Published in final edited form as:
Wiley Interdiscip Rev RNA. 2013 ; 4(5): . doi:10.1002/wrna.1179.

Posttranscriptional regulation of retroviral gene expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA

Jason LeBlanc[†], **Jason Weil**, and **Karen Beemon**[*]
Biology Department, Johns Hopkins University, Baltimore, MD, USA

Abstract

After reverse transcription of the retroviral RNA genome and integration of the DNA provirus into the host genome, host machinery is used for viral gene expression along with viral proteins and RNA regulatory elements. Here, we discuss co-transcriptional and posttranscriptional regulation of retroviral gene expression, comparing simple and complex retroviruses. Cellular RNA polymerase II synthesizes full-length viral primary RNA transcripts that are capped and polyadenylated. All retroviruses generate a singly spliced *env* mRNA from this primary transcript, which encodes the viral glycoproteins. In addition, complex viral RNAs are alternatively spliced to generate accessory proteins, such as Rev, which is involved in posttranscriptional regulation of HIV-1 RNA. Importantly, the splicing of all retroviruses is incomplete; they must maintain and export a fraction of their primary RNA transcripts. This unspliced RNA functions both as the major mRNA for Gag and Pol proteins and as the packaged genomic RNA. Different retroviruses export their unspliced viral RNA from the nucleus to the cytoplasm by either Tap-dependent or Rev/CRM1-dependent routes. Translation of the unspliced mRNA involves frame-shifting or termination codon suppression so that the Gag proteins, which make up the capsid, are expressed more abundantly than the Pol proteins, which are the viral enzymes. After the viral polyproteins assemble into viral particles and bud from the cell membrane, a viral encoded protease cleaves them. Some retroviruses have evolved mechanisms to protect their unspliced RNA from decay by nonsense-mediated RNA decay and to prevent genome editing by the cellular APOBEC deaminases.

INTRODUCTION

Retroviruses were discovered in the beginning of the 20th century and have been used as model systems in the study of a large number of RNA-related cellular functions, including transcription, splicing, nuclear export, and translation. Pioneering work on retroviruses led to the discovery of oncogenes and oncogenic microRNAs and to the establishment of cancer genetics. Retroviral reverse transcriptase has enhanced molecular biology research. Retroviruses are used today as gene therapy vectors. In addition, endogenous retroviruses make up 8% of the human genome but their functions are largely unknown.

Retroviruses are positive strand RNA viruses that reverse transcribe their virion RNA genomes to double-stranded DNA in the cytoplasm of an infected cell. Integration of this DNA copy into the host genome is necessary for efficient viral gene expression. After

[*]Correspondence to: KLB@jhu.edu.
[†]Present address: Analytic Services, Falls Church, VA, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

integration the viral DNA is called the provirus. Reverse transcription and integration, carried out largely by viral encoded and packaged enzymes, allow the virus to use host transcriptional and posttranscriptional machinery for viral gene expression-but with some unconventional twists. The integrated provirus also ensures stable maintenance of the viral genome as part of the host genome as long as the infected cell is alive. Thus, the host DNA polymerase replicates the proviral sequences as part of the host chromosomes. In addition, cellular RNA polymerase II (Pol II) transcribes a full-length primary viral RNA transcript, i.e., capped, polyadenylated, and spliced by cellular factors.

Retroviruses can be classified as either simple or complex viruses; however, both types have genomes ranging from about 7 to 9 kb. The major difference between them is that the complex viruses, such as human immunodeficiency virus-1 (HIV-1) and human T-lymphotropic virus-1 (HTLV-1), encode accessory proteins. These are generated from the primary viral transcripts by alternative splicing, generating additional proteins from alternative reading frames without expanding the genome size. These accessory proteins help influence the virus's infectivity, virulence, and aspects of the viral life cycle, including transcription, RNA export from the nucleus, and protection against host cell antiviral mechanisms. In contrast, the simple retroviruses, such as Rous sarcoma virus (RSV) and murine leukemia virus (MLV), rely on cis-acting viral RNA sequences that interact with host proteins for posttranscriptional regulation of gene expression.

Unlike cellular mRNAs, which are usually spliced to completion before export from the nucleus, a large fraction of both simple and complex retroviral transcripts must remain unspliced but still be exported to the cytoplasm. This unspliced RNA serves as both the genomic RNA, i.e., packaged into progeny viral particles and also as the predominant viral mRNA that encodes the *gag* (structural proteins) and *pol* (enzymes) gene products. This unusual use of unspliced RNA as mRNA requires some unconventional uses of cellular machinery.

First, splicing is incomplete. All retroviruses must divide their primary transcripts into a pool of completely unspliced RNA and a pool of singly spliced *env* (glycoprotein) mRNA. In addition, another fraction of the complex viral primary transcripts are alternatively spliced to generate many additional mRNAs. Thus, the viral primary RNA transcripts have three roles: genomic RNA, mRNA, and pre-mRNA. Second, the unspliced viral RNA must be exported from the nucleus to the cytoplasm, bypassing cellular mechanisms that prevent export of intron-containing RNAs. Different retroviruses have evolved different mechanisms to carry out this process. Some rely on *cis*-acting viral RNA elements that interact with host proteins, while others encode viral accessory proteins that interact with viral RNA and cellular proteins. Third, the unspliced RNA must be translated, which involves either frame-shifting or nonsense codon suppression to generate both *gag* and *pol* protein products, in different amounts, from one mRNA. Finally, the RSV unspliced mRNA is subject to surveillance by the nonsense-mediated mRNA decay (NMD) pathway. In contrast, NMD of cellular mRNAs is usually thought to require splicing, with the concomitant generation of exon-junction complexes, in higher vertebrates. The unspliced viral mRNA has an unusual structure with a long 3′ untranslated region (UTR), which often destabilizes cellular mRNAs, yet the viral RNAs are quite stable. Although, retroviral gene expression is somewhat unconventional, it is very efficient; as much as 5–10% of all cellular polyadenylated RNA is viral RNA in some infected cells.

RETROVIRAL GENOMES

The genomes of simple retroviruses like MLV and Mason-Pfizer monkey virus (MPMV) encode only structural and enzymatic viral proteins. These genomes are organized as: 5′–

gag-pol-env–3′, where *gag* generates the structural proteins (matrix, capsid, and nucleocapsid), *pol* encodes the viral enzymes (protease, reverse transcriptase, and integrase), and *env* synthesizes the viral envelope glycoproteins that are embedded in the membrane. In contrast, the avian viruses, such as avian leukosis virus (ALV) and RSV, synthesize protease as part of the Gag polyprotein. In addition, short cis-acting RNA regulatory sequences either flank the coding regions (e.g., PCE, CTE, DR) or are embedded within open reading frames (NRS and RSE). In addition, the simple avian virus, RSV, has a transduced cellular oncogene, *src*, i.e., located downstream of *env* in the genome; this virus is unusual in having an oncogene but still being competent for replication (Figure 1).[1]

There are also many replication-defective oncogenic retroviruses, isolated mainly from chickens, mice, and cats, which have acquired an oncogene from their host (e.g., myc, ras, myb, fes, fos, jun, ski, ets, erbB). These viruses have lost some of the genes necessary for replication (part of *gag* is usually retained), but they have maintained their cis-acting RNA packaging sequence called *psi*, located near the 5′ end of the primary RNA transcript.[2] These viruses are transmitted with the help of a nondefective 'helper virus', such as ALV, MLV, or feline leukemia virus (FeLV) that supplies viral proteins in *trans*. The concept of an RNA packaging sequence, together with helper proteins supplied in *trans*, provided the basis for use of retroviruses as gene therapy vectors.

In addition to the *gag*, *pol*, and *env* genes shared with simple retroviruses, complex retroviruses encode numerous accessory proteins in alternative reading frames (Figure 2) (Table 1). These accessory proteins include Tat and Tax, important for transcriptional regulation, and Rev and Rex, which promote export and translation of the incompletely spliced viral RNAs. In addition, the HIV-1 accessory protein Vif prevents editing of the HIV-1 genome by APOBEC3G, by targeting it for ubiquitination and degradation. The complex retroviruses also have cis-acting RNA regulatory sequences; most notable are the *trans*-activating region (TAR) at the start of transcription that binds the viral Tat protein, and the Rev Response Element (RRE) within the *env* gene that binds Rev (Figure 2).

THE RETROVIRAL LIFE CYCLE

Retroviruses are enveloped viruses with a protein capsid core and an internal nucleocapsid that surrounds the viral genome: two identical molecules of (+)-stranded RNA. Interactions between cell surface receptors and the viral envelope proteins mediate viral entry into the host cell. Upon binding, the viral and cell membranes fuse, releasing the viral capsid into the cytoplasm.[1]

Once the capsid enters the cytoplasm, the genomic RNA is reverse transcribed by the packaged viral reverse transcriptase. A specific packaged cellular tRNA (varies with type of retrovirus) binds the viral genomic RNA and serves as the primer for synthesis of the first DNA strand. RNase H activity of the reverse transcriptase degrades the RNA template after transcription but leaves behind a fragment of RNA that serves as the primer for the second DNA strand. Because the two primers are inset from each end, the full-length double-stranded DNA provirus is a few hundred nucleotides longer than the RNA template and contains a long terminal repeat (LTR) at each end. The LTRs have *cis*-acting enhancer and promoter sequences that drive viral transcription, as well as sequences required for 3′ end processing and polyadenylation.[1] Since retroviruses generate their own sequences for transcription initiation and termination, flanking the RNA genome, they are not restricted in finding productive integration sites.

Upon completion of reverse transcription, the retroviral DNA complex, coupled with capsid proteins and the viral integrase forms the pre-integration complex (PIC). The PIC travels to the nucleus, where the packaged viral integrase protein facilitates insertion of the viral

genetic material into the host DNA.[1] While integration is fairly random, different viruses prefer to integrate into different types of DNA sequences.[3,4] For example, HIV-1 tends to integrate into active transcription units,[5] while MLV preferentially targets transcription start sites and CpG islands.[6] Avian retroviruses, in contrast, show no strong preference for integration sites.[3,4] After integration, the provirus is dependent upon cellular processes for gene expression.

VIRAL RNA SYNTHESIS

After integration the provirus is transcribed by Pol II, driven by viral promoters and enhancers in the LTR upstream of the transcription start site. The simple retroviruses, such as RSV, have strong enhancers. In contrast, HIV-1 transcription is relatively weak unless aided by a viral accessory protein that enhances its transcription. One of the first HIV-1 genes expressed after integration is the *trans*-activator of transcription (Tat) protein, which is localized to the nucleus where it can bind a U-rich bulge in the HIV nascent TAR RNA (Figure 2).[7] The TAR sequence forms a stable stem-loop structure at the 5′ end of HIV RNA. Upon binding to TAR, Tat associates with and activates the cellular cofactor P-TEFb. The Tat/P-TEFb complex dissociates negative elongation factors from TAR and phosphorylates the CTD of Pol II, which promotes elongation.[8]

Since retroviruses have identical 5′ and 3′ LTRs, bearing transcriptional regulatory sequences, it seemed possible that transcription would also be initiated at the 3′ LTR. This would generate transcripts containing cellular sequences located downstream of the viral insertion site. However, promoter occlusion appears to prevent this. Transcription complexes initiated at the 5′ LTR exert transcriptional interference on the 3′ LTR, disrupting the ability of transcription complexes to form there. In some tumors transcription initiates from the ALV 3′ LTR and reads into downstream oncogenes, due to mutations near the 5′ LTR.[10,11] It is possible that some endogenous retroviruses may also initiate transcription from the 3′ LTR or read through the 3′ poly (A) site into downstream cellular sequences.

5′ CAPS AND INTERNAL METHYLATION

Like all Poll II transcripts, the nascent viral RNA is modified with a 5′ 7-methylguanosine (m7G) ‘cap’ shortly after the start of transcription. This unusual 5′-5′ linkage confers protection from 5′→3′ exonucle-ase degradation and enhances translation.[12,13] Recent studies show that a subset of HIV-1 Rev-dependent RNAs (unspliced and singly spliced) have a trimethyl-guanosine (TMG) cap, as do snRNAs, snoRNAs, and telomerase RNAs.[14] It appears that the Rev protein binds to the RRE on HIV-1 RNA in the nucleus and recruits the cellular methyltransferase, PIMT (peroxisome proliferator-activated receptorinteraction protein with methyltransferase), which hypermethylates these caps. The TMG cap promotes expression and export of HIV RNAs.[14] Since all of the RNAs with TMG caps are transported from the nucleus to the cytoplasm by CRM1,[15] TMG caps could be a marker for the CRM1 pathway.

Many mRNAs are also internally methylated at m^6A residues to varying extents. Early work on RSV RNAs showed the genomic RNA contained 12 sites of m^6A methylation at a consensus sequence (PuGACU)[16] similar to that seen recently in global studies of m^6A in cellular mRNAs.[17] Mutation of the consensus sequence in RSV blocked this methylation but did not detectably affect viral replication.[18]

RETROVIRAL mRNA SPLICING

Duringx2 transcription, a fraction of the viral RNA is subjected to splicing. In general, retroviruses have strong 5′ splice sites and weak 3′ splice sites.[19,20] Retroviral splicing is

driven by the same factors as the splicing of cellular mRNAs, including the splicing snRNPs U1, U2, U4/U6, and U5, other splicing enhancers and suppressors, like SR proteins (ASF/SF2, U2AF, SRp40, and SC35), and several hnRNPs. Given the conserved splice site sequences within the viral RNA, combined with a relatively small transcript, retroviruses are an attractive model for studying RNA splicing.

All retroviruses have at least one spliced mRNA, encoding the *env* product, but are incompletely spliced so that the unspliced primary transcript is retained (Figures 1 and 2). Simple retroviruses employ suboptimal 3′ splice sites as well as splicing enhancers and suppressors[21,22] to regulate their splicing. For example, in RSV, the env 3′ splice has a weak branch point.[23] Improving the branch point sequence results in over-splicing with a concomitant block in viral replication,[21] emphasizing the importance of incomplete splicing for viral replication.

Two other cis-acting sequences that affect splicing of RSV RNA have been characterized. The suppressor of *src* splicing (SSS) is located between *env* and *src* and decreases the amount of spliced *src* mRNA.[24] The SSS effect can be abrogated by addition of competitor RNA, and SSS sequences are unable to inhibit splicing in mammalian cells.[24] Thus, its activity is not mediated via internal base-pairing between the element and a 3′ splice site, and it is presumed instead to interact with an unknown splicing factor specific to avian cells. A second element called the negative regulator of splicing (NRS) is located in the RSV *gag* gene.[25]

RSV Negative Regulator of Splicing

RSV RNA recruits components of the major (U1/U2 snRNPs) and minor (U11/U12 snRNPs) spliceosomes to a 5′ splice site-like sequence within *gag* termed the NRS (Figure 1).[25,26] A splicing construct containing the NRS and a downstream 3′ splice site forms an aberrant splicesomal complex *in vitro* but does not splice.[25,26] This complex contains all five of the major splicing snRNPs, as well as U11 and U12 snRNPs, but the complex is relatively unstable and can be disrupted by the addition of heparin.[26] Furthermore, the U5 snRNP scaffold protein Prp8 cannot be cross-linked to the 5′ splice site, suggesting an aberrant arrangement of snRNPs.[26] The NRS is thought to sequester the 3′ splice site away from the authentic 5′ splice site. Mutation of the NRS diminishes viral replication *in vivo* about twofold and results in a fivefold increase in spliced *src* mRNA but does not alter *env* mRNA levels.[27]

Surprisingly, NRS mutations also increase read through of the 3′ viral poly(A) site from 15 to 25%.[27] Large deletion mutations that remove the NRS cause a similar effect.[28] *In vivo*, the NRS likely acts to inhibit read through and splicing to downstream genes, which can result in their activation. Consistent with this, ALV with a single point mutation in the U1 binding site of the NRS causes lymphomas in chickens much more rapidly than wild type virus, due to activation of downstream oncogenes.[29] In addition, read through of the poly(A) site is thought to be responsible for the generation of transducing viruses like RSV. *In vitro* studies of the mechanism of NRS inhibition of polyadenylation have suggested the importance of SR proteins binding upstream of the pseudo 5′ splice site.[30] In addition, the complex formed between the NRS and viral 3′ splice sites promotes polyadenylation.[31]

HIV-1 RNA Splicing

While the pre-mRNAs of both simple and complex retroviruses are roughly the same size (RSV is 9312 nt while HIV-1 is 9171 nt), extensive alternative splicing of HIV-1 pre-mRNA generates seven additional regulatory proteins (Table 1). HIV-1 has at least four 5′ splice sites and eight 3′ splice sites and generates at least 42 distinct spliced mRNAs.[32] Although

seven different mRNAs would be sufficient to encode all of the HIV-1 proteins, there are many splice variants. These HIV-1 mRNAs vary greatly in abundance and appear to be controlled by *cis*-acting splicing elements. The major HIV-1 5′ splice site, located 288 nucleotides downstream of the transcription start site, is used to generate all spliced mRNAs, while the other three 5′ splice sites are used to make accessory proteins from multiply spliced mRNAs.[32]

HIV-1 splicing is very complex and is regulated by many exonic and intronic splicing enhancers, which typically bind SR proteins, and intronic and exonic splicing silencers, which usually bind hnRNPs. For example, the dominance of the major 5′ splice site has been linked to nearby hnRNP H binding *in vitro*,[33] as well as enhancer elements within the leader sequence and a proximal intronic splicing enhancer.[34] Furthermore, a 20-nucleotide exonic splicing silencer resides within the first coding exon of *tat*. This region suppresses the use of the upstream 3′ splice site, resulting in increased usage of the downstream sites that generate *rev* and *env* mRNAs.[35] A more extensive review of HIV-1 gene expression has been published recently.[36]

HIV-1 undergoes temporal regulation of splicing, which results in the generation of three classes of viral mRNAs (Figure 1).[1] The first class, informally known as the 2 kb class, is the dominant species expressed early in the viral life cycle. This doubly spliced class of mRNAs encodes only accessory proteins that drive viral replication and infectivity. These transcripts include *tat* (promotes viral transcription), *rev* (binds a *cis*-acting element within the viral RNA to promote the nuclear export of incompletely spliced mRNAs), and *nef* (negative regulatory factor that promotes CD4 and CD28 downregulation as well as the survival of infected cells).[1]

The second class, known as the 4 kb class, comprises four singly spliced mRNAs expressed predominantly during the middle stages of HIV infection. *env* encodes the viral envelope proteins on the exterior of mature viral particles.[1] *vif* (viral infectivity factor) mRNA encodes an accessory protein that inhibits the host viral restriction factor APOBEC3G, which deaminates the viral genome during reverse transcription.[37] *vpr* (viral protein R) mRNAs encode an accessory protein that promotes viral replication in nondividing cells, guides the PIC to the nucleus, and regulates cellular apoptotic function. Lastly, *vpu* (viral protein unique) targets CD4-tetherin complexes for degradation, which in turn promotes the release of progeny virions from an infected cell.

The third class, known as the 9 kb class, is the dominant mRNA species expressed during late-stage infection. This class contains only one mRNA: the full-length, unspliced viral RNA. It encodes both the viral structural proteins (Gag) and the viral enzymes (Pol). This RNA species also serves as the genomic RNA, i.e., packaged as a dimer into progeny virions.

Recently, single molecule amplification and long-read sequencing was used to analyze HIV splicing. In this study, 109 different spliced RNAs were identified, with 13 novel splice sites (2 5′ splice sites and 11 3′ splice sites) and a new 1 kb class of viral RNAs. Temporal changes in these transcripts could not be explained by accumulation of a viral export factor (Rev), but may be due either to differences in transcript stability or splicing. This study suggests that splice sites may be evolving in the HIV quasi-species, potentially in response to selective pressure.[38]

VIRAL RNA POLYADENYLATION

Termination of retroviral transcripts occurs by polyadenylation like most cellular mRNAs. The poly(A) signal AAUAAA is present in all retroviral LTRs except for the mouse

mammary tumor virus (MMTV), which uses a near-consensus AGUAAA. These sequences are recognized by CPSF, the cleavage, and polyadenylation specificity factor, which recruits the cleavage stimulation factor (CStF) and cleavage factors I and II (CFI, CF II). These factors then recruit the polyadenylation polymerase (PAP). The RNA is cleaved between the poly(A) signal sequence and a GU-rich motif, and PAP begins polymerizing adenosine residues at the 3′ end of the RNA. After 12–15 residues have been added, poly(A) binding protein N1 (PABPN1) binds the tail, further synergizing adenosine polymerization, resulting in the addition of 200–250 nucleotides.[13]

Some retroviruses have the poly(A) signal in the R (repeat) region of the LTR and thus have potential poly(A) sites at both ends of the viral transcript. These viruses employ various mechanisms to prevent polyadenylation at the 5′ end of the transcript. For example, RSV RNA is only polyadenylated at the 3′ end because of necessary poly(A) sequence elements found only in U3, upstream of the poly(A) site at the 3′ end of the RNA transcript. Similarly, the HIV-1 3′ poly(A) site is occluded by a hairpin structure that is thought to be opened by factors that bind upstream of the site.[39] Interestingly, binding of U1 snRNP to the first 5′ splice site in HIV-1 RNA inhibits polyadenylation at the 5′ end of the transcript.[40] In addition, many cryptic poly(A) sites in cellular mRNAs are thought to be inhibited by U1 snRNP binding.[41]

EXPORT OF UNSPLICED RETROVIRAL mRNA FROM THE NUCLEUS

A major problem that all retroviruses must overcome involves export of their unspliced RNA from the nucleus to the cytoplasm. In addition, the complex retroviruses must also export their incompletely spliced 4 kb mRNAs from the nucleus. In contrast, intron-containing cellular RNAs are not usually exported from the nucleus.[42] Nuclear retention of incompletely spliced mRNAs is caused, in part, by cellular factors that bind unused 3′ and 5′ splice sites.[43] This general rule ensures that mRNAs are export- and translation-competent before being shuttled through the nuclear pore complex. Retroviruses, however, require export of a full-length, unspliced RNA for use as the Gag and Pol mRNA as well as for the genomic RNA packaged into new virions. Thus, these viruses have developed ways to evade the restriction on export of unspliced RNA.

After splicing and polyadenylation, fully processed cellular mRNAs are exported to the cytoplasm by recruitment of REF/Aly, which then directly interacts with the nuclear export factor Tap/NXF1.[44,45] With the aid of hypophosphorylated SR proteins (SRp20, 9G8 and ASF/SF2), Tap/NXF1 and its cofactor p15/NXT1 facilitate the export of the fully spliced RNA to the cytoplasm.[46,47] The DEAD box helicase Dbp5 shuttles between the nucleus and cytoplasm, but it is located at higher concentration around the nuclear rim.[48] Dbp5 interacts with CAN/Nup159 and other nuclear shuttling factors at the nuclear fibrils and can strip proteins off the emerging mRNP complex, including Tap/NXF1.[48,49]

Complex retroviruses encode an accessory protein (Rev for HIV and Rex for HTLV) to aid in the nuclear export of unspliced (9 kb) and incompletely spliced (4 kb) RNA.[1] These Rev/Rex proteins are generated from doubly spliced (2 kb) mRNAs and are imported to the nucleus of infected cells (Figure 2). Upon relocation, these proteins bind to specific RNA elements [RRE or Rex-responsive element (RexRE)] on singly spliced and unspliced RNA to facilitate export. For example, the HIV-1 RRE is a 351 nt RNA element with an elongated stem-loop structure. Rev binding causes a conformational change in the RRE that allows binding of additional Rev proteins to generate oligomers. To facilitate export, Rev or Rex then bind to CRM1 (exportin 1), a cellular export receptor that typically exports proteins containing a leucine-rich nuclear export signal (NES) as well as snRNAs.[15,50] The DEAD box helicase DDX3 binds CRM1 to localize it to the nuclear pore complex (NPC).[51] Then,

CRM1, coupled with RanGTP, transports the viral ribonucleoprotein (RNP) through the NPC. After transport, RanGTP is hydrolyzed, and CRM1 releases the mRNP complex into the cytoplasm (Figure 2). The Rev/RRE interaction starts in the nucleus and promotes export[52]; Rev also promotes translation of HIV RNAs[53] and their encapsidation in the cytoplasm,[54] suggesting multiple roles for this nuclear shuttling protein.

In the absence of Rev, multiple *cis*-acting repressive sequences (CRS) or instability sequences (INS) impair stability, export, and translation of the unspliced and 4 kb HIV-1 mRNAs.[55] When these AU-rich CRS elements are mutated, viral gene expression does not require Rev. The INS elements in HIV-1 RNA associate with p54nrb/PSF *in vivo*, and this inhibits viral expression.[56] Recently, the 23 kb long noncoding RNA, NEAT1 which is found in paraspeckles along with p54nrb/PSF, has been shown to be upregulated by HIV-1 infection. When NEAT1 is knocked down, HIV-1 unspliced RNA export is enhanced.[57]

MMTV also encodes a Rev-like accessory protein, termed Rem, which binds a viral RNA element, RmRE.[58,59] Identification of Rem led to classification of MMTV as the first complex murine retrovirus.

Simple retroviruses do not encode accessory proteins and must rely on *cis*-acting elements within their RNA structure to induce export of the unspliced RNA. MPMV utilizes a nucleotide sequence in a noncoding region near the 3′ end of the unspliced RNA called the constitutive transport element (CTE)[60]. This sequence directly binds Tap/p15, which exports the unspliced RNA.[61] A somewhat analogous element has been found in RSV. There, the direct repeats (DR1 and DR2) are two 100-nucleotide, redundant sequences that flank the *src* gene (Figure 1). Eliminating both DR copies blocks unspliced RNA accumulation in the cytoplasm and packaging of genomic RNA in particles, thus inhibits viral replication.[62] The related ALV contains a single copy of the DR; RSV likely acquired an additional copy when it transduced the *src* gene. Each DR forms a highly stable stem-loop structure critical for its function.[63] Point mutations that disrupt the structure virtually eliminate the export of the unspliced RNA, and compensatory mutations that restore base pairing also restore export. Though RSV DRs do not appear to bind Tap directly,[64] both Tap and the DEAD box helicase Dbp5 are necessary for unspliced viral RNA export (Figure 1).[65] Although spliced RSV mRNAs also contain the DR, mutations in the DR do not affect spliced mRNA export, suggesting that these RNAs go through the normal export process for spliced mRNAs.

THE CYTOPLASMIC FATE OF RETROVIRAL RNAS

Once in the cytoplasm, the retroviral RNAs can finally be translated. Typically, translation initiation factors assemble near the 5′ cap of cellular RNAs and scan the transcript in search of the nearby Kozak sequence (RCCAUGG) that defines the initiating methionine codon. However, retroviruses have rather long 5′ UTRs, which contain structured sequences necessary for dimerization and packaging of RNA into viral particles.[1,2] In addition, most retroviruses have a cap-dependent translation enhancer called a posttranscriptional control element (PCE) near the 5′ end of their RNA transcript. The PCE interacts with RNA helicase A and enhances polysome loading and Gag protein production.[66] In addition, many retroviruses have been reported to have an internal ribosome entry site (IRES),[67] but this work remains controversial.[66]

Some retroviruses contain small upstream open reading frames (uORF) prior to the *gag* initiation site. RSV, e.g., contains three uORfs. Mutational analysis has demonstrated that these short open reading frames are important for viral gene expression and replication.[68,69] After synthesis of these small peptides and translation termination, translation must re-initiate at the *gag* initiation codon to generate the Gag and Gag-Pol polypeptides. MLV has an alternative CUG initiation codon upstream of the *gag* AUG initiation codon, which is

used part of the time to generate a glycosylated Gag protein called glycoGag. If this CUG is mutated to block GlycoGag formation, the virus does not generate infectious virus.[70] In addition, leaky scanning past a weak initiation site is used as another mechanism to generate multiple protein products from a bicistronic mRNA, such as the HIV-1 *vpu*/*env* mRNA.[71]

Surprisingly, HIV-1 unspliced and singly spliced mRNAs in polysomes are predominantly bound by the nuclear cap-binding proteins (CBC) (CBP80/CBP20)[72] rather than eIF4E, which binds the caps of most cellular mRNAs in the cytoplasm after the pioneer round of translation. HIV-1 infection can induce G2/M cell cycle arrest, modulated by vpR, with associated decreases in cellular translation due to reduction in phosphorylated, active eIF4E. HIV-1 is able to maintain translation of its mRNA, and has a competitive advantage over translation of cellular mRNAs, by using the CBC instead of eIF4E.[72]

Termination of *gag* Translation and the Generation of Gag-Pol and Gag-Pro-Pol Products

Gag and Gag-Pol polyproteins are synthesized from the same unsplicedm RNA transcript (Figures 1 and 2). Because viral replication requires many more copies of structural proteins than viral enzymes, retroviruses have developed regulatory mechanisms that maintain the proper ratio of Gag to Gag-Pol products. The precise mechanism for maintaining this balance varies among retroviruses. However, the primary method used by most retroviruses, including RSV and HIV-1, involves ribosomal frameshifting.[1] This frameshift occurs just before the RSV *gag* termination codon. A downstream pseudoknot causes the ribosome to pause on a 'slippery sequence,' typically consisting of an A–U rich region, which allows the ribosome to slip back one nucleotide into a different reading frame (−1). The shift places the *gag* termination codon out of frame, and translation continues into the downstream *pol* sequences. Several other retroviruses, such as HTLV, MMTV, and MPMV, undergo two frameshifting events (Table 1). The first circumvents the *gag* termination codon, placing the ribosomes into the *pro* (protease) reading frame. The second shift places the ribosomes into the *pol* reading frame.[1]

Alternatively, other retroviruses use a method of termination codon suppression to regulate levels of Gag and Pol polyprotein production. MLV, FeLV, and Walleye dermal sarcoma virus (WDSV) each insert a glutamine tRNA at the *gag* stop codon during a small fraction of translation events and read through into the *pol* gene to generate Gag-Pol polyproteins.[73,74] In these genomes, *gag* and *pol* are in the same reading frame but separated by a termination codon.

One exception to the rule regarding translation and the generation of Gag-Pol polyprotein products is observed with foamy virus, a spumaretrovirus. Uniquely, this virus uses alternative splicing to generate separate transcripts that produce Pro and Pol products,[75] avoiding the need for either translational read through or frameshifting.

Retroviruses and Nonsense-Mediated mRNA Decay

Though cellular processes of transcription, splicing, and translation are tightly regulated, occasional errors can occur. For example, a premature termination codon (PTC) in a transcript results in a truncated protein coding sequence. Most eukaryotic cells use the NMD pathway to eliminate these transcripts and prevent production of potentially deleterious truncated proteins.[76] The NMD pathway also regulates the levels of some normal transcripts. In eukaryotic cells there appear to be two processes for inducing NMD. The first involves exon–exon junction complexes (EJCs), and is the method predominantly used in mammalian cells. EJCs are deposited on spliced mRNAs 20–24 nucleotides upstream of introns during splicing.[77] As the ribosome translates the mRNA, it displaces all EJCs in its path. Upon encountering a PTC, however, the ribosome stops, leaving any downstream

EJCs intact. The presence of an EJC downstream of a termination event is thought to trigger the NMD pathway.[76]

Alternatively, in *Saccharomyces cerevisiae* a ribosome that reaches the natural termination codon interacts with an RNP complex located in the 3′ UTR. Upon reaching a PTC, however, the terminating ribosome is generally too far away to interact with this 3′ UTR complex, and the mRNA is targeted for decay-a mechanism known as the 'faux UTR' model.[78] Typically, the normal 3′ UTR is short and the poly(A) sequence and PABP are near the authentic termination codon.[78] Despite their differences, both of these NMD mechanisms rely on the same factors. First, translation is required. NMD in all systems is suppressed when cells are treated with translation inhibitors. Second, the Upf (upframeshift) proteins 1, 2, and 3, along with the Smg proteins (1, 5, 6, and 7) are required.[76,79]

At least some retroviruses are also susceptible to NMD. PTCs in the *gag* gene of unspliced RSV RNA cause degradation of the full-length RNA.[80,81] Both Upf1 and translation are required for this decay.[82] However, since the viral RNA is unspliced, it is unlikely that EJCs are deposited, and therefore unlikely that the RSV RNA goes through EJC-mediated NMD. Termination at the normal *gag* termination codon in the RSV unspliced RNA generates a ~7kb 3′ UTR and would be expected to make these mRNAs a target for the NMD pathway. Nevertheless, the unspliced viral RNAs are very stable with a half-life greater than 10 h.

Recent studies have begun to unravel the mechanism by which some retroviruses might be immune to NMD upon proper termination of the *gag* gene. Sequences (~400 nt) located immediately downstream of the RSV natural *gag* termination codon appear to protect the RNA from NMD and are termed the RSV stability element (RSE).[83] Moving the RSE to a location just 3′ of a PTC also stabilizes the RNA. The secondary structure of the RSE includes several critical structures.[84,85] To date, it remains unclear which, if any, cellular factors bind the RSE to facilitate stability or whether the RSE forms RNA–RNA or protein–protein mediated long range interactions that bring the poly(A) tail into close proximity with the terminating ribosome to stabilize the RNA. Generation of a looped 3′ UTR with a cellular reporter has been shown to stabilize mRNAs in other systems.[86]

Alternatively, the RSE may bind an inhibitor of NMD. Since it has been reported that Upf1 binds to mRNAs with long 3′ UTRs,[87] it is possible that the RSE disrupts this binding. Surprisingly, Upf1, together with Staufen, appears to play a positive role in HIV-1 gene expression.[88] HTLV-1 inhibits NMD of its viral RNAs by a different mechanism. Its trans-activator Tax interacts with Upf1 and INT6/EIF3E to inhibit NMD in infected cells.[89] In addition, the HTLV-1 Rex protein has been reported to downregulate NMD of viral and cellular mRNAs in infected cells.[90] The structure of the normal HTLV-1 genome is thought to make it a target for NMD.[90]

CONCLUSIONS

Retroviruses make unconventional use of host cell machinery to regulate their posttranscriptional gene expression. Once their RNA genomes are reverse transcribed and integrated into the host DNA, they use cellular machinery for transcription of primary RNA transcripts. These RNAs are incompletely spliced, and a fraction of these unspliced transcripts are exported to the cytoplasm, translated, and monitored by NMD. In addition, a fraction of all retroviral transcripts are spliced once to generate the *env* mRNA, and the complex viruses are spliced many times to generate accessory proteins. Some of these proteins regulate posttranscriptional RNA processing. Translation of the viral RNA entails ribosomal frameshifts or termination codon suppression to generate different amounts of structural proteins and enzymes from a single unspliced mRNA. HIV-1 unspliced RNAs

often have trimethylated G caps, and these bind the nuclear CBP80 protein instead of eIF4E. This permits viral translation while cell translation is repressed by a G2/M cell cycle block that inactivates eIF4E. One of the host defense mechanisms against HIV-1 and other retroviruses involves the cytosine deamination proteins in the APOBEC family. APOBEC3G is packaged into HIV-1 particles and deaminates the viral genome during reverse transcription. The virus counters this restriction with its Vif accessory protein, which is also packaged and targets APOBEC3G for ubiquitylation. Further, in mice APOBEC3 provides partial protection against MMTV and MLV infection and pathogenesis *in vivo*.[91]

Acknowledgments

We thank Kirsten Birst for assistance with the figures, and Scott Johnson for review of the manuscript. This work was supported by NIH grants R01CA048746 and P50 GM103297.

REFERENCES

1. Coffin, JM.; Hughes, SH.; Varmus, HE. Retroviruses. New York: Cold Spring Harbor Laboratory Press; 1997.
2. Shank PR, Linial M. Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: characterization of a deletion in the provirus. J Virol. 1980; 36:450–456. [PubMed: 6253670]
3. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, Collins F, Shinn P, Leipzig J, Hannenhalli S, et al. Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS Pathog. 2006; 2:e60. [PubMed: 16789841]
4. Narezkina A, Taganov KD, Litwin S, Stoyanova R, Hayashi J, Seeger C, Skalka AM, Katz RA. Genome-wide analyses of avian sarcoma virus integration sites. J Virol. 2004; 78:11656–11663. [PubMed: 15479807]
5. Liu H, Dow EC, Arora R, Kimata JT, Bull LM, Arduino RC, Rice AP. Integration of human immunodeficiency virus type 1 in untreated infection occurs preferentially within genes. J Virol. 2006; 80:7765–7768. [PubMed: 16840357]
6. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003; 300:1749–1751. [PubMed: 12805549]
7. Berkhout B, Silverman RH, Jeang KT. Tat trans-activates the human immunodeficiency virus through a nascent RNA target. Cell. 1989; 59:273–282. [PubMed: 2478293]
8. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol Cell Biol. 2004; 24:787–795. [PubMed: 14701750]
9. Cullen BR, Lomedico PT, Ju G. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. Nature. 1984; 307:241–245. [PubMed: 6363938]
10. Hayward WS, Neel BG, Astrin SM. Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature. 1981; 290:475–480. [PubMed: 6261142]
11. Goodenow MM, Hayward WS. 5′ long terminal repeats of *myc*-associated proviruses appear structurally intact but are functionally impaired in tumors induced by avian leukosis viruses. J Virol. 1987; 61:2489–2498. [PubMed: 3037111]
12. Mitchell SF, Walker SE, Algire MA, Park EH, Hinnebusch AG, Lorsch JR. The 5′-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. Mol Cell. 2010; 39:950–962. [PubMed: 20864040]
13. Shatkin AJ, Manley JL. The ends of the affair: capping and polyadenylation. Nat Struct Biol. 2000; 7:838–842. [PubMed: 11017188]
14. Yedavalli VS, Jeang KT. Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. Proc Natl Acad Sci USA. 2010; 107:14787–14792. [PubMed: 20679221]
15. Cullen BR. Nuclear mRNA export: insights from virology. Trends Biochem Sci. 2003; 28:419–424. [PubMed: 12932730]

16. Kane SE, Beemon K. Precise localization of m^6A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. Mol Cell Biol. 1985; 5:2298–2306. [PubMed: 3016525]
17. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, et al. Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature. 2012; 485:201–206. [PubMed: 22575960]
18. Kane SE, Beemon K. Inhibition of methylation at two internal N^6-methyladenosine sites caused by GAC to GAU mutations. J Biol Chem. 1987; 262:3422–3427. [PubMed: 3029112]
19. McNally MT, Beemon K. Intronic sequences and 3′ splice sites control Rous sarcoma virus RNA splicing. J Virol. 1992; 66:6–11. [PubMed: 1309264]
20. O'Reilly MM, McNally MT, Beemon KL. Two strong 5′ splice sites and competing, suboptimal 3′ splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. Virology. 1995; 213:373–385. [PubMed: 7491762]
21. Katz RA, Skalka AM. Control of retroviral RNA splicing through maintenance of suboptimal processing signals. Mol Cell Biol. 1990; 10:696–704. [PubMed: 2153921]
22. Zhang L, Stoltzfus CM. A suboptimal *src* 3′ splice site is necessary for efficient replication of Rous sarcoma virus. Virology. 1995; 206:1099–1107. [PubMed: 7856084]
23. Fu XD, Katz RA, Skalka AM, Maniatis T. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. 1991; 5:211–220. [PubMed: 1847346]
24. Amendt BA, Simpson SB, Stoltzfus CM. Inhibition of RNA splicing at the Rous sarcoma virus *src* 3′ splice site is mediated by an interaction between a negative cis element and a chicken embryo fibroblast nuclear factor. J Virol. 1995; 69:5068–5076. [PubMed: 7609076]
25. Gontarek RR, McNally MT, Beemon K. Mutation of an RSV intronic element abolishes both U11/U12 snRNP binding and negative regulation of splicing. Genes Dev. 1993; 7:1926–1936. [PubMed: 8405999]
26. Giles KE, Beemon KL. Retroviral splicing suppressor sequesters a 3′ splice site in a 50S aberrant splicing complex. Mol Cell Biol. 2005; 25:4397–4405. [PubMed: 15899846]
27. O'Sullivan CT, Polony TS, Paca RE, Beemon KL. Rous sarcoma virus negative regulator of splicing selectively suppresses *src* mRNA splicing and promotes polyadenylation. Virology. 2002; 302:405–412. [PubMed: 12441084]
28. Miller JT, Stoltzfus CM. Two distant upstream regions containing cis-acting signals regulating splicing facilitate 3′-end processing of avian sarcoma virus RNA. J Virol. 1992; 66:4242–4251. [PubMed: 1318403]
29. Polony TS, Bowers SJ, Neiman PE, Beemon KL. Silent point mutation in an avian retrovirus RNA processing element promotes c-*myb*-associated shortlatency lymphomas. J Virol. 2003; 77:9378–9387. [PubMed: 12915553]
30. Wilusz JE, Beemon KL. The negative regulator of splicing element of Rous sarcoma virus promotes polyadenylation. J Virol. 2006; 80:9634–9640. [PubMed: 16973567]
31. Hudson SW, McNally MT. Juxtaposition of two distant, serine-arginine-rich protein-binding elements is required for optimal polyadenylation in Rous sarcoma virus. J Virol. 2011; 85:11351–11360. [PubMed: 21849435]
32. Purcell DF, Martin MA. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J Virol. 1993; 67:6365–6378. [PubMed: 8411338]
33. Schaub MC, Lopez SR, Caputi M. Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. J Biol Chem. 2007; 282:13617–13626. [PubMed: 17337441]
34. Asang C, Erkelenz S, Schaal H. The HIV-1 major splice donor D1 is activated by splicing enhancer elements within the leader region and the p17-inhibitory sequence. Virology. 2012; 432:133–145. [PubMed: 22749061]
35. Amendt BA, Si ZH, Stoltzfus CM. Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: evidence for inhibition mediated by cellular factors. Mol Cell Biol. 1995; 15:4606–4615. [PubMed: 7623852]

36. Karn J, Stoltzfus CM. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med. 2012; 2:a006916. [PubMed: 22355797]
37. Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature. 2002; 418:646–650. [PubMed: 12167863]
38. Ocwieja KE, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, Wang S, Link DR, Olson J, Travers K, et al. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. Nucleic Acids Res. 2010; 40:10345–10355. [PubMed: 22923523]
39. Das AT, Klaver B, Berkhout B. A hairpin structure in the R region of the human immunodeficiency virus type 1 RNA genome is instrumental in polyadenylation site selection. J Virol. 1999; 73:81–91. [PubMed: 9847310]
40. Ashe MP, Furger A, Proudfoot NJ. Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. RNA. 2000; 6:170–177. [PubMed: 10688356]
41. Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. U1 snRNP determines mRNA length and regulates isoform expression. Cell. 2012; 150:53–64. [PubMed: 22770214]
42. Erkmann JA, Kutay U. Nuclear export of mRNA: from the site of transcription to the cytoplasm. Exp Cell Res. 2004; 296:12–20. [PubMed: 15120988]
43. Chang DD, Sharp PA. Regulation by HIV Rev depends upon recognition of splice sites. Cell. 1989; 59:789–795. [PubMed: 2686839]
44. Stutz F, Bachi A, Doerks T, Braun IC, Séraphin B, Wilm M, Bork P, Izaurralde E. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA. 2000; 6:638–650. [PubMed: 10786854]
45. Zhou Z, Luo MJ, Straesser K, Katahira J, Hurt E, Reed R. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature. 2000; 407:401–405. [PubMed: 11014198]
46. Jin L, Guzik BW, Bor YC, Rekosh D, Hammarskjöld ML. Tap and NXT promote translation of unspliced mRNA. Genes Dev. 2003; 17:3075–3086. [PubMed: 14701875]
47. Huang Y, Yario TA, Steitz JA. A molecular link between SR protein dephosphorylation and mRNA export. Proc Natl Acad Sci USA. 2004; 101:9666–9670. [PubMed: 15210956]
48. Schmitt C, von Kobbe C, Bachi A, Panté N, Rodrigues JP, Boscheron C, Rigaut G, Wilm M, Séraphin B, Carmo-Fonseca M, et al. Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO J. 1999; 18:4332–4347. [PubMed: 10428971]
49. Lund MK, Guthrie C. The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Mol Cell. 2005; 20:645–651. [PubMed: 16307927]
50. Fornerod M, Ohno M, Yoshida M, Mattaj IW. CRM1 is an export receptor for leucine-rich nuclear export signals. Cell. 1997; 90:1051–1060. [PubMed: 9323133]
51. Yedavalli VS, Neuveut C, Chi YH, Kleiman L, Jeang KT. Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function. Cell. 2004; 119:381–392. [PubMed: 15507209]
52. Cochrane AW, McNally MT, Mouland AJ. The retrovirus RNA trafficking granule: from birth to maturity. Retrovirology. 2006; 3:18. [PubMed: 16545126]
53. Arrigo SJ, Chen IS. Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 vif, vpr, and env/vpu 2 RNAs. Genes Dev. 1991; 5:808–819. [PubMed: 1827422]
54. Blissenbach M, Grewe B, Hoffmann B, Brandt S, Uberla K. Nuclear RNA export and packaging functions of HIV-1 Rev revisited. J Virol. 2010; 84:6598–6604. [PubMed: 20427541]
55. Schneider R, Campbell M, Nasioulas G, Felber BK, Pavlakis GN. Inactivation of the human immunodeficiency virus type 1 inhibitory elements allows Rev-independent expression of Gag and Gag/protease and particle formation. J Virol. 1997; 71:4892–4903. [PubMed: 9188551]
56. Zolotukhin AS, Michalowski D, Bear J, Smulevitch SV, Traish AM, Peng R, Patton J, Shatsky IN, Felber BK. PSF acts through the human immunodeficiency virus type 1 mRNA instability elements to regulate virus expression. Mol Cell Biol. 2003; 23:6618–6630. [PubMed: 12944487]

57. Zhang Q, Chen C-Y, Yedavalli VS, Jeang KT. NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. mBio. 2013; 4:e00596–e00512. [PubMed: 23362321]
58. Indik S, Günzburg WH, Salmons B, Rouault F. A novel, mouse mammary tumor virus encoded protein with Rev-like properties. Virology. 2005; 337:1–6. [PubMed: 15914215]
59. Mertz JA, Simper MS, Lozano MM, Payne SM, Dudley JP. Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. J Virol. 2005; 79:14737–14747. [PubMed: 16282474]
60. Bray M, Prasad S, Dubay JW, Hunter E, Jeang KT, Rekosh D, Hammarskjold ML. 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 USA. 1994; 91:1256–1260. [PubMed: 8108397]
61. Pasquinelli AE, Ernst RK, Lund E, Grimm C, Zapp ML, Rekosh D, Hammarskjöld ML, Dahlberg JE. The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. EMBO J. 1997; 16:7500–7510. [PubMed: 9405378]
62. Ogert RA, Lee LH, Beemon KL. Avian retroviral RNA element promotes unspliced RNA accumulation in the cytoplasm. J Virol. 1996; 70:3834–3843. [PubMed: 8648719]
63. Yang J, Cullen BR. Structural and functional analysis of the avian leukemia virus constitutive transport element. RNA. 1999; 5:1645–1655. [PubMed: 10606274]
64. Paca RE, Ogert RA, Hibbert CS, Izaurralde E, Beemon KL. Rous sarcoma virus DR posttranscriptional elements use a novel RNA export pathway. J Virol. 2000; 74:9507–9514. [PubMed: 11000220]
65. LeBlanc JJ, Uddowla S, Abraham B, Clatterbuck S, Beemon KL. Tap and Dbp5, but not Gag, are involved in DR-mediated nuclear export of unspliced Rous sarcoma virus RNA. Virology. 2007; 363:376–386. [PubMed: 17328934]
66. Bolinger C, Boris-Lawrie K. Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. Retrovirology. 2009; 6:8. [PubMed: 19166625]
67. Balvay L, Lopez Lastra M, Sargueil B, Darlix JL, Ohlmann T. Translational control of retroviruses. Nat Rev Microbiol. 2007; 5:28–40.
68. Donzé O, Damay P, Spahr PF. The first and third uORFs in RSV leader RNA are efficiently translated: implications for translational regulation and viral RNA packaging. Nucleic Acids Res. 1995; 23:861–868. [PubMed: 7708504]
69. Moustakas A, Sonstegard TS, Hackett PB. Alterations of the three short open reading frames in the Rous sarcoma virus leader RNA modulate viral replication and gene expression. J Virol. 1993; 67:4337–4349. [PubMed: 7685415]
70. Low A, Datta S, Kuznetsov Y, Jahid S, Kothari N, McPherson A, Fan H. Mutation in the glycosylated gag protein of murine leukemia virus results in reduced in vivo infectivity and a novel defect in viral budding or release. J Virol. 2007; 81:3685–3692. [PubMed: 17267509]
71. Schwartz S, Felber BK, Pavlakis GN. Mechanism of translation of monocistronic and multicistronic human immunodeficiency virus type 1 mRNAs. Mol Cell Biol. 1992; 12:207–219. [PubMed: 1729599]
72. Sharma A, Yilmaz A, Marsh K, Cochrane A, Boris-Lawrie K. Thriving under stress: selective translation of HIV-1 structural protein mRNA during Vpr-mediated impairment of eIF4E translation activity. PLoS Pathog. 2012; 8:e1002612. [PubMed: 22457629]
73. Hatfield DL, Levin JG, Rein A, Oroszlan S. Translational suppression in retroviral gene expression. Adv Virus Res. 1992; 41:193–239. [PubMed: 1575083]
74. Fodor SK, Vogt VM. Characterization of the protease of a fish retrovirus, walleye dermal sarcoma virus. J Virol. 2002; 76:4341–4349. [PubMed: 11932400]
75. Yu SF, Baldwin DN, Gwynn SR, Yendapalli S, Linial ML. Human foamy virus replication: a pathway distinct from that of retroviruses and hepadnaviruses. Science. 1996; 1271:1579–1582. [PubMed: 8599113]
76. Isken O, Maquat LE. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. 2007; 21:1833–1856. [PubMed: 17671086]

77. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. 2000; 19:6860–6869. [PubMed: 11118221]

78. Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, Jacobson A. A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature. 2004; 432:112–118. [PubMed: 15525991]

79. Page MF, Carr B, Anders KR, Grimson A, Anderson P. SMG-2 is a phosphorylated protein required for mRNA surveillance in *Caenorhabditis elegans* and related to Upf1p of yeast. Mol Cell Biol. 1999; 19:5943–5951. [PubMed: 10454541]

80. Barker GF, Beemon K. Nonsense codons within the Rous sarcoma virus *gag* gene decrease the stability of unspliced viral RNA. Mol Cell Biol. 1991; 11:2760–2768. [PubMed: 1850103]

81. Barker GF, Beemon K. Rous sarcoma virus RNA stability requires an open reading frame in the *gag* gene and sequences downstream of the *gag-pol* junction. Mol Cell Biol. 1994; 14:1986–1996. [PubMed: 8114730]

82. LeBlanc JJ, Beemon KL. Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. J Virol. 2004; 78:5139–5146. [PubMed: 15113896]

83. Weil JE, Beemon KL. A 3′ UTR sequence stabilizes termination codons in the unspliced RNA of Rous sarcoma virus. RNA. 2006; 12:102–110. [PubMed: 16301601]

84. Weil JE, Hadjithomas M, Beemon KL. Structural characterization of the Rous sarcoma virus RNA stability element. J Virol. 2009; 83:2119–2129. [PubMed: 19091866]

85. Withers JB, Beemon KL. The structure and function of the Rous sarcoma virus RNA stability element. J Cell Biochem. 2011; 112:3085–3092. [PubMed: 21769913]

86. Eberle AB, Stalder L, Mathys H, Orozco RZ, Mühlemann O. Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region. PLoS Biol. 2008; 6:e92. [PubMed: 18447580]

87. Hogg JR, Goff SP. Upf1 senses 3′ UTR length to potentiate mRNA decay. Cell. 2010; 143:379–389. [PubMed: 21029861]

88. Ajamian L, Abrahamyan L, Milev M, Ivanov PV, Kulozik AE, Gehring NH, Mouland AJ. Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. RNA. 2008; 14:914–927. [PubMed: 18369187]

89. Mocquet V, Neusiedler J, Rende F, Cluet D, Robin JP, Terme JM, Duc Dodon M, Wittmann J, Morris C, Le Hir H, et al. The human T-lymphotropic virus type 1 tax protein inhibits nonsense-mediated mRNA decay by interacting with INT6/EIF3E and UPF1. J Virol. 2012; 86:7530–7543. [PubMed: 22553336]

90. Nakano K, Ando T, Yamagishi M, Yokoyama K, Ishida T, Ohsugi T, Tanaka Y, Brighty DW, Watanabe T. Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect. 2013 pii: S1286-4579(13)00064-6.

91. Ross SR. Are viruses inhibited by APOBEC3 molecules from their host species? Plos Pathog. 2009; 5:e1000347. [PubMed: 19390611]

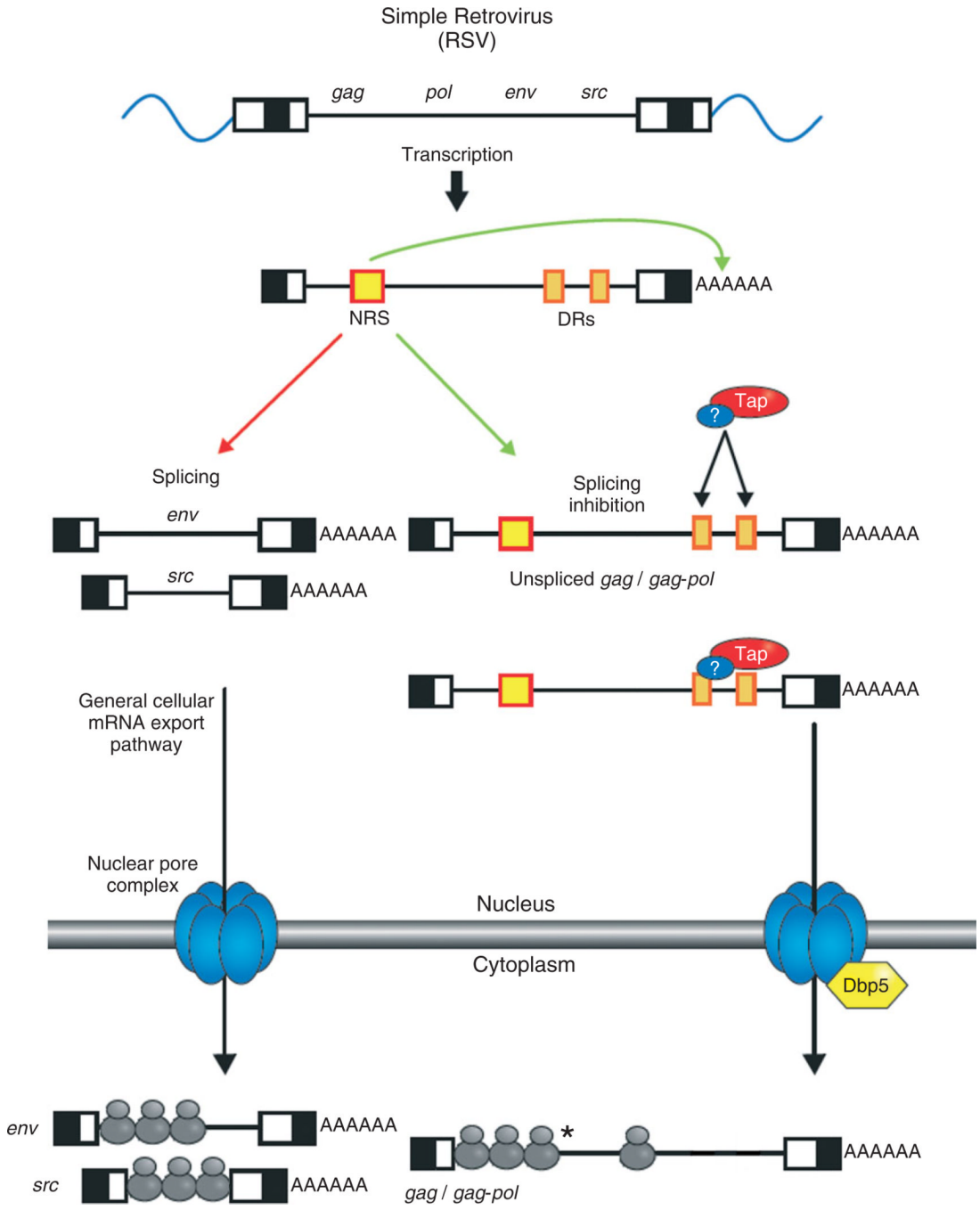

FIGURE 1.

Simple retroviruses, such as Rous sarcoma virus, transcribe a 9 kb primary RNA transcript. A fraction of these transcripts are spliced to generate sub-genomic *env* and *src* mRNAs (left side), which are exported to the cytoplasm like spliced cellular mRNAs. The remaining primary transcripts are not spliced (right side), due to inhibition by the viral NRS element and by inefficient 3′ splice sites. The viral DR sequences interact with cellular proteins to facilitate export of the unspliced RNA to the cytoplasm. The cytoplasmic translocation of the unspliced viral RNAs through the nuclear pore complex is facilitated by the DEAD box helicase Dbp5. The major translation product of the unspliced mRNA is the Gag protein. In

addition, an infrequent −1 frame-shift (*) before the *gag* termination codon allows translation of the Gag-Pol polyprotein.

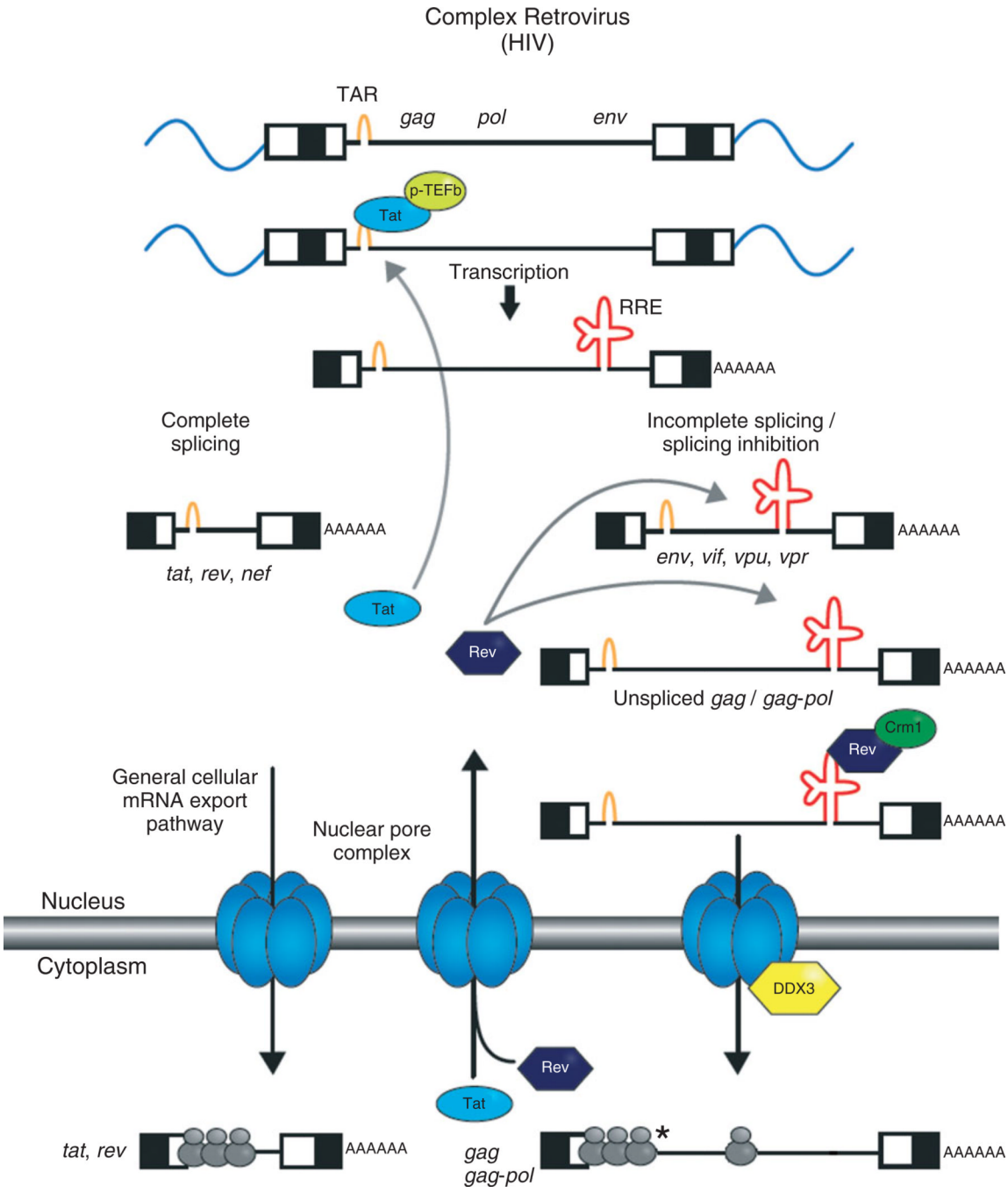

FIGURE 2.

Complex retroviruses, such as HIV-1, generate 3 classes of mRNA: unspliced (9 kb), singly spliced (4 kb), and completely spliced (2 kb). Early HIV gene expression results in the generation of completely spliced mRNAs (left side), including *tat* and *rev*. Following translation, Tat and Rev are imported into the nucleus. Tat binds the TAR RNA element to promote transcriptional elongation, while Rev binds the RRE of singly spliced and unspliced RNAs (right side) to mediate nuclear export of the viral RNA after recruiting the cellular factor CRM1. The cytoplasmic translocation of the partially spliced/unspliced viral RNA through the nuclear pore complex is facilitated by the DEAD box helicase DDX3.

Translation of the unspliced HIV-1 RNA is similar to that of RSV shown in Figure 1. * Marks the site of a frame shift from the *gag to* the *pol* reading fame.

TABLE 1

Viral and Cellular Factors Affecting Retroviral Gene Expression

Virus	HIV	HTLV	MMTV	MPMV	MLV	RSV
Simple or complex virus	Complex	Complex	Complex	Simple	Simple	Simple
Transcription						
Viral effector	Tat	Tax	None	None	None	None
cis-acting RNA sequence	TAR	TAR	None	None	None	None
Splicing						
Alternative splicing	Yes	Yes	Yes	No	No	Yes
Protein-coding mRNAs[1]	7	4	6[2]	2	2	3
Nuclear export						
Viral effector	Rev	Rex	Rem	None	None	None
cis-acting element	RRE	RexRE	RemRE	CTE	Unknown	DR
Cellular factors	CRM1	CRM1	CRM1	Tap	Unknown	Tap
NPC elements	DDX3	Unknown	Unknown	Unknown	Unknown	Dbp5
Translation						
Frame shift sites	One	Two	Two	Two	None[3]	One
5′ UTR PCE	Present	Present	Unknown	Present	Unknown	Unknown

CTE, constitutive transport element; HIV, human immunodeficiency virus; HTLV, human T-lymphotropic virus; MMTV, mouse mammary tumor virus; MPMV, Mason-Pfizer monkey virus; MLV, murine leukemia virus; RSV, Rous sarcoma virus; RRE, Rev Response Element; RexRE, Rex-responsive element; TAR, *trans*-activating region.

[1] Includes only the abundantly expressed mRNAs; minor splice variants are omitted.

[2] Includes the three distinct *sag* mRNAs, which encode the viral superantigens.

[3] MLV read through across the *gag*-pol junction occurs via termination codon suppression using a glutamine tRNA.