

Can the HIV-1 splicing machinery be targeted for drug discovery?

Zodwa Dlamini
Rodney Hull

Research, Innovation & Engagements
Portfolio, Mangosuthu University of
Technology, Durban, South Africa

Abstract: HIV-1 is able to express multiple protein types and isoforms from a single 9 kb mRNA transcript. These proteins are also expressed at particular stages of viral development, and this is achieved through the control of alternative splicing and the export of these transcripts from the nucleus. The nuclear export is controlled by the HIV protein Rev being required to transport incompletely spliced and partially spliced mRNA from the nucleus where they are normally retained. This implies a close relationship between the control of alternate splicing and the nuclear export of mRNA in the control of HIV-1 viral proliferation. This review discusses both the processes. The specificity and regulation of splicing in HIV-1 is controlled by the use of specific splice sites as well as exonic splicing enhancer and exonic splicing silencer sequences. The use of these silencer and enhancer sequences is dependent on the serine arginine family of proteins as well as the heterogeneous nuclear ribonucleoprotein family of proteins that bind to these sequences and increase or decrease splicing. Since alternative splicing is such a critical factor in viral development, it presents itself as a promising drug target. This review aims to discuss the inhibition of splicing, which would stall viral development, as an anti-HIV therapeutic strategy. In this review, the most recent knowledge of splicing in human immunodeficiency viral development and the latest therapeutic strategies targeting human immunodeficiency viral splicing are discussed.

Keywords: alternative splicing, exonic splicing enhancer, exonic specific silencer, splicing based therapies, SR proteins, hnRNP, Rev, Tat, Vpr

Introduction

Like all retroviruses, HIV-1 displays an absolute need for alternative splicing in order to form mature infectious viral particles, as this is necessary to form the viral envelope transcript. This is because like all retroviruses, HIV-1 encodes all of its necessary protein on a single 9 kb genomic mRNA transcript. Previously, it was thought that >40 different mRNAs are encoded by this transcript;¹ however, the next generation sequencing has revealed that there are ~109 differently spliced mRNAs.² Retroviral genomes characteristically consist of a 5'-long terminal repeat (5'-LTR) followed by the *gag*, *pol*, and *env* genes and a 3'-LTR. These LTR regions are the parts of the viral genome that are joined to the genomic DNA of the host cell (Figure 1).¹ The expression of multiple proteins and multiple isoforms of these proteins are achieved through the use of alternative splicing combined with the specific release of particular species of spliced mRNA from the nucleus. The HIV-1 splicing events can be divided based on the phases of virus expression, which correlates to the level of splicing.³ As a result, three different classes of spliced transcript are formed. The 2 kb completely spliced the transcripts, 4 kb partially spliced the transcripts, and 9 kb unspliced the

Correspondence: Zodwa Dlamini
Research, Innovation & Engagements
Portfolio, Mangosuthu University of
Technology, 511 Griffiths Mxenge
Highway, Umlazi, Durban 4031,
South Africa
Email dlaminiz@mut.ac.za

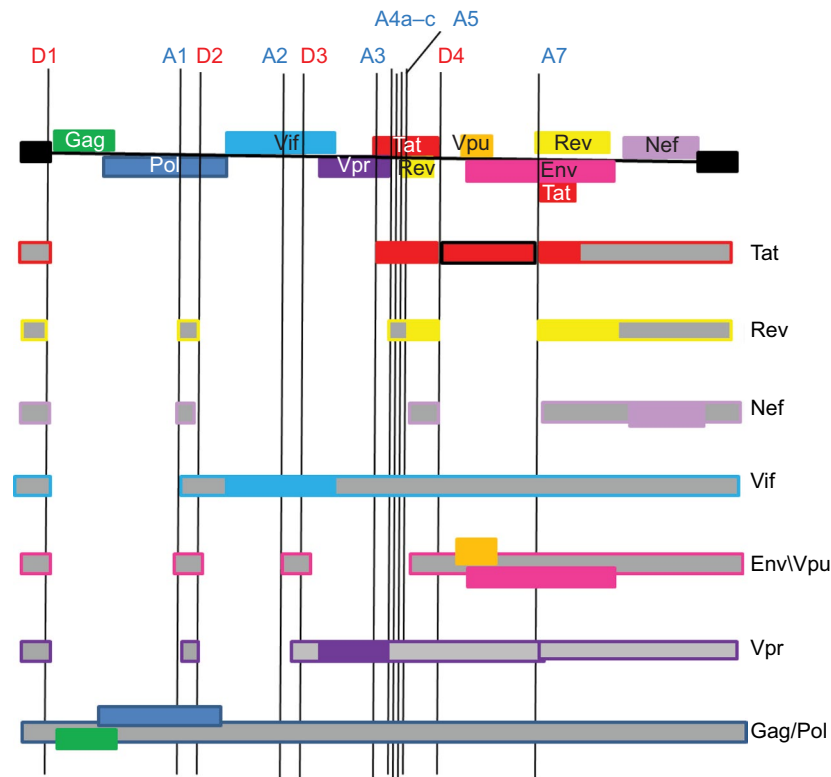


Figure 1 Organization of the HIV genome.

Note: The 5' ss (D1–4) and 3' ss (A1–7) are indicated as well as the resulting mRNA transcripts with non-coding sequences marked in gray.

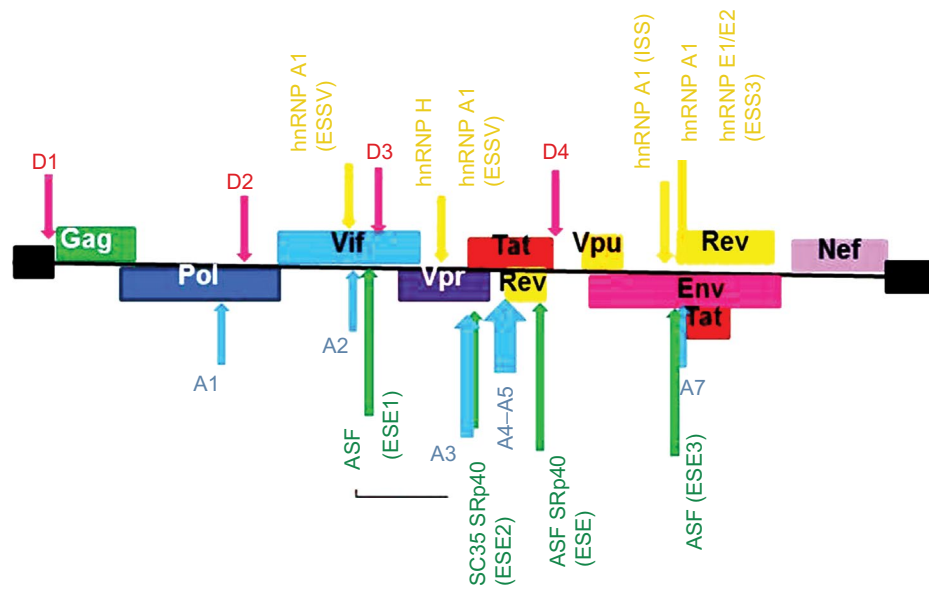
class. Immediately after infection, short completely spliced mRNAs are produced, which encode the viral regulatory proteins Tat, Rev, and Nef. These proteins are required in the early stages of viral infection. Their expressions are followed by the production and nuclear export of the partially spliced mRNA coding for Vif, Vpr, Vpu, and Env. Finally, the 9 kb transcript encodes the viral structural proteins Gag and GagPol.^{4,5} The control of these splicing activities is vital to virus maturation and multiplication. This complex pattern of transcription is controlled by the proteins Rev and Tat and is orchestrated by four alternative 5' splice sites (ss; SD1–4) and eight alternative 3' splice acceptor sites (SA1, SA2, SA3, SA4a, SA4b, SA4c, SA5, and SA7; Figures 1 and 2). The 3' ss vary in their usage depending on various regulatory signals including exon splicing enhancing (ESE) signals and exon splicing silencing (ESS) signals. ESE signals regulate splicing by interacting with the members of the SR protein family. The ESS signals bind the members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family that block the assembly of the spliceosome complex. Splicing within the HIV-1 genome generally occurs in a 5' to 3' direction, resulting in the sequential removal of introns from the 5' end.^{4,5}

HIV-1 genome ss

The splice acceptors in HIV-1 contain short polypyrimidine tracts, unusual branch points, and regulatory sequences that up- or down-regulate the usage of ss generate a large number of transcripts from a single mRNA. All these factors lead to the splicing being sub-optimal. The levels of various mRNAs differ with Env and Nef being abundant and Tat, Vif, and Vpr occurring at lower amounts. It is thought that these differences in mRNA levels arise because of the differences in ss efficiency.⁶ Most of the 5' ss contain the conventional splice consensus sequence.⁷ However, the 3' ss consist of non-consensus splice sequences, and these ss generate the inefficient and delayed splicing of HIV-1 mRNA.⁸ These 5' ss are locations for the binding of U1 snRNP, whereas the 3' ss are locations for the binding of the U2AF splicing factor. Different components of this splicing factor bind to different regions of the 3' ss. The U2AF65 domain binds to a polypyrimidine tract, whereas the U2AF35 domain binds to a conserved AG sequence.⁵

ss efficiency

These ss are described as either strong or weak, and this description depends on the level of base pair homology



Acceptor site	Regulatory elements	Regulatory factors involved	Sequences	Positions
A1			AAUUUUCGGUUUAUUACAG	
A2	ESSV	hnRNP A1	UUAGGACAUUAGUUAGCCCUAGG	4995–5017
	ESE1	ASF/SF2		
A3	ESSp	hnRNP H	UGGGU	5362–5366
	ESS2	hnRNP A1	CUAGACUAGA	5428–5437
	ESE2	SC35, SRp40	CCAGUAGAUCCUAGACUAGA	5418–5437
A4			AGUUUGUUUCACAACAAAAG	
			AGUUUGUUUCACAACAAAAGCCUAG	
			AAGUGUUGCUUUCUUGCCAAG	
A5	ESE GAR	ASF/SF2 SRp40	GAAGAAGCGGAGACAGCGACGAAGA	5558–5582
	ESS3 E1/E2	hnRNP A1, hnRNP	AGAUCCAUUCGAUUAG/unknown	8047–8062
A7	ISS	hnRNP A1	UAGUGAAUAGAGUUAGGCAGGGA	7928–7950
	ESE3	ASF/SF2	GAAGAAGAA	8016–8025
		hnRNP A1	UAGAAGAAGAA	8018–8025

Figure 2 Position of splicing enhancers and silencers as well as the proteins that target these sites.

Notes: Negative regulation sites are marked in orange, whereas positive regulators of splicing are marked in green. The sequences of the acceptor sites as well as their positions and the regulatory elements and factors associated with each site are given in the table below the figure.

Abbreviations: ESS, exonic splicing silencer; ESE, exonic splicing enhancer; ESSV, exonic splicing silencer; hnRNP, heterogeneous nuclear ribonucleoprotein; ISS, intronic splicing silencer; GAR, guanosine-adenosine rich.

between the 5' ss and the 3' ss have for either U1 snRNP or U2AF, respectively.⁵ Studies that aimed to determine the efficiency of splicing were based on comparing the ability of ss to induce splicing when they were inserted into the exon-one, intron one, exon two region of human β globulin constructs. In comparison with a control β globulin 5' ss, all the HIV 5' ss D1 and D4 were just as efficient as the control 5' ss. The efficiency of the 5' ss could also be further ranked

with D1 and D4 being two to three times more efficient than D2 and D3.⁷ The same study indicated that all the HIV-1 3' ss were less efficient than the β globulin 3' ss.⁷ The importance of negative and positive regulation of 3' ss by their downstream exonic sequences was demonstrated Kammler et al.⁹ Removal of the downstream exonic sequences also decreased the efficiency of all the 3' ss except for ss A2 and A3 that maintained most of their splicing efficiency. However, the

splicing efficiency of the A2 and A3 ss is decreased when the downstream exonic sequences were included.⁵

A1 ss

The A1 ss is used to form the Vif mRNA (Figure 2) through the use of D1 5' ss and A1 3' ss, and therefore suboptimal splicing at this ss is necessary for viral progression. The appropriate splicing activity of this ss requires a suboptimal 5' ss (D2) downstream of A1. More efficient splicing at D2 results in increased splicing at the A1 site and a decrease in viral particle production as there is less unspliced RNA.^{5,10}

A2 ss

The A2 ss is used to form the Vpr mRNA (Figure 2). This is negatively regulated through the activity of the ESS silencer. Inhibition of this silencer results in reduced production of viral particles because of the increased splicing at A2 and a decrease in unspliced mRNA levels. The SR protein ASF increases splicing at this site (Figure 2).^{5,10}

A3 ss

This site has a low splicing efficiency even though it has a highly optimized polypyrimidine tract, and this is probably due to the presence of two silencer sequences, one upstream ESS2p and another downstream ESS2. At the same time, the enhancer site ESE2 is a target site for multiple SR proteins. Splicing at this site gives rise to the Tat protein that is essential for viral transcription (Figure 2).^{5,10} The multiple Tat mRNAs are formed by splicing 3' ss A3 to D1, D2, or D3, and these transcripts are either completely spliced consisting of two exons or are incompletely spliced and consist of one exon.⁵

A4 ss

Splicing at all three of the A4 sites gives rise to multiple isoforms of the Rev protein.^{5,10} Splicing at the 4c ss occurs less frequently than splicing at the 4a and 4b 3' ss. However, all these ss are relatively weak and require the activity of enhancer sequences. The strength of splicing is also increased by the strength of the downstream 5' D4 ss (Figure 2).⁵

A5 ss

Utilization of this ss gives rise to env mRNA (Figure 2).^{5,10} Splicing at the A5 site occurs more frequently than at all three A4 ss. This is a relatively weak ss and requires the activity of splice enhancer sequences. This includes the guanosine-adenosine-rich (GAR) ESE. As with the 4' ss, the strength of splicing is increased by the strength of the downstream D4 5' ss.⁵

A7 ss

Splicing at this site is negatively regulated by three silencing sequences.¹⁰ Inactivation of this ss led to an absence of splicing within the tat/rev intron; however, this did not affect the number of unspliced REV-dependent transcripts.⁹ Negative regulation also occurs as a result of an intronic splicing silencer (ISS) sequence (Figure 2). Positive regulation of splicing at this site is the result of the ESE3 enhancer element. This element is known as the Janus GAA3 element. It has been given the name Janus sequence because of its dual nature of being able to act as a silencer sequence as well as an enhancer sequence.⁵

Splicing enhancers and silencers

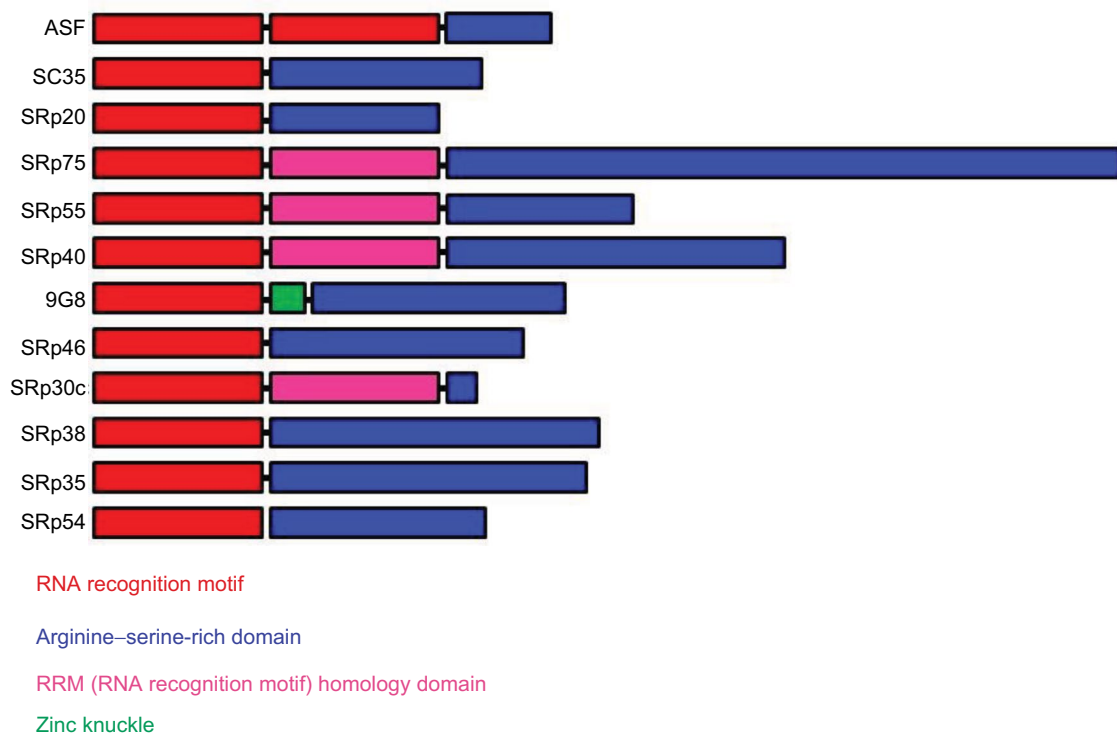
As mentioned earlier, HIV mRNAs also contain cis-elements that increase or decrease splicing of these RNAs. These ESEs or ESSs are generally binding sites for the members of the SR protein family and the hnRNP protein family, respectively.⁵

SR proteins

This family of proteins is highly conserved, and each member contains one or two RNA recognition motifs (RRMs) at their N terminus and a serine-arginine-rich domain at their C terminus. SR proteins interact with RNA through their N terminus and with proteins through their C terminus (Figure 3).^{5,11} In alternative splicing, they function to increase the activity of a 3' ss by recruiting the members of the splicing complex. In addition, this may result in an increase in the usage of the downstream 5' sites, which may also block the binding of inhibitory factors. SR proteins can also inhibit splicing by binding to splicing silencer elements and preventing the formation of the spliceosome. The activity of SR proteins is determined by their site of phosphorylation, with phosphorylation being required for ss recognition and de-phosphorylation being required for splicing catalysis.^{5,12,13} SRp75 acts on the splice acceptor 2 element, which results in the inclusion of exon 2 to produce the mRNA for vif. This is down-regulated by a splicing silencer downstream of exon 2.¹⁴ SRp55 also binds a splice silencer within exon 5 to inhibit splicing at SD3.¹⁵ In exon 5, there is a splice enhancer known as GAR, which interacts with SRp40 and ASF/SF2 to increase the splicing activity at SA5, resulting in the increased transcription of nef.¹⁵ SR proteins also influence the splicing of tat by binding to sequences within exon 4.¹⁶ SR proteins also increase the expression of HIV proteins by stabilizing viral mRNAs and promoting their nuclear export.¹⁷

hnRNPs

The hnRNP A/B protein family was first identified as inhibitors of HIV-1 pre-mRNA splicing, when it was discovered



Protein	Function	HIV binding sites
ASF	Prevents exon skipping, ensures splicing accuracy. Isoforms repress splicing	A2-ESE1 A5-ESE GAR A7-ESE
SC35	Splicing of pre-mRNA. Spliceosome assembly, Formation of the splicing complex and interactions of both U1 and U2 snRNPs with pre-mRNA. Apoptosis. Exerts transcription corepressor activity	
SRp20 SRp75	RNA processing in cellular proliferation AS site selection. Represses the splicing of MAPT/Tau exon 10	
SRp40 SRp55	Splicing and selection of AS sites Splicing and selection of AS sites. Represses the splicing of MAPT/Tau exon 10. RNA binding	A5-ESE GAR
9G8	Required for pre-mRNA splicing. Can also modulate AS in vitro. Represses the splicing of MAPT/Tau exon 10. RNA binding; zinc ion binding	
SRp35 SRp54	Seems to antagonize SR proteins in pre-mRNA splicing regulation May function in pre-mRNA splicing	
SRp46	AS regulator, Represses the splicing of MAPT/Tau exon 10	
SRp30c	Splicing and selection of AS sites. Represses the splicing of MAPT/Tau exon 10	A3-ESE2
SRp38	Involved in constitutive and regulated RNA splicing	
SRp54	Involved in constitutive and regulated RNA splicing	

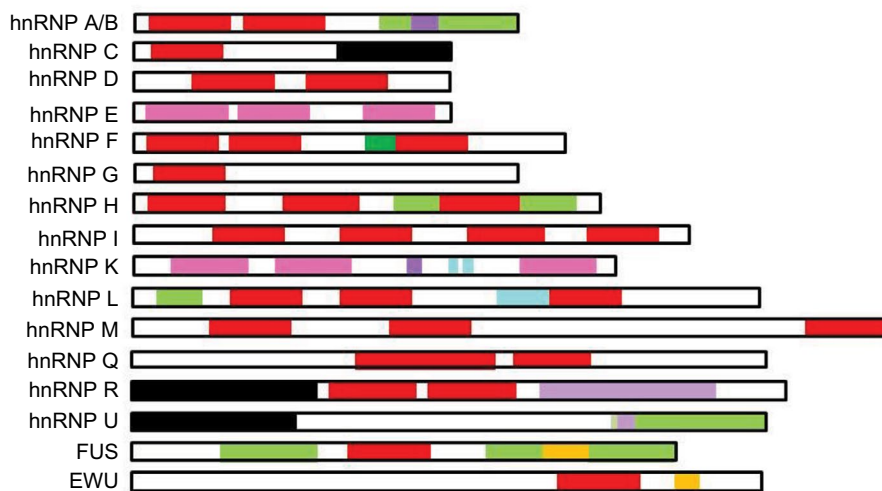
Figure 3 List of SR proteins involved in splicing and their domain structure.

Note: The function of each SR protein and its HIV binding sites are given in the table below the figure.

Abbreviations: ESE, exonic splicing enhancer; GAR, guanosine-adenosine rich.

that they associate with the exonic silencer sequences located in exon 2, the area coding for the Tat protein, and control the expression of Tat. These proteins consist of two RRM and a glycine-rich domain at their C-terminus (Figure 4).¹⁸ However, not all of these RNA binding domains are necessary for the regulation of HIV mRNA splicing.¹⁹ The ability of hnRNP A/B proteins to inhibit splicing seems to involve them in

counteracting the ability of SR proteins to recognize and bind to ESE sequences.¹⁸ Two further splice silencers that control tat mRNA splicing are found in exon 4. They also function to inhibit the production of tat mRNA. These silencers are known as ESS2 and ESSp2 and are bound by hnRNP A1 and hnRNP H, respectively.²⁰ However, the silencing function of ESS2 can be inhibited by the binding of the SR protein SC35 to a splicing



Domain	Function
RNA recognition motif (RRM)	Bind single-stranded RNA
K homology (KH) domain	Nucleic acid-binding. RNA binding and recognition
Arginine-Glycine-Glycine box (RRG box)	RNA binding
Glycine-rich domain	Nuclear localization, protein binding
Proline-rich domain	Protein interaction domain
Zinc finger	DNA binding
Asp Glu-rich acidic domain	DNA/RNA mimicry

Figure 4 Members of the hnRNP protein family and their domain structure.
Note: The general function served by each of these domains in hnRNP protein family is given in table below the figure.
Abbreviation: hnRNP, heterogeneous nuclear ribonucleoprotein.

enhancer located within exon 4.²¹ The hnRNP family was later found to be involved in the regulation of alternative splicing of other HIV-1 transcripts. For example, hnRNP A/B prevents splicing to form vpr by binding to a silencer within exon 3 known as exonic splicing silencer. This binding prevents the binding of U2AF65 to the SA3 ss. U2AF65 is required for the binding of U2AF.²² The splicing factor ASF/SF2 also competes with hnRNP A/B for binding positions that lead to splicing into SA3.¹⁶ All 2 kb fully spliced mRNAs depend on SA7. Here, hnRNP A1 binds to the ISS, located upstream of SA7, inhibiting splicing. Splicing involving SA7 is also blocked through the binding of hnRNP A1 to the ESS3 within exon 7.²³ Partially spliced viral mRNAs contain negative nuclear retention sequences known as instability sequence elements (INS). Regulatory proteins bind these mRNAs at these sequences sequestering them and preventing their nuclear export. The list of proteins capable of binding these negative regulatory sequences include hnRNP A1.²⁴ In addition to the hnRNP A/B

family, the hnRNP H family also plays a role in the control of splicing in HIV. This family includes hnRNPs H, H', F, and GRSF-1, all of which have been shown to interact with splicing enhancers and silencers. The minimal RNA recognition sequence of this family was found to be able to control the splicing of tat in an ADP-dependent manner.²⁵

HIV RNA secondary structures

The secondary structure adopted by the target viral mRNA plays a role in the regulation of splicing through the conformation that it adopts either exposing or hiding sites.⁵ The 5' splice donor RNA is able to fold into a stem loop structure. The ability of the RNA to maintain this structure is a key means of splicing regulation as this structure restricts the access of the splicing machinery to the RNA sequence.²⁶ For instance, the region around the ESS3 splicing silencer adopts a hairpin loop structure. The loop contains the binding site for the hnRNP-A1 protein.²⁷ On the other hand, the

secondary RNA structure cannot only interfere with the access of the spliceosome to splice donor and acceptor sites on mRNA but also result in increased access when compared to their secondary RNA structures. For example, the stem-loop structures of viral mRNA will result in unpaired nucleotides. The unpaired nucleotides give better access to the ss resulting in the increase in the efficiency of splicing.²⁸ The effect of various RNA secondary structures on splicing is shown in Table 1.

Methylation of adenosines N⁶ (m⁶A) occurs within the 3' untranslated region of the HIV genome. These methylation sites lead to the recruitment of cellular YTH domain family proteins (YTHDF1–3). YTHDF proteins bind to m⁶A-modified cellular RNAs and affect RNA metabolism and processing. These proteins then promote the translation of mRNA enhancing the expression of viral proteins.²⁹ Infection of CD4 T cells also leads to an increase in the levels of methylation of both host and viral RNA. Methylation of two conserved adenosines in the stem loop II region of Rev response element (RRE) increases the RNA binding ability of RRE.³⁰

Viral proteins associated with alternative splicing

Rev

Rev is one of the first proteins produced by the virus. It is an 18 kDa protein and localizes to the nucleus because of its nuclear localization signal (NLS). Here, it binds to a structural motif in the env gene known as the RRE. At the same time, it interacts with the cellular protein Crm1. This results in the nuclear export of the partially spliced 4 kb transcripts and the un-spliced 9 kb transcript.³¹ The nuclear export of all

partially spliced mRNAs is unusual as they contain functional introns. The primary function of the Rev protein is its direct interaction with the RRE, a cis acting target that is found on incompletely spliced viral mRNAs, facilitating the export of unspliced viral mRNA from the nucleus.³² Rev is a 116 amino acid long protein containing an arginine-rich NLSs and leucine-rich nuclear export signals (NESs). The protein is transported into the nucleus through the interaction of the NLS domain nuclear import factors such as importin β .²⁴ The RRE is 350 nucleotides in length and forms a stem loop-rich conformation. At least six Rev monomers bind to the RRE and form RRE-RNA-Rev dimers. The two Rev proteins form a V shape in this dimer exposing the arginine-rich RNA binding domain of each protein to the RRE of the RNA. The RRE-RNA-Rev dimers then oligomerize to form a larger RNP complex.²⁴ The NES of the protein interacts with CRM1 leading to the nuclear export of the RNP complex. However, CRM1 is not the only protein capable of interacting with Rev, resulting in nuclear export. Multiple other host co-factors can perform this function, including a range of RNA helicases whose expression levels are known to increase following HIV infection.²⁴ The association with CRM1 is regulated by a GTPase. This is achieved by a RAN-GTP gradient established between the nucleus and the cytoplasm through the dephosphorylation of RAN-GTPase into RAN-GDPase in the cytoplasm. The triphosphate form favors the binding of CRM1 to the NES of Rev, whereas the diphosphate form favors the release of the transcript and Rev from CRM1.^{32,33} Rev also recruits the methyltransferase PIMT (peroxisome proliferative activated protein with methyltransferase domain) to act on RRE containing mRNAs. PIMT selectively hypermethylates the normally monomethylated RNA cap into a trimethylated cap. This results in the hypermethylation of partially spliced mRNAs but not fully spliced mRNAs, and these mRNAs with the altered RNA cap structure are recognized as substrates for the CRM1 transporter protein.²⁴ The nuclear export of Rev–RRE complexes also requires the eukaryotic initiation factor (eIF-5A) which is thought to target the Rev–RRE complex to the nuclear pore complex (NPC) and increases the binding efficiency to CRM1.³⁴ Rev function is increased but not reliant upon its association with the FG repeat nucleoporin protein Rip/Rab. This protein aids in NPC localization.³² The export of large mRNA fragments is also aided by the association of the Rev–RRE complex with the RNA helicase DDX3.³⁵ In addition to the instability (INS) sequences that act as cis-regulatory elements and that prevent nuclear export, some viral mRNAs also contain AU-rich sequences that repress nuclear export. Rev and RRE are able to counteract these

Table 1 The influence of RNA secondary structure on specific splice sites

Splice site	3D structure	Affected enhancers or silencers
A2-D3	Single stem loop	ESSV exposed
A3	Extended stem loop	ESS2 and ESE2 exposed and bound simultaneously by hnRNP A1
D1	Stable stem loop	Role is not established
A7	Three stem loops	Loop 1 contains the ISS Loop 2 contains the A7 site and ESE3 Loop 3 contains ESE3

Notes: This table lists the secondary structural conformations adopted by HIV-1 mRNA and the influence that this secondary structure has on splicing enhancers and silencers. Adapted from Stoltzfus.⁵

Abbreviations: 3D, three dimensional; ESS, exonic splicing silencer; ESE, exonic splicing enhancer; ESSV, exonic splicing silencer; hnRNP, heterogeneous nuclear ribonucleoprotein; ISS, intronic splicing silencer.

retention signals, leading to nuclear export.³⁶ A number of other proteins have been identified as associated with Rev. These interactors are listed in Table 2. Rev also enhances the translation of RRE containing mRNAs, by enhancing the association of these mRNAs with the polyribosome, by interacting with the co-factors that aid in the association of RRE containing mRNAs with the translational machinery. It does not seem to be able to interact directly with ribosomal proteins. Alternative explanations for this enhanced translation include increased RNA stability.²⁴ Rev is known to enhance the stability of RRE containing pre-mRNA in the nucleus.³²

Tat

Early HIV-1 expression studies confirmed that the expression of reporter genes under the control of HIV-1 LTR required a transactivating factor. This factor was found to be a viral protein and was thus given the name Tat.³⁷ The functions of Tat can thus be summarized as a molecular adapter between components of the cellular machinery involved in transcription and the viral mRNA.³³ The activity of Tat is dependent on a sequence known as the transactivating response region (TAR), which is located downstream of the initiator site 3' of the HIV promoter. The TAR assumes a U-rich bulge confirmation, and this bulge is specifically recognized by Tat. In addition, the G34 residue in TAR is required to form a nucleotide bulge by forming a base pair association with residue C30 of TAR.³³ Tat is a small 101 amino acid protein that is the result of a multiple spliced, RNA fragment and contains a co-factor binding domain, a basic RNA binding domain, which is also responsible for nuclear export, and a C terminal domain that is involved in mRNA capping.³³ If

Tat is deleted or inhibited, the result is short RNA sequences that are not polyadenylated. This implies that without Tat, RNA polymerase II has poor processivity of viral mRNA, most likely because of hypophosphorylation of the carboxy terminal domain of RNA polymerase II.³³ Phosphorylation of RNA polymerase II is performed by the CDK9 component of P-TEFb, which in turn can associate with multiple cyclins, including Cyclin T1. Tat selectively recruits cyclin T1 into this Tat-P-TEFb complex, activating the HIV-1 LTR promoter.³⁸ This association is thought to allow Tat to increase the efficiency of elongation of RNA polymerase II.³³

Tat also plays a role in inhibiting histone acetylases, resulting in chromatin remodeling near the transcriptional start site.³³ Tat has been shown to be involved in the phosphorylation of the SR protein SF2, which leads to an increase in the splicing of viral RNA. However, acetylated Tat leads to the increase of unspliced mRNA, and the acetylation of Tat increases as infection proceeds.³⁹ A sequence in exon 5 known as the GAR sequence is utilized through Tat to increase the expression of Env.⁴⁰

Vpr

Vpr plays a role in LTR activation, nuclear localization, and regulation of transcription. In general, it acts as an inhibitor of cellular pre-mRNA splicing.^{5,10} Vpr achieves this in a number of ways. First, it interacts with the splicing factor SAP145 (SF3B2) inhibiting it from binding to another splicing factor SAP49 and pre-mRNA, thus preventing the formation of a pre-mRNA SAP49 spliceosomal complex.^{41,42} Second, Vpr mutants result in a decrease in the levels of single spliced and unspliced mRNA that is associated with oversplicing and defective replication.⁴³

Alternative splicing-based therapies

The alternative splicing of viral mRNA is an absolute requirement for viral maturation, expression of viral proteins, and assembly of new viral particles (Figure 5). As such interfering with the splicing of the HIV-1 pre-mRNA is a tempting drug target. This could be accomplished by inhibiting HIV-1 splicing or by enhancing HIV-1 splicing. There are two basic strategies to interfere with alternative splicing in HIV-1 and basically rely on targeting splicing sites or splicing factors. These involve the masking of regulatory mRNA sequences by blocking ss on viral RNA or by altering the expression level or activity of SR or hnRNP proteins. The first approach can involve the use of modified antisense RNA or peptide nucleic acid fusion molecules that block viral ss. Another useful way of blocking ss is through the use of bifunctional

Table 2 Specific proteins that interact with Rev

Possible interactor	Function
B23	Co-localizes with Rev in the nucleoli, helps Rev to maintain its presence in this location
P32	Overcome an import post-translational block of Rev
Nucleosome assembly protein	Nuclear import of Rev
Human I-mfa domain containing protein	Regulate Rev nuclear import
Poly(A) binding protein I	Rev-related translation
Human Rev interacting protein	Promotes the release of incompletely spliced mRNA from the perinuclear region
Matrix-associated RNA binding protein	Stabilizes Rev-RRE interaction, a component of the nuclear pore proteome, frees mRNA from INS-mediated retention

Notes: This table lists the various cellular proteins that interact with the HIV-1 Rev protein and the resulting effect of this interaction. Adapted from Kula A, Marcello A. Dynamic post-transcriptional regulation of HIV-1 gene expression. *Biology*. 2012;1:116–133.³⁵

Abbreviations: RRE, Rev response element; INS, instability.

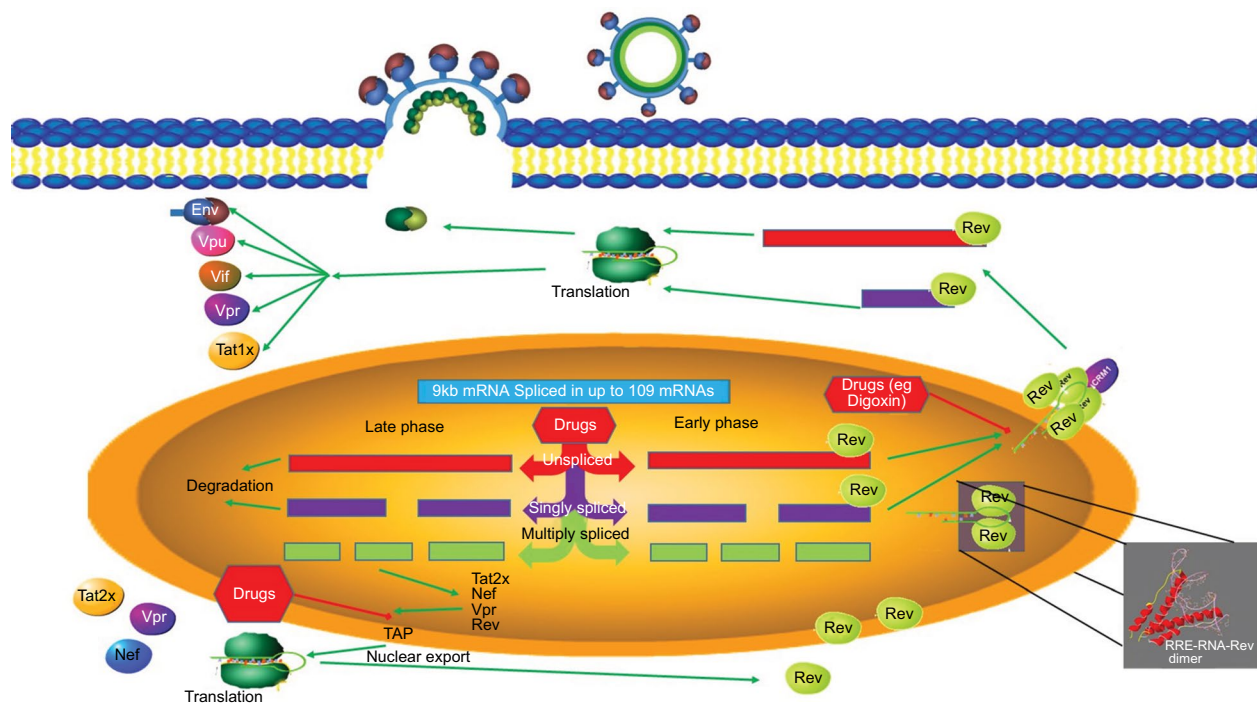


Figure 5 A summary of the splicing pathways in HIV.

Notes: During the early phase, unspliced RNA is exported through association with Rev, where it associates with the ribosome and is translated into Gagpol and Gag proteins. Partially spliced mRNAs are also exported and translated into proteins Env, Vpr, Vif, and Tat. During the late phase of viral infection, multiple spliced mRNAs are exported to the cytoplasm where they are translated into Tat, Vpu, Nef, and Rev.

oligonucleotides. These molecules consist of two separate oligonucleotides: one portion is the antisense of the target, which leads to the molecule being a specific blocker, and a second molecule that does not bind the target mRNA but recruits proteins or RNA–protein complexes that alter ss selection.⁴⁴ The second strategy is viable because the blocking of a single SR protein can be compensated for by other SR protein in the host cell without a significant loss of function while HIV requires specific SR proteins.

The use of anti-sense and siRNA molecules to block or interfere with splicing is well established in research studies and could be applied to therapeutic uses. In a similar fashion, the recent advances in CRISPR-Cas9 technologies have been successfully used to disrupt viral replication. However, the virus was able to adapt and use coupled non-homologous end joining processes to produce viral variants that could evade inhibition of replication.⁴⁵

SR proteins could be targeted by interfering with their phosphorylation. Indole derivatives interfere with the phosphorylation of SR proteins by binding directly to their RS domains, blocking phosphorylation and altering the function of the SR protein. They bind directly to the RS domain of the SR protein that blocks phosphorylation impeding the splicing activity and nuclear trafficking of the protein.⁴⁶ One of these compounds IDC16 was found to inhibit viral production

in macrophages and peripheral blood mononuclear cells. It was also able to inhibit viral production in isolates from patients where the infection had developed resistance to HIV multi-therapies. This small fused tetracyclic indole molecule functions by inhibiting the splicing activity of ASF.^{47,48} Other advantages to IDC16 include low cellular toxicity, based on initial reports, and the fact that the virus would struggle to develop resistance to the drug. This is because resistance would involve mutations in enhancer sequences or mutations that improve the binding of SR proteins, both of which would likely result in non-viable viruses.⁴⁶ In order to improve the activity and decrease the cytotoxicity of these compounds, a parallel synthesis of IDC16 mimics was performed and an active diheteroarylamide-type compound was identified. This compound was named compound 9 (1C8), and although it had the same antiviral activity as IDC1, it had a lower level of cytotoxicity.⁴⁸ Another inhibitor of SR proteins is the isonicotinamide compound SRPIN340, which inhibits the SR protein kinases SRPK1 and SRPK2. This compound inhibits SRp75 phosphorylation, making the protein less stable. However, upon further study, it was discovered that this compound did not effectively inhibit HIV replication.⁵ Some of these therapies are summarized in Table 3.

Since Tat and Rev are necessary for HIV mRNA splicing and expression of viral proteins, both these proteins are

Table 3 Some of the compounds under development as anti-splicing HIV therapies

Compound	Activity	Reference
SR protein targeting treatments		
Tetracyclic indole molecule (IDC16)	Inhibiting the splicing activity of ASF	48
DHA-type compound 9 (1C8)	Inhibiting the splicing activity of ASF	48
Isonicotinamide compound SRPIN340	Inhibitor of the SR protein kinases SRPK1 and SRPK2	
Rev targeting treatments		
Aminoglycoside antibiotics	Inhibit the association of Rev with RRE	32
Aromatic heterocyclic compounds	Inhibit the association of Rev with RRE	32
Clomiphene	Targets the formation of Rev-RRE	49
8-Azaguanine	Target the formation of Rev-RRE	50,55
2-(2-[5-Nitro-2-thienyl]vinyl)quinolone	Target the formation of Rev-RRE	50,55
Indole derivative KH-BD19	CLK inhibitory activity	55
Digoxin	Over-splicing of HIV mRNA and blocking Rev expression	50
Tat targeting treatments		
Antisense U7snRNP	Tat/Rev exon skipping	10
RNA binding domain of ASF	Competing with Tat for binding the TAR domain	53
Dominant negative forms of Tat	Inhibit Tat function	52
MCEF fragments	Repress RNA polymerase II transcription of HIV-1 LTR-directed Tat-transactivation	46

Note: This table lists some of the compounds that are currently being assessed as anti-HIV drugs due to their ability to interfere with or block HIV mRNA splicing.

Abbreviations: DHA, diheteroarylamine; RRE, Rev response element; TAR, transactivating response region; LTR, long terminal repeat.

prime targets to disrupt viral gene expression by interfering with alternative splicing. The most logical target is the RRE sequence on viral RNA. If the association of Rev with this sequence can be blocked or inhibited, viral maturation will be impeded. It is already known that aminoglycoside antibiotics and aromatic heterocyclic compounds inhibit the association of Rev with RRE.³² The absolute requirement for Rev to export partially spliced transcripts from the nucleus has led many to consider this protein as an attractive therapeutic target. The activity of this protein may be targeted using antisense oligonucleotides or ribozymes.³² However, more specific methods to inhibit Rev are being explored. General inhibition of the exportin-1-mediated pathway or NES-containing proteins may lead to undesired side effects and cellular toxicity as these are vital pathways in normal cellular function. Some of these strategies are listed in Table 3. Another strategy has been to identify bioavailable HIV inhibitors that target the gene expression of the virus by interfering with REV activity. This led to the identification of the estrogen receptor modulator, clomiphene. These drugs target the formation of REV–RRE complexes as the anti-splicing effects of clomiphene and reduce the levels of unspliced and singly spliced mRNA to a greater extent than multiply spliced mRNA. This results in a reduction in the level of viral transcription.⁴⁹ Similarly, two other small molecules were found to affect HIV gene expression and altered the levels of viral mRNAs other than Rev and Tat. These compounds were 8-azaguanine

and 2-(2-[5-nitro-2-thienyl] vinyl) quinolone. This ability to decrease the levels of viral mRNAs other than Tat and Rev implied that these molecules were targeting the activity of Rev.⁵⁰ The cardiac glycoside digoxin was identified as having the ability to inhibit HIV replication. This drug accomplishes this by inducing oversplicing of HIV mRNA and by blocking Rev expression (Figure 5).⁵⁰ Dominant negative mutated forms of Rev or proteins essential to Rev function such as eIF-5A are also viable targets for drug development. However, drug delivery remains a problem with the use of mutant proteins as effective drugs.³² A final means of targeting Rev involves depleting the protein within the infected cell by creating sinks and reservoirs of sequestered or inactivated Rev. This can be achieved by using antibodies that sequester Rev within the cytoplasm. Another strategy involves the use of RNA decoys and fragments of non-viral RNA that contain the REV sequence and therefore compete for binding. These two strategies also present difficulties with drug delivery.³²

Tat is also a prime target for drug development and some of the methods being used to target Tat are listed in Table 3. Tat activity has been targeted through the use of proteins or peptides. Portions of the Aff1 transcription factor family member MCEF were found to be able to repress RNA polymerase II transcription of HIV-1 LTR-directed Tat transactivation and could conceivably be used as a transcription repressor or HIV genes.⁵¹ Tat has also been targeted through the use of dominant negative forms of the Tat protein.

Dominant negative domains of Tat have been created and successfully used to inhibit Tat function. These domains were fused with Arg-Ser splicing factor motifs to increase the targeting efficiency of these dominant negative Tat.⁵² Tat function is also inhibited by ASF competing for binding the TAR domain, and overexpression of ASF inhibits viral transcription in the later stages of infection. However, ASF activates transcription in the early stages of infection.⁵³ The RNA binding domain could therefore be used as a therapy way to block Tat activity. The binding of the TAR domain has also been inhibited through the use of specifically designed tripeptides that are able to bind to the TAR RNA stem loop.⁵⁴

Vif has recently been targeted in combination with an antisense U7snRNP. This leads to Tat/Rev exon skipping.¹⁰ Targeting hnRNPs for drug development may not be viable, as they are widely expressed in the human body and therefore are most likely to be involved in many cellular processes.¹⁸ Any attempt to interfere with their activities may therefore result in disastrous consequences for the patient.

The sequential removal of introns from the 5' end of the HIV-1 pre-RNA is itself an attractive therapeutic target. This requirement of the 5' to 3' directionality of splicing is so important that splicing of a 3' intron is inhibited when the 5' intron is not removed. In addition, the mutation of the first 5' ss inhibits all further splicing events, leading to accumulation of unspliced transcripts and blocking the progression of viral maturation.⁴

Conclusion

Most of the current antiviral medication is based on the inhibition of viral protein activity. This can lead to resistance against these drugs by mutation of these protein targets. Targeting alternative splicing pathways can avoid this problem in two ways. First, many of these targets are host cellular proteins that are less prone to mutation due to rapid replication. Second, it would be harder for the virus to adapt to these compounds as changes in the splicing sequences would result in the virus not being able to complete the transcription of all the mRNAs that are required to be replicated. In addition to this, by targeting the splicing of HIV mRNA, the impact on the host cell can be decreased, avoiding toxicity and side effects. This can be achieved through the specific selection of HIV-1 ss. Third, since gene expression is controlled through the regulation of mRNA splicing and export by the viral proteins Tat and Rev, the specific inhibition of these proteins is an attractive therapeutic target. Therefore, inhibition of alternative splicing is a promising target for the development of new therapeutic drugs. The current strategies

being explored from the use of RNAi and CRISPR to small molecules that inhibit SR protein function or block the association of Rev with RRE are diverse and promising in their specificity toward the splicing and eventual gene expression process of HIV.

Disclosure

The authors report no conflicts of interest in this work.

References

- Hryckiewicz K, Bura M, Kowala-Piaskowska A, Bolewska B, Mozer-Lisewska I. HIV RNA splicing. *HIV AIDS Rev.* 2011;10(3):61–64.
- Ocwieja KE, Sherrill-Mix S, Mukherjee R, et al. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res.* 2012;40(20):10345–10355.
- Karn J, Stoltzfus CM. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med.* 2012; 2(2):a006916.
- Bohne J, Wodrich H, Kräusslich H-G. Splicing of human immunodeficiency virus RNA is position-dependent suggesting sequential removal of introns from the 5' end. *Nucleic Acids Res.* 2005;33(3):825–837.
- Stoltzfus CM. Chapter 1 Regulation of HIV-1 Alternative RNA Splicing and Its Role in Virus Replication. *Adv Virus Res.* 2009;74:1–40.
- Purcell DF, Martin MA. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol.* 1993;67(11):6365–6378.
- 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(2):373–385.
- Si Z-h, Rauch D, Stoltzfus CM. The exon splicing silencer in human immunodeficiency virus type 1 tat exon 3 is bipartite and acts early in spliceosome assembly. *Mol Cell Biol.* 1998;18(9):5404–5413.
- Kammler S, Otte M, Hauber I, Kjems J, Hauber J, Schaal H. The strength of the HIV-1 3' splice sites affects REV function. *Retrovirology.* 2006;3:89.
- Tazi J, Bakkour N, Marchand V, Yadi L, Aboufirassi A, Branlani C. Alternative splicing: regulation of HIV-1 replication as a target for therapeutic action. *FEBS J.* 2010;277:867–876.
- Manley JL, Krainer AR. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.* 2010;24(11):1073–1074.
- Graveley BR. Sorting out the complexity of SR protein functions. *RNA.* 2000;6(9):1197–1211.
- Mermoud JE, Cohen P, Lamond AI. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res.* 1992;20(20):5263–5269.
- Exline CM, Feng Z, Stoltzfus CM. Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. *J Virol.* 2008;82(8):3921–3931.
- Tranell A, Fenyo EM, Schwartz S. Serine- and arginine-rich proteins 55 and 75 (SRp55 and SRp75) induce production of HIV-1 vpr mRNA by inhibiting the 5'-splice site of exon 3. *J Biol Chem.* 2010;285(41): 31537–31547.
- Ropers D, Ayadi L, Gattoni R, et al. Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. *J Biol Chem.* 2004;279(29):29963–29973.
- Tranell A, Tingsborg S, Fenyo EM, Schwartz S. Inhibition of splicing by serine-arginine rich protein 55 (SRp55) causes the appearance of partially spliced HIV-1 mRNAs in the cytoplasm. *Virus Res.* 2011;157(1): 82–91.
- Caputi M, Mayeda A, Krainer AR, Zahler AM. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J.* 1999; 18(14):4060–4067.

19. Jean-Philippe J, Paz S, Michael L, Lu, Caputi M. A truncated hnRNP A1 isoform, lacking the RGG-box RNA binding domain, can efficiently regulate HIV-1 splicing and replication. *Biochim Biophys Acta*. 2014;1839(4):251–258.
20. Jacquenet S, Mâreau As, Bilodeau PS, Damier L, Stoltzfus CM, Branlant C. A second exon splicing silencer within human immunodeficiency virus type 1 tat exon 2 represses splicing of tat mRNA and binds protein hnRNP H. *J Biol Chem*. 2001;276(44):40464–40475.
21. Zahler AM, Damgaard CK, Kjems J, Caputi M. SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing. *J Biol Chem*. 2004;279(11):10077–10084.
22. Domsic JK, Wang Y, Mayeda A, Krainer AR, Stoltzfus CM. Human immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV antagonizes binding of U2AF65 to viral polypyrimidine tracts. *Mol Cell Biol*. 2003;23(23):8762–8772.
23. 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(8):4606–4615.
24. Jeang K-T. Multi-faceted post-transcriptional functions of HIV-1 Rev. *Biology*. 2012;1(2):165–174.
25. 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(18):13617–13626.
26. Mueller N, van Bel N, Berkhout B, Das AT. HIV-1 splicing at the major splice donor site is restricted by RNA structure. *Virology*. 2014;468–470:609–620.
27. Leventgood JD, Rollins C, Mishler CHJ, et al. Solution structure of the HIV-1 exon splicing silencer 3. *J Mol Biol*. 2012;415:680–698.
28. Shunaeva A, Potashnikova D, Pichugin A, et al. Improvement of HIV-1 and human T cell lymphotropic virus type 1 replication-dependent vectors via optimization of reporter gene reconstitution and modification with intronic short hairpin RNA. *J Virol*. 2015;89(20):10591–1601.
29. Kennedy EM, Bogerd HP, Kornepati AV, et al. Posttranscriptional m(6) A editing of HIV-1 mRNAs enhances viral gene expression. *Cell Host Microbe*. 2016;19(5):675–685.
30. Lichinchi G, Gao S, Saletore Y, et al. Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol*. 2016;1:16011.
31. Chang DD, Sharp PA. Messenger RNA transport and HIV rev regulation. *Science*. 1990;249(4969):614–615.
32. Pollard VW, Malim MH. The HIV-1 REV protein. *Ann Rev Microbiol*. 1998;52(1):491–532.
33. Strebel K. Virus-host interactions: role of HIV proteins Vif, Tat, and Rev. *Aids*. 2003;17(Suppl 4):S25–S34.
34. Bevec D, Jaksche H, Off M, et al. Inhibition of HIV-1 replication in lymphocytes by mutants of the Rev cofactor eIF-5A. *Science*. 1996;271(5257):1858–1860.
35. Kula A, Marcello A. Dynamic post-transcriptional regulation of HIV-1 gene expression. *Biology*. 2012;1:116–133.
36. Graf M, Bojak A, Deml L, Bieler K, Wolf H, Wagner R. Concerted action of multiple cis-acting sequences is required for Rev dependence of late human immunodeficiency virus type 1 gene expression. *J Virol*. 2000;74(22):10822–10826.
37. Sodroski J, Rosen C, Goh WC, Haseltine W. A transcriptional activator protein encoded by the x-lor region of the human T-cell leukemia virus. *Science*. 1985;228(4706):1430–1434.
38. Bieniasz PD, Grdina TA, Bogerd HP, Cullen BR. Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. *Proc Natl Acad Sci U S A*. 1999;96(14):7791–7796.
39. Berro R, Kehn K, de la Fuente C, et al. Acetylated tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. *J Virol*. 2006;80(7):3189–3204.
40. Jablonski JA, Amelio AL, Giacca M, Caputi M. The transcriptional transactivator Tat selectively regulates viral splicing. *Nucleic Acids Res*. 2010;38(4):1249–1260.
41. Hashizume C, Kuramitsu M, Zhang X, Kurosawa T, Kamata M, Aida Y. Human immunodeficiency virus type 1 Vpr interacts with spliceosomal protein SAPI45 to mediate cellular pre-mRNA splicing inhibition. *Microbes Infect*. 2007;9(4):490–497.
42. Terada Y, Yasuda Y. Human immunodeficiency virus type 1 Vpr induces G2 checkpoint activation by interacting with the splicing factor SAPI45. *Molecular and cellular biology*. 2006;26(21):8149–8158.
43. Baeyens A, Naessens E, Van Nuffel A, et al. HIV-1 Vpr N-terminal tagging affects alternative splicing of the viral genome. *Sci Rep*. 2016;6:34573.
44. Brosseau JP, Lucier JF, Lamarche AA, et al. Redirecting splicing with bifunctional oligonucleotides. *Nucleic Acids Res*. 2014;42(6):e40.
45. Wang G, Zhao N, Berkhout B, Das AT. CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. *Mol Ther*. 2016;24(3):522–526.
46. Bakkour N, Lin Y-L, Maire S, et al. Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. *PLoS Pathog*. 2007;3(10):e159.
47. Ohe K, Hagiwara M. Modulation of alternative splicing with chemical compounds in new therapeutics for human diseases. *ACS Chem Biol*. 2015;10(4):914–924.
48. Cheung PK, Horhant D, Bandy LE, et al. A parallel synthesis approach to the identification of novel diheteroaryl amide-based compounds blocking HIV replication: potential inhibitors of HIV-1 Pre-mRNA alternative splicing. *J Med Chem*. 2016;59(5):1869–1879.
49. Prado S, Beltran M, Coiras M, Bedoya LM, Alcami J, Gallego J. Bioavailable inhibitors of HIV-1 RNA biogenesis identified through a Rev-based screen. *Biochem Pharmacol*. 2016;107:14–28.
50. Wong R, Balachandran A, Haaland M, Stoilov P, Cochrane A. HIV-1 RNA processing as therapeutic target: characterization of small molecule modulators of HIV-1 RNA processing and transport. *Retrovirology*. 2013;10(1):O45.
51. Niedzielski MF, Hopewell R, Ismail Z, Estable MC. MCEF is localized to the nucleus by protein sequences encoded within three distinct exons, where it represses HIV-1 Tat-transactivation of LTR-directed transcription. *Int J Biol Sci*. 2007;3(4):225–236.
52. D'Orso I, Grunwell JR, Nakamura RL, Das C, Frankel AD. Targeting tat inhibitors in the assembly of human immunodeficiency virus type 1 transcription complexes. *J Virol*. 2008;82(19):9492–9504.
53. Paz S, Krainer AR, Caputi M. HIV-1 transcription is regulated by splicing factor SRSF1. *Nucleic Acids Res*. 2014;42(22):13812–13823.
54. Hwang S, Tamilarasu N, Ryan K, et al. Inhibition of gene expression in human cells through small molecule-RNA interactions. *Proc Natl Acad Sci U S A*. 1999;96(23):12997–13002.
55. Wong RW, Balachandran A, Haaland M, Stoilov P, Cochrane A. Characterization of novel inhibitors of HIV-1 replication that function via alteration of viral RNA processing and rev function. *Nucleic Acids Res*. 2013;41(20):9471–9483.

HIV/AIDS - Research and Palliative Care

Dovepress

Publish your work in this journal

HIV/AIDS - Research and Palliative Care is an international, peer-reviewed open access journal focusing on advances in research in HIV, its clinical progression and management options including antiviral treatment, palliative care and public healthcare policies to control viral spread. The journal is included in PubMed. The manuscript man-

agement system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/hivaids---research-and-palliative-care-journal>