferred to a 50 ml conical tube and 1/20 volume of 3M sodium acetate (NaOAc) and 2.5 volume of 100% ethanol was added. The precipitate was stored as a slurry at -80°C.

Cytoplasmic RNA fractionation

This procedure was used for fractionating cytoplasmic RNA from 293T cells cultured in 4 to 6 15-cm plates. The plates were removed from the incubator and placed on ice. Medium was aspirated off of each plate and each plate was then washed with 10 ml of ice cold PBS twice. Cells were then scraped off into a 6 ml ice cold PBS and added to a 50-ml screw cap conical tube.

The cells were then centrifuged at 1400 rpm for 3 minutes at 4°C, and the supernatant was aspirated off. The cell pellet was placed on ice and 1.5 ml of ice cold RSB (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 1.5 mM MgCl₂) was added using a 5 ml plastic pipet. While gently swirling the tube, an additional 1.5 ml of RSB containing 0.1% IgepalTM CA-630 [Octylphenoxy]polyethoxyethanol (USB Corporation Cat. No. 19626) was added into the resuspended cell pellet. Nuclei were then immediately pelleted by centrifugation at 1200 rpm for 3 minutes at 4°C. . The supernatant was then transferred to a 15-ml screw cap conical tube which was centrifuged at 4°C, 1200 rpm for 3 minutes.. The supernatant was removed into a new 15ml conical tube and centrifuged for the second time to

assure removal of the nuclei. The supernatant was then transferred to a new 15 ml conical tube and 3 ml of 2X PK buffer (200 mM Tris-HCl, pH7.5, 25 mM EDTA, 300 mM NaCl) containing 2% SDS and freshly added 400 µg/ml Proteinase K was added. The 6ml mixture was then incubated at 37°C for 30 min.

At room temperature, 6 ml of phenol (pH6.8)/CHISAM (1:1), was added to the mixture which was vigorously shaken for 1 minute, followed by centrifugation at 3000 rpm for 5 minutes at room temperature. The extraction was repeated with 6 ml CHISAM (chloroform:isoamyl alcohol = 49:1). After the second CHISAM extraction the aqueous phase was transferred to a 50 ml conical tube and 1/20 volume of 3M sodium acetate (NaOAc) and 2.5 volume of 100% ethanol was added. The precipitate was stored as a slurry at -80°C.

To precipitate either total or cytoplasmic RNA samples from the stored slurries, and a portion of the NaOAc/EtOH/RNA slurries were centrifuged at 4°C, 3000 rpm for 60 minutes. Following centrifugation, the supernatant was carefully removed, and the RNA pellets were washed twice with 80% Ethanol and air dried. All liquid on the side of the tubes was carefully removed by capillary action using a sterile pipette tip. The RNA pellets were then resuspended in an appropriate amount of nuclease free ddH₂O. After use, the remaining NaOAc/EtOH/RNA slurries were stored at -80°C.

H. cDNA synthesis and PCR amplification of antisense target sequence and antisense RNA vector:

Total RNA from 293T cells was harvested, 48 hours post-transfection. cDNAs of a target region (AST) and an antisense vector were synthesized from 1 µg of the total RNAs using sequence specific primers. The SuperScript™ III-First-Strand Synthesis System for RT-PCR (Invitrogen Cat. No: 18080-051) was used in the cDNA synthesis. The cDNAs were PCR amplified using custom primers compatible to Illumina's MiSeq system and a Platinum Taq DNA Polymerase High Fidelity. The PCR amplification cycles and the sequences of primers used in this study are shown below.

PCR amplification cycles

Step 1: 94°C for 2 minutes
Step 2: 94°C for 15 seconds
Step 3: 54°C for 30 seconds
Step 4: 68°C for 1 minute and 5 seconds
Step 5: Go to step 2, 24 more times
Step 6: Hold at 4°C

Primers used in the cDNA synthesis and PCR amplification of AST and AS-RNA:

RT-PCR oligo 2527 (antisense RNA target region; AST):
CACTCTTCTCTTTGCCTTGGTGGG

RT-PCR oligo 2531 (antisense RNA vector; AS-RNA):
TGTCCTCATATCGCCTCCTCC

Illumina MiSeq system compatible forward AST primer; oligo 3095:
TCGTCCGCGAGCGTCAGATGTGTATAAGAGACAG AGCCATGTGTAAAATTAACC

Illumina MiSeq system compatible reverse AST primer; oligo 3100:
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTGTCCCTCATATCGCCTCC

Illumina MiSeq system compatible forward AS-RNA primer; oligo 3129:
TCGTCCGCGAGCGTCAGATGTGTATAAGAGACAGTAGAACTTTAAAAATTATGC

Illumina MiSeq system compatible reverse AS-RNA primer; oligo 3132:
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACTAGTAGTTTAAAGTGCAC

Oligo 2527 and oligo 2531 were used for cDNA synthesis of AST and AS-RNA, respectively. Oligo 3095 and oligo 3100 were used for PCR amplification from the AST cDNA. Oligo 3129 and oligo 3132 were used for PCR amplification of AS-RNA cDNA (corresponding to the antisense env sequence of the AS RNA vector).

I. NGS library preparation and sequencing

Tagmentation using a homemade Tn5 Transposase

Amplified PCR products of either AST or antisense env sequence of AS-RNA vector were fragmented using a homemade Tn5 Transposase. Illumina adapters and indices were also added. The tagmentation protocol was developed by Picelli et al²²⁹ and adjusted by Jing Huang for use in Myles H. Thaler Center for AIDS and Human Retrovirus Research, Department of Microbiology, Immunology and Cancer Biology (MIC), the University of Virginia. The procedure is described below.

I1. PCR product purification

AMPure XP beads (Beckman Coulter Life Sciences Cat. No. A63880, 5 ml) to room temperature (RT) were used to purify the PCR products. The AMPure XP beads were first brought to room temperature for at least 30 minutes. They were then vortexed for at least 30 sec and 1.8X volume of the beads was added to each well of a 96-well plate. The slurry was well mixed by pipetting up and down 10 times followed by incubation at RT for 10 minutes. The mixture was then placed on the magnetic stand for 2 minutes or longer until the supernatant

was cleared. On the magnetic stand, the supernatant was removed and discarded and the beads were washed twice with 200 μ l of freshly prepared 80% Ethanol. (Incubate the sample with the ethanol for 30 seconds) and air dried for 15 minutes. The plate was then removed from the magnetic stand, 22.5 μ l of EB buffer (10 mM Tris-Cl, pH 8.5) was added. The slurry was mixed well by pipetting 10 times and incubated at RT for 2 minutes. The plate was then placed back on the magnetic stand for 2 min or longer to clear the beads from the supernatant.. 20 μ l of all the eluted samples were then transferred to a new 96-well plate and the DNA concentration in each well was quantitated using a Qubit high sensitivity ds-DNA kit (Thermo Fisher Scientific Cat. No. Q32854). Then, the purified PCR products were fragmented by a homemade Tn5 transposase, as described below.

12. Tn5 Transposase Tagmentation

5X TAPS-DMF buffer was first prepared by combining 1 volume of the 100 mM Tris(hydroxymethyl)methylAminoPropaneSulfonic (TAPS) (SIGMA-ALDRICH Cat. No. T5130) and 1 volume of N,N-Dimethylformamide (DMF) (SIGMA Life Science Cat. No.D4551-250ML). It was then added to 20- μ l tagmentation reaction consisting of the following:

Tagment buffer (5X TAPS-DMF buffer)	4 μ l
Tn5 transposase	1-5 μ l
<u>10 ng of PCR products</u>	<u>X μl</u>
Nuclease-free ddH ₂ O	add up to a final volume of 20 μ l

The reaction mix was mixed well well by pipetting up and down 10 times, and incubated at at 55°C for 6 minutes, after which it was held at 10°C. The fragmented products were,then purified in order to remove the Tn5 transposase using Zymo-columns (Genesee Scientific Cat# 11-302). The purified PCR products were subsequently indexed in order to prepare NGS sequencing library. The indexing procedure is described below.

13. PCR Enrichment (Index PCR)

The Illumina p5 and p7 indexes and Kapa HiFi HotStart Ready PCR mix (KAPABIOSYSTEMS Cat. No.KK2601) were thawed on ice. 25 µl of the Kapa HiFi HotStart mix was added to each well of a new 96-well PCR plate, together with 5 µl of each index. 10 µl of water and 5 µl of the purified fragmented PCR product (approximately 2.5 ng) were then added to a final volume of 50 µl. The reagents were well mixed, placed on the thermal cycler and run for the following cycles.

Index PCR:

Step 1: 72°C for	3 minutes
Step 2: 95°C for	30 seconds
Step 3: 12 cycles of	
95°C for	10 seconds
55°C for	30 seconds
72°C for	30 seconds
Step 4: 72°C for	5 minutes
Step 5: Hold at 10°C	

Then, the indexed products (NGS sequencing library) were purified by repeating the first PCR product purification. The library was stored at -20°C for

up to a week. The quality of the library was evaluated using high sensitivity DNA chips (AGILENT TECHNOLOGIES Cat. No.5067-4626). Quantification of the samples was assessed using Qubit high sensitivity dsDNA kit (ThermoFisher Scientific Cat. No. Q32854). The DNA library was sequenced using Illumina MiSeq benchtop sequencer and MiSeq Reagent Kit v2 500cycle (Illumina Cat. No. MS-102-2003).

J. Small RNA library preparation

The NEBNext Multiplex Small RNA Library Prep Sets for Illumina (Set 1: E7300S or E7300L; Set 2: E7580S or E7580L; S = 24 reactions and L = 96 reactions) were used to prepare the libraries used for the analysis of the small RNA.

J1. Ligate the 3' SR Adaptor

In a sterile nuclease-free PCR tube, 1 µg of the total RNA, 1 µl of the 3' SR Adaptor for Illumina, and nuclease free water were mixed to the final volume of 7 µl. The reaction was incubated in a preheated thermal cycler for 2 minutes at 70°C. Then, the reaction tube was placed on ice. 10 µl of the 3' Ligation Reaction Buffer (2X) and 3 µl of 3' Ligation Enzyme Mix to the 7-µl reaction were then added, the reaction was mixed well and incubated for 1 hour at 25°C in a thermal cycler.

J2. Hybridize the Reverse Transcription Primer¹

One μl of the SR RT Primer for Illumina and 4.5 μl of the nuclease free water were then added to the 20- μl ligation reaction from the previous step. The sample was heated at 75°C for 5 minutes and then incubated at 37°C for 15 minutes, followed by the 25-°C incubation for 15 minutes. (This step is important to prevent adaptor-dimer formation. The SR RT Primer hybridizes to the excess of 3' SR Adaptor (that remains free after the 3' ligation reaction) and transforms the ssDNA adaptor into a dsDNA molecule. dsDNAs are not substrates for ligation mediated by T4 RNA Ligase 1 and therefore do not ligate to the 5' SR Adaptor in the subsequent ligation step).

J3. Ligate the 5' SR Adaptor

The 5'SR adaptor was resuspended in 120 μl of nuclease-free water. 1.1x N μl of the 5'SR Adaptor were aliquoted into separate, nuclease-free 200 μl PCR tubes, with N equal to the number of samples being processed for a given experiment. The adaptor was incubated in the thermal cycler at 70 °C for 2 minutes and immediately placed ice. The tube was placed on ice and the denatured adaptor was used within 30 minutes of denaturation. The remaining resuspended 5' SR adaptor was stored at -80°C. One μl of the denatured 5' SR Adaptor for Illumina, 1 μl of the 5' Ligation Reaction Buffer (10X), and 2.5 μl of the 5' Ligation Enzyme Mix was then added to the 25.5- μl ligation mixture from the previous step. The 30- μl reaction mixture was incubated at 25°C for 1 hour.

J4. Reverse Transcription

8 μ l of the First Strand Synthesis Reaction Buffer, 1 μ l of the Murine RNase Inhibitor, 1 μ l of the ProtoScript II Reverse Transcriptase were to the 30- μ l of the Adaptor Ligated RNA from the previous step and mixed. The 40- μ l reaction mixture was then incubated at 50°C for 1 hour. Then, the samples were either immediately subjected to PCR amplification. The samples could be heated at at 70°C for 15 minutes in order to inactivate the RT reaction. The heat inactivated samples could be stored at -20°C before proceeding to the PCR amplification step.

J5. PCR Amplification

Fifty μ l of the LongAmp Taq 2X Master Mix, 2.5 μ l of the SR Primer for Illumina, 2.5 μ l of the Index (X) Primer*, and 5 μ l of the nuclease free water were added to the 40- μ l Reverse Transcription product from the previous step and mixed well making the final volume of the PCR reaction 100 μ l. PCR amplification was carried out using the following cycles.

PCR cycling conditions:

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	94°C	30 sec	1
Denaturation	94°C	15 sec	
Annealing	62°C	30 sec	12
Extension	70°C	15 sec	
Final Extension	70°C	5 min	1
Hold	4°C	∞	

The 100- μ l PCR amplified cDNA construct was purified using the QIAQuick PCR Purification Kit. Before eluting the DNA from the column, the

column was centrifuged with the lid of the spin column open for 5 minutes at 13,200 rpm. Centrifugation with the opened lid ensured that no ethanol remained during DNA elution. It was important to dry the spin column membrane of any residual ethanol that may interfere with the correct loading of the sample on the PAGE gel. DNA was eluted in 27.5 μ l of Nuclease-free Water. The eluted samples were checked by loading 1 μ l of the purified PCR reaction on the Bioanalyzer using DNA 1000 chip according to the manufacturer's instruction.

J5. Size Selection using 6% Polyacrylamide Gel

Twenty-five μ l of the purified PCR products was mixed with 5 μ l of the 6X Gel Loading Dye, Blue (vortex the Gel Loading Dye before using) and 15 μ l X 2each of mixed amplified cDNA construct and loading dye were loaded on a 6% PAGE 10-well gel. 5 μ l of Quick-Load pBR322 DNA-MspI Digest was also loaded in one well. The gel was run for 1 hour at 120 V until the blue dye reached the bottom of the gel. The gel was removed from the apparatus and stained the with SYBR Gold nucleic acid gel stain in a clean container for 2-3 minutes and viewed on a UV transilluminator.

Note: The 140 and 150 nucleotide bands correspond to the adaptor-ligated constructs derived from the 21 and 30 nucleotide RNA fragments, respectively. For the miRNAs, isolate the bands corresponding to ~140 bp. For piRNAs, isolate the band corresponding to ~150 bp.

To elute the PCR products from the gel, the two gel slices were placed in a 1.5-ml tube and crushed with the RNase-free Disposable Pellet Pestles. They were then soaked in 250 μ l of the 1X DNA Gel Elution buffer (The NEBNext Multiplex Small RNA Library Prep Sets for Illumina). The tube was rotated end-

to-end for at least 2 hours at room temperature. The eluate and the gel debris was then transferred to the top of a gel filtration column which was centrifuge for 2 minutes at > 13,200 rpm. The eluate was recovered and 1 μ l of the linear Acrylamide, 25 μ l of the 3M sodium acetate, pH 5.5, and 750 μ l of 100% ethanol was added. The solution was vortexed well and. DNA was precipitated by placing the tube at -80°C for at least 30 minutes. To pellet the precipitated DNA the tube was centrifuged at 14,000 x g for 30 minutes at 4°C. The supernatant was carefully removed, taking care not to disturb the pellet. The DNA pellet was washed with 80% ethanol by vortexing vigorously. The DNA was re-pelleted by centrifugation at 14,000 x g for 30 minutes at 4°C and the resulting DNA pellet was air-dried for 10 minutes at room temperature to remove residual ethanol. The DNA pellet was resuspended in 12 μ l of TE Buffer (10mM Tris-HCl, 1mM EDTA). The quality of the library was evaluated using high sensitivity DNA chips (AGILENT TECHNOLOGIES Cat. No.5067-4626). Quantification of the samples was assessed using Qubit high sensitivity dsDNA kit (ThermoFisher Scientific Cat. No. Q32854). The DNA library was sequenced using Illumina MiSeq benchtop sequencer and MiSeq Reagent Kit v2 300cycle (Illumina Cat. No. MS-102-2002).

K. NGS data analysis

NGS data of the antisense env RNA (AS RNA) sequence and the target region (AsT)

FASTQ files, derived from the output of the Illumina MiSeq, were uploaded to Galaxy (<https://usegalaxy.org>; Galaxy version 17.01). The FASTQ quality format of the FASTQ files was converted to that of FASTQSANGER using FASTQ Groomer. Then, filtered the sequences by quality. Quality cut-off value was set to 20, and Percent of bases in sequence that must have quality equal to / higher than cut-off value was set to 90. Then, sequencing artifacts were removed. The quality controlled FASTA files were exported from Galaxy and imported into Geneious version 8.1.9 (Biomatters Ltd, Auckland, NZ). Paired reads of both read 1 and 2 were set and the paired reads were merged using the Geneious “set paired reads and merge pair reads” command. The AS RNA and the AsT sequences were then mapped to the reference sequence of the antisense env RNA vector and the HIV-1 reference genome (M19921.2 HIV NL4-3 Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant clone pNL4-3), respectively. The DNA sequences were mapped to the reference sequence using “Geneious” mapping algorithm. The frequencies of nucleotide changes were obtained from the Geneious Variations/SNPs finder.

NGS data analysis of Small RNA libraries

Only read 1 of FASTQ files, derived from the output of the Illumina MiSeq, were uploaded to Galaxy (<https://usegalaxy.org>; Galaxy version 17.01) to be

filtered by sequence quality and subjected to sequencing artifact removal, as previously described. Then, the quality controlled FASTA files were exported from Galaxy and subjected to the adaptor sequence removal, using “clip_adapters.pl” of mirDeep2008 package^{230,231}. Then, the FASTA sequences were imported into Geneious version 8.1.9 (Biomatters Ltd, Auckland, NZ). To remove human sequences, the NGS reads of the small RNA libraries were first mapped against the human reference genome (NCBI GRCh38) using the Bowtie algorithm version 1.1.2 developed by Ben Langmead and Cole Trapnell, which is available in GENEIOUS bioinformatics software. The unmapped sequences were then mapped against the HIV-1 reference genome (M19921.2 HIV NL4-3 Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant clone pNL4-3), using the same Bowtie algorithm version 1.1.2, in order to detect the small sequences that aligned to the antisense target region.

- End of Dissertation -

References:

1. Rouzioux, C. *et al.* Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome {(AIDS)}. *Science* (80-.). **220**, 868–871 (1983).
2. Barré-Sinoussi, F., Ross, A. L. & Delfraissy, J.-F. Past, present and future: 30 years of HIV research. *Nat. Rev. Microbiol.* **11**, 877–883 (2013).
3. Engelman, A. & Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat. Rev. Microbiol.* **10**, 279–90 (2012).
4. Freed, E. O. HIV-1 replication. *Somat. Cell Mol. Genet.* **26**, 13–33 (2001).
5. Wolf, D. & Goff, S. P. Host restriction factors blocking retroviral replication. *Annu. Rev. Genet.* **42**, 143–163 (2008).
6. Rajsbaum, R. & García-Sastre, A. Viral evasion mechanisms of early antiviral responses involving regulation of ubiquitin pathways. *Trends Microbiol.* **21**, 421–429 (2013).
7. Kar, A. K., Diaz-Griffero, F., Li, Y., Li, X. & Sodroski, J. Biochemical and biophysical characterization of a chimeric TRIM21-TRIM5alpha protein. *J. Virol.* **82**, 11669–81 (2008).
8. Le Grice, S. F. J. Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. *J. Biol. Chem.* **287**, 40850–40857 (2012).
9. Guo, F., Cen, S., Niu, M., Saadatmand, J. & Kleiman, L. Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. *J. Virol.* **80**, 11710–22 (2006).
10. Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M. & Malim, M. H. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog.* **4**, 13–20 (2008).
11. Chiu, Y.-L. *et al.* Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells. *Nature* **435**, 108–114 (2005).
12. Harris, R. S. *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803–809 (2003).
13. Mangeat, B. *et al.* Broad Antiretroviral Defence by Human APOBEC3G Through Lethal Editing of Nascent Reverse Transcripts. *Nature* **424**, 99–103 (2004).
14. Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94–8 (2003).
15. Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat. Med.* **9**, 1404–1407 (2003).
16. Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056–1060 (2003).
17. Nathans, R. *et al.* Small-molecule inhibition of HIV-1 Vif. **26**, 1–6 (2008).
18. Cen, S. *et al.* Small molecular compounds inhibit HIV-1 replication through specifically stabilizing APOBEC3G. *J. Biol. Chem.* **285**, 16546–16552 (2010).

19. Hughes, S. H. & Coffin, J. M. What Integration Sites Tell Us about HIV Persistence. *Cell Host Microbe* **19**, 588–598 (2016).
20. Schröder, A. R. W. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–9 (2002).
21. Mitchell, R. S. *et al.* Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol.* **2**, (2004).
22. Llano, M. *et al.* An Essential Role for LEDGF/p75 in HIV Integration. *Science (80-.)*. **314**, 461–464 (2006).
23. Kammler, S. *et al.* The strength of the HIV-1 3' splice sites affects Rev function. *Retrovirology* **3**, 89 (2006).
24. Emery, A., Zhou, S., Pollom, E. & Swanstrom, R. Characterizing HIV-1 Splicing Using Next Generation Sequencing. *J. Virol.* **91**, (2017).
25. Purcell, D. F. & Martin, M. Alternative Splicing of Human Immunodeficiency Virus Type 1 mRNA Modulates Viral Protein Expression, Replication, and Infectivity. *J. Virol.* **67**, 6365–6378 (1993).
26. Chukkapalli, V. & Ono, A. Molecular determinants that regulate plasma membrane association of HIV-1 Gag. *J. Mol. Biol* **410**, 512–524 (2011).
27. Votteler, J. & Sundquist, W. I. Virus Budding and the ESCRT Pathway. *Cell Host Microbe* **14**, (2013).
28. Neil, S. J. D., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425–430 (2008).
29. Van Damme, N. *et al.* The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein. *Cell Host Microbe* **3**, 245–252 (2008).
30. Mangeat, B. *et al.* HIV-1 Vpu neutralizes the antiviral factor tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation. *PLoS Pathog.* **5**, (2009).
31. Dubé, M. *et al.* Antagonism of tetherin restriction of HIV-1 release by Vpu involves binding and sequestration of the restriction factor in a perinuclear compartment. *PLoS Pathog.* **6**, e1000856 (2010).
32. Sundquist, W. I. & Krausslich, H.-G. HIV-1 Assembly , Budding , and Maturation. *Cold Spring Harb Perspect Med* **2012**; (2012).
33. Pettit, S. C. *et al.* The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J. Virol.* **68**, 8017–27 (1994).
34. Pokorná, J., Machala, L., Řezáčová, P. & Konvalinka, J. *Current and novel inhibitors of HIV protease.* *Viruses* **1**, (2009).
35. Freed, E. O. HIV-1 assembly, release and maturation. *Nat Rev Microbiol* **13**, 484–496 (2015).
36. Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V & Cullen, B. R. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**, 254–257 (1989).
37. Emerman, M., Vazeux, R. & Peden, K. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell* **57**, 1155–1165 (1989).
38. Hammariskjold, M. L. *et al.* Regulation of human immunodeficiency virus

- env expression by the rev gene product. *J Virol* **63**, 1959–1966 (1989).
39. Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. & Pavlakis, G. N. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 1495–9 (1989).
 40. Malim, M. H., Böhnlein, S., Hauber, J. & Cullen, B. R. Functional dissection of the HIV-1 Rev trans-activator-Derivation of a trans-dominant repressor of Rev function. *Cell* **58**, 205–214 (1989).
 41. Fischer, U., Huber, J., Boelens, W. C., Mattajt, L. W. & Lührmann, R. 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 (1995).
 42. Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I. W. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051–1060 (1997).
 43. Neville, M., Stutz, F., Lee, L., Davis, L. I. & Rosbash, M. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* **7**, 767–775 (1997).
 44. Cullen, B. R. Nuclear mRNA export: Insights from virology. *Trends Biochem. Sci.* **28**, 419–424 (2003).
 45. Emerman, M. & Malim, M. H. HIV-1 Regulatory / Accessory Genes : Keys to Unraveling Viral and Host Cell Biology. *Science (80-)*. **280**, 1880–1884 (1998).
 46. Bray, M. *et al.* A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 1256–60 (1994).
 47. Ernst, R. K., Bray, M., Rekosh, D. & Hammarskjöld, M. L. A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Mol. Cell. Biol.* **17**, 135–44 (1997).
 48. Ernst, R. K., Bray, M., Rekosh, D. & Hammarskjöld, M.-L. Secondary structure and mutational analysis of the Mason-Pfizer monkey virus RNA constitutive transport element. *RNA* **3**, 210–222 (1997).
 49. Zolotukhin, A., Valentin, A., Pavlakis, G. & Felber, B. Continuous propagation of RRE(-) and Rev(-)RRE(-) human immunodeficiency virus type 1 molecular clones containing a cis-acting element of simian retrovirus type 1 in human peripheral blood lymphocytes. *J. Virol.* **68**, 7944–7952 (1994).
 50. Pasquinelli, A. E. *et al.* The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J.* **16**, 7500–7510 (1997).
 51. Grüter, P. *et al.* TAP, the Human Homolog of Mex67p, Mediates CTE-Dependent RNA Export from the Nucleus. *Mol. Cell* **1**, 649–659 (1998).
 52. Clouse, K. N., Luo, M. J., Zhou, Z. & Reed, R. A Ran-independent pathway for export of spliced mRNA. *Nat. Cell Biol.* **3**, 97–99 (2001).
 53. Li, Y. *et al.* An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature* **443**, 234–237 (2006).
 54. Li, Y., Bor, Y.-C., Fitzgerald, M. P., Lee, K. S. & Rekosh, D. An NXF1

- mRNA with a retained intron is expressed in hippocampal and neocortical neurons and is translated into a protein that functions as an Nxf1 co-factor Myles H. Thaler Center for AIDS and Human Retrovirus Research and the Department of Microbiology. *Mol. Biol. Cell* **27**, 1–37 (2016).
55. Wang, B., Rekosh, D. & Hammarskjold, M.-L. Evolutionary conservation of a molecular machinery for export and expression of mRNAs with retained introns. *RNA* 426–437 (2015). doi:10.1261/rna.048520.114
 56. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.* **8**, 761–773 (2007).
 57. Carmody, S. R. & Wentz, S. R. mRNA nuclear export at a glance. *J. Cell Sci.* **122**, 1933–1937 (2009).
 58. Masuda, S. *et al.* Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.* **19**, 1512–1517 (2005).
 59. Cheng, H. *et al.* Human mRNA Export Machinery Recruited to the 5' End of mRNA. *Cell* **127**, 1389–1400 (2006).
 60. Björk, P. & Wieslander, L. Integration of mRNP formation and export. *Cell. Mol. Life Sci.* (2017). doi:10.1007/s00018-017-2503-3
 61. Brennan, C. M., Gallouzi, I. E. & Steitz, J. A. Protein ligands to HuR modulate its interaction with target mRNAs in vivo. *J. Cell Biol.* **151**, 1–13 (2000).
 62. Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M. & Borden, K. L. B. eIF4E is a central node of an RNA regulon that governs cellular proliferation. *J. Cell Biol.* **175**, 415–426 (2006).
 63. Prechtel, A. T. *et al.* Expression of CD83 is regulated by HuR via a novel cis-active coding region RNA element. *J. Biol. Chem.* **281**, 10912–10925 (2006).
 64. Siddiqui, N. & Borden, K. L. B. mRNA export and cancer. *Wiley Interdiscip. Rev. RNA* **3**, 13–25 (2012).
 65. Katayama, S. *et al.* Antisense transcription in the mammalian transcriptome. *Sci. (New York, NY)* **309**, 1564–1566 (2005).
 66. Kawaji, H. *et al.* Hidden layers of human small RNAs. *BMC Genomics* **9**, 157 (2008).
 67. Katayama, S. Antisense Transcription in the Mammalian Transcriptome. *Science (80-)*. **1564**, 1564–6 (2012).
 68. Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–93 (2013).
 69. Georg, J. & Hess, W. R. cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.* **75**, 286–300 (2011).
 70. Terryn, N. & Rouzé, P. The sense of naturally transcribed antisense RNAs in plants. *Trends Plant Sci.* **5**, 394–396 (2000).
 71. Georg, J., Honsel, A., Voß, B., Rennenberg, H. & Hess, W. R. A long antisense RNA in plant chloroplasts. *New Phytol.* **186**, 615–622 (2010).
 72. Faghihi, M. A. & Wahlestedt, C. Regulatory roles of natural antisense transcripts. *Nat. Rev. Mol. Cell Biol.* **10**, 637–43 (2009).
 73. Faghihi, M. A. *et al.* Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β -

- secretase. *Nat. Med.* **14**, 723–730 (2008).
74. Lyle, R. *et al.* The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat. Genet.* **25**, 19–21 (2000).
 75. Zhao, J., Sun, B. K., Erwin, J. A., Song, J.-J. & Lee, J. T. Polycomb Proteins Targeted by a Short Repeat RNA to the Mouse X Chromosome. *Science (80-.)*. **322**, 750–756 (2008).
 76. Gupta, R. a *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–6 (2010).
 77. Tsai, M.-C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689–93 (2010).
 78. Yap, K. L. *et al.* Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Mol. Cell* **38**, 662–674 (2010).
 79. Hatzoglou, A. *et al.* Natural antisense RNA inhibits the expression of BCMA, a tumour necrosis factor receptor homologue. *BMC Mol. Biol.* **3**, 1–7 (2002).
 80. Tufarelli, C. *et al.* Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat. Genet.* **34**, 157–165 (2003).
 81. Luther, H. P. Role of endogenous antisense RNA in cardiac gene regulation. *J. Mol. Med.* **83**, 26–32 (2005).
 82. Miller, R. Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand. *Science (80-.)*. **239**, 1420–1422 (1988).
 83. Michael, N. L. *et al.* Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J. Virol.* **68**, 979–987 (1994).
 84. Ludwig, L. B. *et al.* Human Immunodeficiency Virus-Type 1 LTR DNA contains an intrinsic gene producing antisense RNA and protein products. *Retrovirology* **3**, 80 (2006).
 85. Vanhée-Brossollet, C. *et al.* A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals. *Virology* **206**, 196–202 (1995).
 86. Kobayashi-Ishihara, M. *et al.* HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. *Retrovirology* **9**, 38 (2012).
 87. Landry, S. *et al.* Detection, characterization and regulation of antisense transcripts in HIV-1. *Retrovirology* **4**, 1–16 (2007).
 88. Clerc, I. *et al.* Polarized expression of the membrane ASP Polarized expression of the membrane ASP protein derived from HIV-1 antisense transcription in T cells. *Retrovirology* **8**, (2011).
 89. Matsuoka, M. & Jeang, K. T. Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer* **7**, 270–280 (2007).
 90. Briquet, S., Richardson, J., Vanhée-Brossollet, C. & Vaquero, C. Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus. *Gene* **267**, 157–164 (2001).
 91. Rasmussen, M. H. *et al.* Antisense transcription in gammaretroviruses as a

- mechanism of insertional activation of host genes. *J. Virol.* **84**, 3780–3788 (2010).
92. Manghera, M., Magnusson, A. & Douville, R. N. The sense behind retroviral anti-sense transcription. *Viol. J.* **14**, 9 (2017).
 93. Kobayashi-ishihara, M. *et al.* HIV-1-encoded antisense RNA suppresses viral replication for a prolonged period. ??? **9**, 1 (2012).
 94. Peeters, a, Lambert, P. F. & Deacon, N. J. A fourth Sp1 site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription. *J. Virol.* **70**, 6665–6672 (1996).
 95. Lin, S., Zhang, L., Luo, W. & Zhang, X. Characteristics of antisense transcript promoters and the regulation of their activity. *Int. J. Mol. Sci.* **17**, 1–17 (2015).
 96. Michael, N. L. *et al.* Negative-Strand RNA Transcripts Are Produced in Human Immunodeficiency Virus Type 1-Infected Cells and Patients by a Novel Promoter Downregulated by Tat. 979–987 (1994).
 97. Sandelin, A. *et al.* Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat. Rev. Genet.* **8**, 424–436 (2007).
 98. Anderson, W. F., Blaese, R. M. & Culver, K. The ADA human gene therapy clinical protocol: points to consider response with clinical protocol. *Hum Gene Ther* **1**, 331–362 (1990).
 99. Cavazzana-Calvo, M. *et al.* Gene Therapy of Human Severe Combined Immunodeficiency (SCID)– X1 Disease. *Science (80-)*. **288**, 669–672 (2000).
 100. Howe, S. J. *et al.* Insertional mutagenesis in combination with acquired somatic mutations leads to leukemogenesis following gene therapy of SCID-X1. *J. Clin* **118**, 3143–50 (2008).
 101. Modlich, U. *et al.* Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol. Ther.* **17**, 1919–28 (2009).
 102. Lewis, P. F. & Emerman, M. Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *J. Virol.* **68**, 510–6 (1994).
 103. Escors, D. & Breckpot, K. Lentiviral vectors in gene therapy: Their current status and future potential. *Arch. Immunol. Ther. Exp. (Warsz)*. **58**, 107–119 (2010).
 104. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72**, 8463–71 (1998).
 105. Liechtenstein, T., Perez-Janices, N. & Escors, D. Lentiviral vectors for cancer immunotherapy and clinical applications. *Cancers (Basel)*. **5**, 815–837 (2013).
 106. Shaw, A. & Cornetta, K. Design and Potential of Non-Integrating Lentiviral Vectors. *Biomedicines* **2**, 14–35 (2014).
 107. Yu, S. F. *et al.* Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3194–8 (1986).
 108. Miyoshi, H., Blömer, U., Takahashi, M., Gage, F. H. & Verma, I. M.

- Development of a self-inactivating lentivirus vector. *J. Virol.* **72**, 8150–7 (1998).
109. Zufferey, R. *et al.* Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J. Virol.* **72**, 9873–80 (1998).
 110. Coil, D. A. & Miller, A. D. Phosphatidylserine Is Not the Cell Surface Receptor for Vesicular Stomatitis Virus Phosphatidylserine Is Not the Cell Surface Receptor for Vesicular Stomatitis Virus. *J. Virol.* **78**, 10920–10926 (2004).
 111. Cronin, J., Zhang, X.-Y. & Reiser, J. Altering the tropism of lentiviral vectors through pseudotyping. *Curr. Gene Ther.* **5**, 387–398 (2005).
 112. Burns, J. C., Friedmann, T., Driever, W., Burrascano, M. & Yee, J.-K. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: Concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells (gene therapy/zebrafish). *Genetics* **90**, 8033–8037 (1993).
 113. Morgan, R. A. *et al.* Cancer Regression in Patients After Transfer of Genetically Engineered Lymphocytes. *Science (80-.)*. **314**, 126–129 (2006).
 114. Cavalieri, S. *et al.* Human T lymphocytes transduced by lentiviral vectors in the absence of TCR activation maintain an intact immune competence. *Blood* **102**, 497–505 (2003).
 115. Breckpot, K., Aerts, J. L. & Thielemans, K. Lentiviral vectors for cancer immunotherapy: transforming infectious particles into therapeutics. *Gene Ther.* **14**, 847–862 (2007).
 116. Arce, F., Breckpot, K., Collins, M. & Escors, D. Europe PMC Funders Group Targeting lentiviral vectors for cancer immunotherapy. **7**, 248–260 (2012).
 117. Porter, D. L., Levine, B. L., Kalos, M., Bagg, A. & June, C. H. Chimeric Antigen Receptor–Modified T Cells in Chronic Lymphoid Leukemia. *N Engl J Med* **365**, 725–733 (2011).
 118. Rodgers, D. T. *et al.* Switch-mediated activation and retargeting of CAR-T cells for B-cell malignancies. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E459-68 (2016).
 119. Johnson, L. A. & June, C. H. Driving gene-engineered T cell immunotherapy of cancer. *Cell Res.* **27**, 38–58 (2016).
 120. Dias, N. & Stein, C. a. Minireview Antisense Oligonucleotides : Basic Concepts and Mechanisms. *Mol. Cancer Ther.* **1**, 347–355 (2002).
 121. Stephenson, M. L. & Zamecnik, P. C. Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proc Natl Acad Sci U S A* **75**, 285–288 (1978).
 122. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 280–284 (1978).
 123. Zamecnik, P. C., Goodchild, J., Taguchi, Y. & Sarin, P. S. Inhibition of replication and expression of human T-cell lymphotropic virus type III in cultured cells by exogenous synthetic oligonucleotides complementary to

- viral RNA. *Proc Natl Acad Sci U S A* **83**, 4143–4146 (1986).
124. Baltimore, D. Gene therapy. Intracellular immunization. *Nature* **335**, 395–396 (1988).
 125. Shimada, T. *et al.* Trial of antisense RNA inhibition of HIV replication and gene expression. *Antivir. Chem. Chemother.* **2**, 133–142 (1991).
 126. Lu, X. *et al.* Antisense-mediated inhibition of human immunodeficiency virus (HIV) replication by use of an HIV type 1-based vector results in severely attenuated mutants incapable of developing resistance. *J. Virol.* **78**, 7079–7088 (2004).
 127. Ward, A. M., Rekosh, D. & Hammarskjold, M.-L. Trafficking through the Rev/RRE pathway is essential for efficient inhibition of human immunodeficiency virus type 1 by an antisense RNA derived from the envelope gene. *J. Virol.* **83**, 940–952 (2009).
 128. Tebas, P. *et al.* Antiviral effects of autologous CD4 T cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to HIV. *Blood* **121**, 1524–1533 (2013).
 129. Elbashir, S. M. *et al.* Duplexes of 21 ± nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
 130. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–854 (1993).
 131. Kavi, H. H., Fernandez, H., Xie, W. & Birchler, J. A. Genetics and biochemistry of RNAi in *Drosophila*. *Curr. Top. Microbiol. Immunol.* **320**, 37–75 (2008).
 132. Skalsky, R. L. & Cullen, B. R. Viruses, microRNAs, and host interactions. *Annu. Rev. Microbiol.* **64**, 123–141 (2010).
 133. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**, 126–139 (2009).
 134. Czech, B. & Hannon, G. J. Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet* **12**, 19–31 (2011).
 135. Zeng, Y. & Cullen, B. R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res.* **32**, 4776–4785 (2004).
 136. Park, J.-E. *et al.* Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **475**, 201–205 (2011).
 137. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744 (2005).
 138. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–41 (2004).
 139. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185–197 (2004).
 140. Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* **79**, 351–379 (2010).
 141. Bass, B. L. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **71**, 817–846 (2002).

142. Kazuko Nishikura. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **79**, 321–349 (2010).
143. Wulff, B. E. & Nishikura, K. Substitutional A-to-I RNA editing. *Wiley Interdiscip. Rev. RNA* **1**, 90–101 (2010).
144. Kim, U., Wang, Y., Sanford, T., Zeng, Y. & Nishikura, K. Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11457–61 (1994).
145. Nishikura, K. Editor meets silencer: RNA editing and RNA interference. *Nat. Rev. Mol. Cell Biol.* **7**, 919–931 (2006).
146. Doria, M., Neri, F., Gallo, A., Farace, M. G. & Michienzi, A. Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. *Nucleic Acids Res.* **37**, 5848–5858 (2009).
147. Doria, M. *et al.* ADAR2 editing enzyme is a novel human immunodeficiency virus-1 proviral factor. *J. Gen. Virol.* **92**, 1228–1232 (2011).
148. Phuphuakrat, A. *et al.* Double-stranded RNA adenosine deaminases enhance expression of human immunodeficiency virus type 1 proteins. *J Virol* **82**, 10864–10872 (2008).
149. Gélinas, J.-F., Clerzius, G., Shaw, E. & Gatignol, A. Enhancement of Replication of RNA Viruses by ADAR1 via RNA Editing and Inhibition of RNA-Activated Protein Kinase. *J. Virol.* **85**, 8460–8466 (2011).
150. Zhang, Z. & Carmichael, G. G. The fate of dsRNA in the Nucleus: A p54nrb-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106**, 465–475 (2001).
151. Scadden, A. D. J. The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nat. Struct. Mol. Biol.* **12**, 489–496 (2005).
152. Caudy, A. *a et al.* A micrococcal nuclease homologue in RNAi effector complexes. *Nature* **425**, 411–414 (2003).
153. Chadwick, D. R. & Lever, a M. Antisense RNA sequences targeting the 5' leader packaging signal region of human immunodeficiency virus type-1 inhibits viral replication at post-transcriptional stages of the life cycle. *Gene Ther.* **7**, 1362–8 (2000).
154. Davis, B. M., Humeau, L. & Dropulic, B. In vivo selection for human and murine hematopoietic cells transduced with a therapeutic MGMT lentiviral vector that inhibits HIV replication. *Mol. Ther.* **9**, 160–172 (2004).
155. Donahue, R. E. *et al.* Reduction in SIV replication in rhesus macaques infused with autologous lymphocytes engineered with antiviral genes. *Nature* **4**, 181–186 (1998).
156. Kim, J. H. *et al.* Inhibition of HIV replication by sense and antisense rev response elements in HIV-based retroviral vectors. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirology* **12**, 343–351 (1996).
157. Kusunoki, A., Saitou, T., Miyano-Kurosaki, N. & Takaku, H. Inhibition of the human chemokine receptor CXCR4 by antisense phosphorothioate oligodeoxyribonucleotides. *FEBS Lett.* **488**, 64–68 (2001).
158. Mautino, M. R. & Morgan, R. A. Inhibition of HIV-1 replication by novel

- lentiviral vectors expressing transdominant Rev and HIV-1 env antisense. *Gene Ther* **9**, 421–431 (2002).
159. Morel, F., Salimi, S., Markovits, J., Austin, T. W. & Plavec, I. Hematologic recovery in mice transplanted with bone marrow stem cells expressing anti-human immunodeficiency virus genes. *Hum. Gene Ther.* **10**, 2779–2787 (1999).
 160. Morgan, R. A. & Walker, R. Gene Therapy for AIDS Using Retroviral Mediated Gene Transfer to Deliver HIV-1 Antisense TAR and Transdominant Rev Protein Genes to Syngeneic Lymphocytes in HIV-1 Infected Identical Twins. *Hum Gene Ther* **7**, 1281–1306 (1996).
 161. Shahabuddin, M. & Khan, A. S. Inhibition of human immunodeficiency virus type 1 by packageable, multigenic antisense RNA. *Antisense Nucleic Acid Drug Dev.* **10**, 141–51. (2000).
 162. Tung, F. Y. T. & Tung, M. H. Characterization of antisense RNA-Mediated inhibition of SIV replication. *J. Med. Virol.* **48**, 321–325 (1996).
 163. Veres, G. *et al.* Comparative Analyses of Intracellularly Expressed Antisense RNAs as Inhibitors of Human Immunodeficiency Virus Type 1 Replication. *J. Virol.* **72**, 1894–1901 (1998).
 164. Sadler, H. A., Stenglein, M. D., Harris, R. S. & Mansky, L. M. APOBEC3G Contributes to HIV-1 Variation through Sublethal Mutagenesis. *J. Virol.* **84**, 7396–7404 (2010).
 165. Huntley, M. A. *et al.* Complex regulation of ADAR-mediated RNA-editing across tissues. *BMC Genomics* **17**, 61 (2016).
 166. Patterson, J. B., Thomis, D. C., Hans, S. L. & Samuel, C. E. Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology* **210**, 508–11. (1995).
 167. Patterson, J. B. & Samuel, C. E. Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. *Mol. Cell. Biol.* **15**, 5376–5388 (1995).
 168. Liu, Y. & Samuel, C. E. Mechanism of interferon action: functionally distinct RNA-binding and catalytic domains in the interferon-inducible, double-stranded RNA-specific adenosine deaminase. *J. Virol.* **70**, 1961–8 (1996).
 169. Nishikura, K. A-to-I editing of coding and non-coding RNAs by ADARs. *Nat. Rev. Mol. Cell Biol.* **17**, 83–96 (2015).
 170. Chilibeck, K. A. *et al.* FRET analysis of in vivo dimerization by RNA-editing enzymes. *J. Biol. Chem.* **281**, 16530–16535 (2006).
 171. Samuel, C. E. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411**, 180–193 (2011).
 172. Jayan, G. C. & Casey, J. L. Inhibition of Hepatitis Delta Virus RNA Editing by Short Inhibitory RNA-Mediated Knockdown of ADAR1 but Not ADAR2 Expression Inhibition of Hepatitis Delta Virus RNA Editing by Short Inhibitory RNA-Mediated Knockdown of ADAR1 but Not ADAR2 Expression. **76**, 12399–12404 (2002).
 173. Heale, B. S. E. *et al.* Editing independent effects of ADARs on the

- miRNA/siRNA pathways. *EMBO J.* **28**, 3145–56 (2009).
174. Polson, A. G. & Bass, B. L. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.* **13**, 5701–11 (1994).
 175. Eggington, J. M., Greene, T. & Bass, B. L. Predicting sites of ADAR editing in double-stranded RNA. *Nat. Commun.* **2**, 319 (2011).
 176. Kuttan, A. & Bass, B. L. Mechanistic insights into editing-site specificity of ADARs. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E3295–304 (2012).
 177. Nishikura, K. Editor meets silencer : crosstalk between RNA editing and RNA interference. *7*, 919–931 (2006).
 178. Michienzi, A., Cagnon, L., Bahner, I. & Rossi, J. J. Ribozyme-mediated inhibition of HIV 1 suggests nucleolar trafficking of HIV-1 RNA. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8955–60 (2000).
 179. Desterro, J. M. P. *et al.* Dynamic association of RNA-editing enzymes with the nucleolus. *J. Cell Sci.* **116**, 1805–1818 (2003).
 180. Nasioulas, G. *et al.* Elements distinct from human immunodeficiency virus type 1 splice sites are responsible for the Rev dependence of env mRNA. *J Virol* **68**, 2986–2993 (1994).
 181. Pfaller, C. K., Li, Z., George, C. X. & Samuel, C. E. Protein Kinase PKR and RNA Adenosine Deaminase ADAR1: New Roles for Old Players as Modulators of the Interferon Response. *Curr Opin Immunol* **23**, 573–582 (2011).
 182. Samuel, C. E. ADARs, Viruses and Innate Immunity. *Curr Top Microbiol Immunol* **353**, 1–28 (2012).
 183. Tomaselli, S., Galeano, F., Locatelli, F. & Gallo, A. Adars and the balance game between virus infection and innate immune cell response. *Curr. Issues Mol. Biol.* **17**, 37–52 (2015).
 184. Chen, L. L. & Carmichael, G. G. Altered Nuclear Retention of mRNAs Containing Inverted Repeats in Human Embryonic Stem Cells: Functional Role of a Nuclear Noncoding RNA. *Mol. Cell* **35**, 467–478 (2009).
 185. Scadden, A. D. J. & Smith, C. W. J. A ribonuclease specific for inosine-containing RNA: A potential role in antiviral defence? *EMBO J.* **16**, 2140–2149 (1997).
 186. Scadden, A. D. J. & Smith, C. W. J. RNAi is antagonized by AAel hyper-editing. *EMBO Rep.* **2**, 1107–1111 (2001).
 187. Morita, Y. *et al.* Human endonuclease V is a ribonuclease specific for inosine-containing RNA. *Nat. Commun.* **4**, 2273 (2013).
 188. Vik, E. S. *et al.* Endonuclease V cleaves at inosines in RNA. *Nat. Commun.* **4**, 2271 (2013).
 189. Alseth, I., Dalhus, B. & Bjørås, M. Inosine in DNA and RNA. *Curr. Opin. Genet. Dev.* **26**, 116–123 (2014).
 190. Das, A. K. & Carmichael, G. G. ADAR Editing Wobbles the MicroRNA World. *ACS Chem. Biol.* **2**, 217–220 (2007).
 191. Engeland, C. E. *et al.* Proteome analysis of the HIV-1 Gag interactome. *Virology* **460–461**, 194–206 (2014).
 192. Li, C. L., Yang, W. Z., Chen, Y. P. & Yuan, H. S. Structural and functional

- insights into human Tudor-SN, a key component linking RNA interference and editing. *Nucleic Acids Res.* **36**, 3579–3589 (2008).
193. Yeung, M. L. *et al.* Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: Evidence for the processing of a viral-cellular double-stranded RNA hybrid. *Nucleic Acids Res.* **37**, 6575–6586 (2009).
 194. Klase, Z. *et al.* HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR. *BMC Mol. Biol.* **8**, 63 (2007).
 195. Narayanan, A., Kehn-Hall, K., Bailey, C. & Kashanchi, F. Analysis of the roles of HIV-derived microRNAs. *Expert Opin. Biol. Ther.* **11**, 17–29 (2011).
 196. Ouellet, D. L. *et al.* Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res.* **36**, 2353–2365 (2008).
 197. Lamersa, S. L., Fogel, G. B. & McGrath, M. S. HIV-miR-H1 evolvability during HIV pathogenesis. *Biosystems* **101**, 88–96 (2010).
 198. Omoto, S. *et al.* HIV-1 nef suppression by virally encoded microRNA. *Retrovirology* **1**, 44 (2004).
 199. Schopman, N. C. T. *et al.* Deep sequencing of virus-infected cells reveals HIV-encoded small RNAs. *Nucleic Acids Res.* **40**, 414–427 (2012).
 200. Klase, Z., Houzet, L. & Jeang, K.-T. Replication competent HIV-1 viruses that express intragenomic microRNA reveal discrete RNA-interference mechanisms that affect viral replication. *Cell Biosci.* **1**, 38 (2011).
 201. Klase, Z. *et al.* HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. *Retrovirology* **6**, 18 (2009).
 202. Narayanan, A. *et al.* Exosomes derived from HIV-1-infected cells contain trans-activation response element RNA. *J. Biol. Chem.* **288**, 20014–20033 (2013).
 203. Zhang, Y. *et al.* A novel HIV-1-encoded microRNA enhances its viral replication by targeting the TATA box region. *Retrovirology* **11**, 23 (2014).
 204. Bernard, M. A. *et al.* Novel HIV-1 MiRNAs stimulate TNF α release in human macrophages via TLR8 signaling pathway. *PLoS One* **9**, (2014).
 205. Hundley, H. A., Krauchuk, A. A. & Bass, B. L. C. *C. elegans* and *H. sapiens* mRNAs with edited 3' UTRs are present on polysomes. *RNA* **14**, 2050–2060 (2008).
 206. Legiewicz, M. *et al.* Resistance to RevM10 inhibition reflects a conformational switch in the HIV-1 Rev response element. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14365–70 (2008).
 207. Carter, B. J. Adeno-associated virus and the development of adeno-associated virus vectors: A historical perspective. *Mol. Ther.* **10**, 981–989 (2004).
 208. Rolling, F. Recombinant AAV-mediated gene transfer to the retina: gene therapy perspectives. *Gene Ther.* **11 Suppl 1**, S26-32 (2004).
 209. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr. Opin. Mol. Ther.* **11**, 442–447 (2009).
 210. Naldini, L. Gene therapy returns to centre stage. *Nature* **526**, 351–360 (2015).

211. Maguire, A. M. *et al.* Safety and Efficacy of Gene Transfer for Leber's Congenital Amaurosis. *J. Med.* **358**, 2240–2248 (2008).
212. Horsch, M. *et al.* Requirement of the RNA-editing enzyme ADAR2 for normal physiology in mice. *J. Biol. Chem.* **286**, 18614–18622 (2011).
213. Zhang, Z., Hao, Z., Wang, Z., Li, Q. & Xie, W. Structure of human endonuclease V as an inosine-specific ribonuclease. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 2286–2294 (2014).
214. Kennedy, E. M. *et al.* Production of functional small interfering RNAs by an amino-terminal deletion mutant of human Dicer. *Proc. Natl. Acad. Sci.* **113**, E6547–E6547 (2016).
215. Hafner, M. *et al.* Resource Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
216. Mukherjee, N. *et al.* Integrative Regulatory Mapping Indicates that the RNA-Binding Protein HuR Couples Pre-mRNA Processing and mRNA Stability. *Mol. Cell* **43**, 327–339 (2011).
217. Lebedeva, S. *et al.* Transcriptome-wide Analysis of Regulatory Interactions of the RNA-Binding Protein HuR. *Mol. Cell* **43**, 340–352 (2011).
218. Aalto, A. P. & Pasquinelli, A. E. Small non-coding RNAs mount a silent revolution in gene expression. *Curr. Opin. Cell Biol.* **24**, 333–340 (2012).
219. Weiss, B., Davidkova, G. & Zhou, L. Antisense RNA gene therapy for studying and modulating biological processes. **55**, 334–358 (1999).
220. Matsui, M. & Corey, D. R. Non-coding RNAs as drug targets. *Nat Rev Drug Discov* **16**, 167–179 (2017).
221. Kung, J. T. Y., Colognori, D. & Lee, J. T. Long Noncoding RNAs: Past, Present, and Future. *Genetics* **193**, 651–669 (2013).
222. D'Ydewalle, C. *et al.* The Antisense Transcript SMN-AS1 Regulates SMN Expression and Is a Novel Therapeutic Target for Article The Antisense Transcript SMN-AS1 Regulates SMN Expression and Is a Novel Therapeutic Target for Spinal Muscular Atrophy. *Neuron* **93**, 66–79 (2017).
223. Adachi, A. *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59**, 284–91 (1986).
224. Connor, R. I., Chen, B. K., Choe, S. & Landau, N. R. Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes. *Virology* **206**, 935–44 (1995).
225. Välineva, T., Yang, J., Palovuori, R. & Silvennoinen, O. The Transcriptional Co-activator Protein p100 Recruits Histone Acetyltransferase Activity to STAT6 and Mediates Interaction between the CREB-binding Protein and STAT6. *J. Biol. Chem.* **280**, 14989–14996 (2005).
226. Coyle, J. H., Bor, Y. C., Rekosh, D. & Hammarskjöld, M. L. The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. *Rna* **17**, 1344–1356 (2011).
227. Graham, F. L. & van der Eb, A. J. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–467 (1973).
228. Wehrly, K. & Chesebro, B. p24 Antigen Capture Assay for Quantification of

- Human Immunodeficiency Virus Using Readily Available Inexpensive Reagents. *Methods* **12**, 288–293 (1997).
229. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).
230. Friedländer, M. R. *et al.* Discovering microRNAs from deep sequencing data using miRDeep. *Nat. Biotechnol.* **26**, 407–15 (2008).
231. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).
232. Cullen, B. R. Nuclear RNA export pathways. *J. Cell Sci.* **116**, 587–597 (2003).