[Frontiers in Bioscience 14, 2714-2729, January 1, 2009]

Role of RNA structure and protein factors in the control of HIV-1 splicing

Jean-Michel Saliou[1], Cyril F. Bourgeois[2,3,4,5], Lilia Ayadi-Ben Mena[1], Delphine Ropers[1], Sandrine Jacquenet[1], Virginie Marchand[1], James Stevenin[2,3,4,5], Christiane Branlant[1]

[1] MAEM, UMR 7567 CNRS-UHP, Nancy Universite, Faculte des sciences et techniques, BP239, Vandoeuvre Les Nancy CEDEX, France, [2] IGBMC, Department of Functional Genomics, Illkirch, F-67400 France, [3] INSERM, U596, Illkirch, F-67400 France, [4] CNRS, UMR7104, Illkirch, F-67400 France, [5] Universite Louis Pasteur, Strasbourg, F-67000 France

TABLE OF CONTENTS

1. ABSTRACT

Alternative splicing plays a key role in the production of numerous proteins by complex lentiviruses such as HIV-1. The study of HIV-1 RNA splicing has provided useful information not only about the physiology of the virus, but also about the general mechanisms that regulate mammalian pre-mRNA alternative splicing. Like all retroviruses, a fraction of HIV-1 transcripts remains intact to serve as genomic RNA and to code for Gag and Gag-Pol protein precursors. In addition, splicing is important for controlling the production of some viral proteins, which could otherwise have a negative effect on the infected cell. Here, we summarize how the utilization of HIV-1 splicing sites is limited by the binding of nuclear factors to *cis*-acting silencer elements, taking into account the role of RNA secondary structure in these mechanisms. We also describe how the poorly efficient HIV-1 acceptor sites are nevertheless activated by serine/arginine-rich proteins. Finally, we discuss how nuclear factors that interact with both the transcription and splicing machineries also participate in the control of HIV-1 RNA splicing.

2. INTRODUCTION

The choice between alternative splice sites is prevalent in higher eukaryotic genomes, since about 75% of the genes are regulated by alternative splicing of their pre-messenger RNAs (1). On the basis of data accumulated over the last 15-20 years, a number of general features have started to emerge to explain how alternative splicing is regulated in cells, leading the scientists from this field to believe that the elucidation of a "splicing code" can now be envisaged (2). One of those general rules is the existence of antagonistic mechanisms that allow either the activation or the repression of alternative splice sites, driven by a variety of *cis*-acting splicing enhancers and silencers. Two main families of proteins are considered as mediators between the *cis* regulatory elements and the spliceosomal machinery: serine/arginine-rich (SR) proteins and hnRNP proteins that recognize splicing enhancers and splicing silencers, respectively, although there are exceptions to this general view (3). In addition, other parameters are known to influence alternative splicing, including the dynamic coupling between transcription and other RNA processing events (4), or the presence of RNA secondary structure (5).

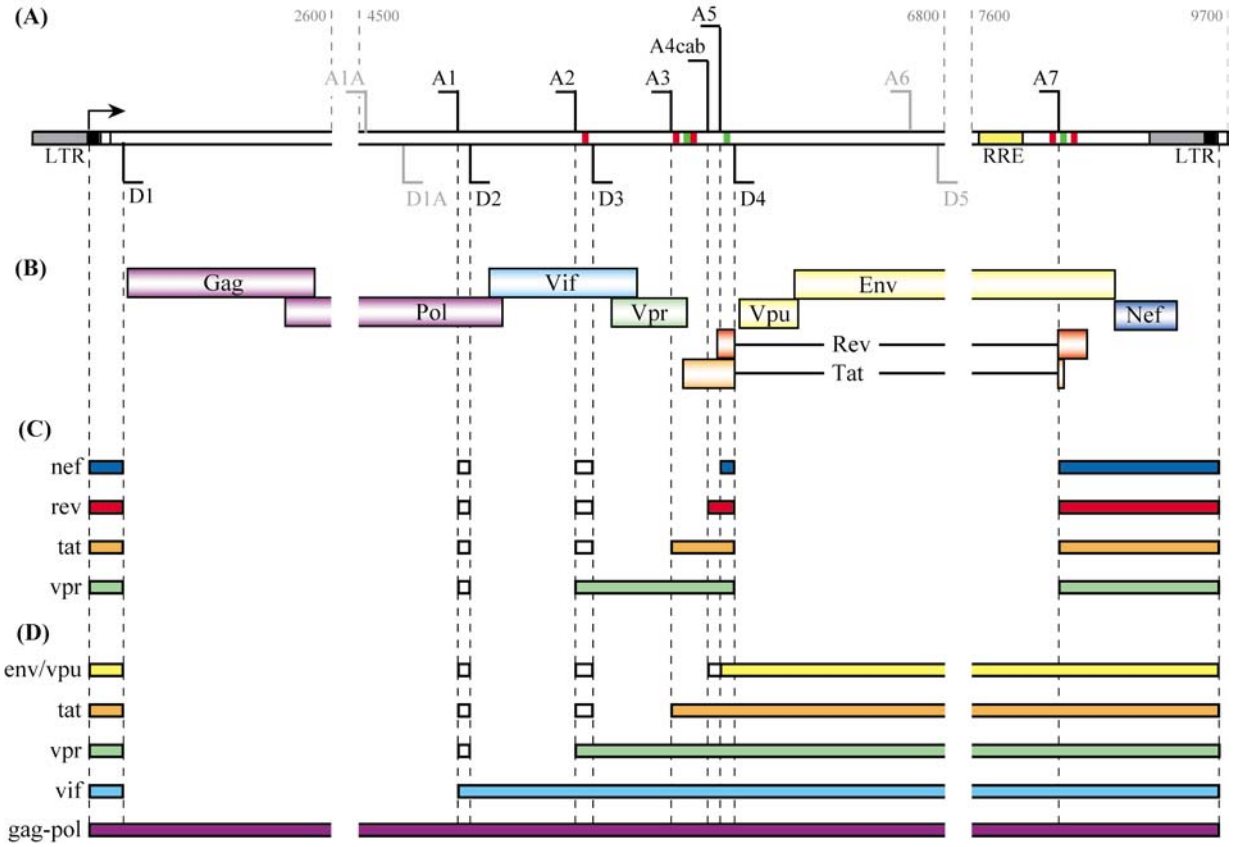

Figure 1. Schematic organisation of the HIV-1 proviral genome. (**A)** HIV-1 genome contains 5 donor splice sites (D1, D1A, D2, D3 and D4) and 9 acceptor splice sites (A1, A1A, A2, A3, A4c,a,b, A5 and A7) (8,11). Sites A1A and D1A have been reported to be preferentially involved in pre-mRNA stabilization (11). In addition, strains from the IIIB family of HIV viruses contain the additional A6 and D5 pair of sites (95). The LTRs which flank HIV-1 DNA in the host genome are composed of U5 (grey box), R (black box) and U3 (white box). The Rev Response Element (RRE) is represented by a yellow box. Little red boxes represent the ESSV, ESS2p, ESS2, ISS and ESS3 splicing silencer elements. Little green boxes represent ESE2, GAR ESE and ESE3 splicing enhancer elements (ESE3 is both an enhancer and a silencer) (38). Regions 2600 to 4500 (downstream from site A1A) and 6800 to 7600 (upstream of site A7) were not represented to simplify the figure. The nucleotide numbering corresponds to the NL4-3 HIV-1 strain. (B) Organization of the open reading frames of the HIV-1 genome. The nine ORFs code for protein precursors Gag, Gag-Pol and Env, regulatory proteins Tat, Rev and Nef, and auxiliary proteins Vif , Vpr and Vpu. (C) and (D) Messenger RNAs produced during early (C) and late (D) phases of HIV-1 infection. White boxes represent exons which are not systematically included in messenger RNAs.

3. OVERVIEW OF HIV-1 ALTERNATIVE SPLICING IN THE CONTEXT OF VIRAL INFECTION

Alternative splicing plays a key role in the complex process of HIV-1 multiplication. After integration of the proviral cDNA into the host genome, RNA polymerase II (RNA Pol II) recognizes the LTR promoter. In the absence of the viral protein Tat, the vast majority of the transcripts that are produced are very short (TAR RNA), because of transcriptional pausing of RNA Pol II (6,7). In contrast, when Tat is present, its association with the TAR element, located at the 5' extremity of the transcripts, allows the recruitment of cyclin T1 and CDK9, components of p-TEFb. Consequently CDK9 phosphorylates the CTD of the large subunit of RNA Pol II and other transcription factors, resulting in the synthesis of full length viral RNA (9229 nucleotides in the case of the HIV-1 Bru strain). This RNA has numerous functions in virus multiplication. It is encapsidated as genomic RNA to produce new viruses and also used as messenger RNA for translation of the Gag and Gag-Pol precursor proteins. On the other hand, it has to be spliced to produce mRNAs encoding the Env precursor protein as well as numerous regulatory and auxiliary proteins (Figure 1B). One of the regulatory proteins, Rev, plays a key role in the transport of a fraction of the unspliced transcripts into the cytoplasm.

HIV-1 RNA splicing is very complex, as HIV-1 RNA contains at least 5 splicing donor sites (5' ss) and 8 to 9 splicing acceptor sites (3' ss) (8-11) (Figure 1A). The combined utilization of these splice sites leads to the production of at least 40 distinct mRNAs in infected cells (8,12). In early stages of the multiplication cycle, transcripts are multiply spliced. They predominantly

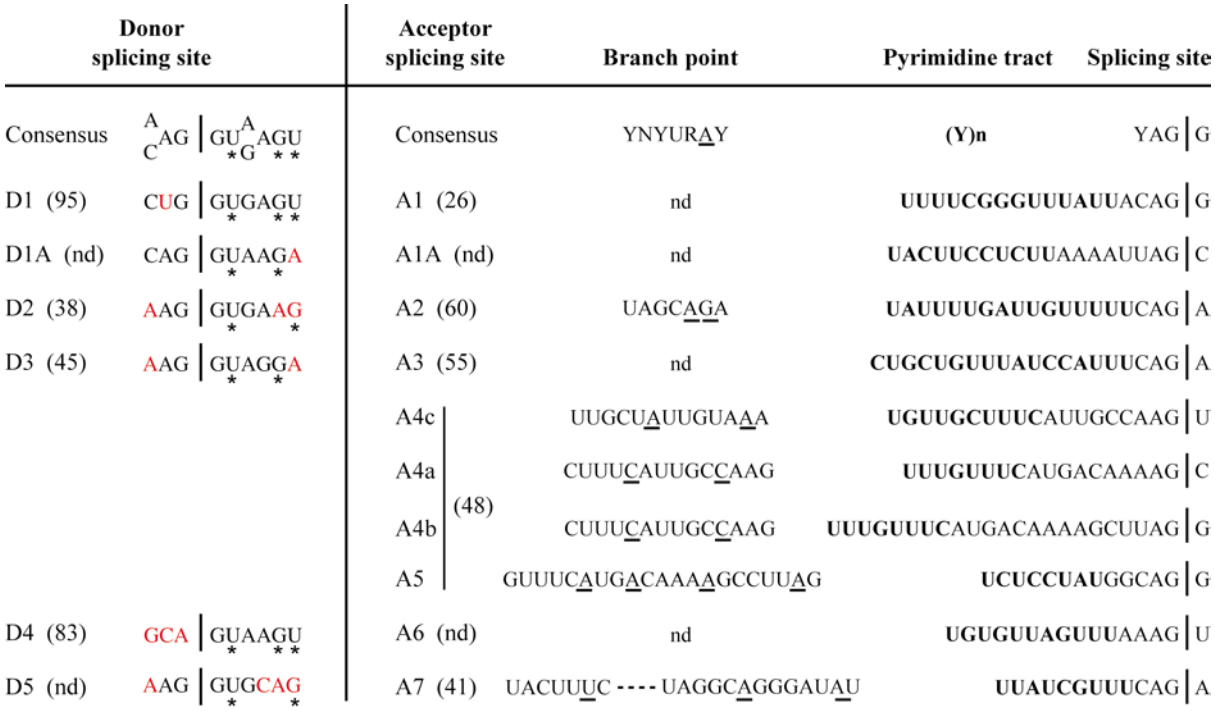

Figure 2. Sequence of the HIV-1 5' (donor) and 3' (acceptor) splice sites. Nucleotides of 5' ss which do not base-pair with U1 snRNA are indicated in red, and those which are expected to base-pair with U6 snRNA are marked by an asterisk. Exon-intron and intron-exon junctions are represented by verticals bars. Numbers in brackets indicate the relative efficiency of utilization (in %) of splice sites *in vivo* (17). Underlined nucleotides represent identified or putative branch points (18,34,35). N=A, C, G or U; Y=C or U; R=A or G.

encode the regulatory Tat and Rev proteins, which are absolutely required for virus multiplication, as well as Nef, whose absence results in a lower efficiency of multiplication (13). During this early stage, alternative splicing of HIV-1 RNA takes place on the basis of the relative strengths of the various 5' and 3' splice sites (Figure 1C).

In a second step of cell infection by virus HIV-1, the Rev protein plays a key role in the fate of viral transcripts. Indeed, when enough Rev has been produced, it recognizes a strong binding site within the Rev Response Element (RRE) and multimerizes on this element, allowing the CRM1-dependent export of unspliced or singly spliced HIV-1 RNAs into the cytoplasm (14). At this late stage of the multiplication cycle, a large fraction of the transcripts is transported in an intact form into the cytoplasm. The singly spliced mRNAs encode the Env precursor and the Vif, Vpr and Vpu auxiliary proteins (Figure 1). Vpr, which has an early function in the transport of the pre-integration complex into the nucleus, plays several important functions for virus multiplication (15). Vif is involved in virus assembly (16). Therefore, at this stage, both the Rev action and the relative strengths of the splice sites D1, D2 and D3 and A1 to A5, define the splicing pattern of HIV-1 RNA (Figure 1D).

Although the HIV-1 5' ss are not optimal, most of them have a rather good complementary with U1 snRNA, especially D1 and D4 (Figure 2). The relative strengths of the donor sites were evaluated both *in vitro* and *in vivo* (17,18), and they correlate well with the level of complementarity with U1 snRNA. This suggests that regulation at the 5' ss is limited and up to now, only a few splicing enhancers were found to act at 5' ss: for example the GAR element located upstream of site D4, binds the SR proteins ASF/SF2 and SRp40, and activates site D4 utilization (19,20).

Site D1 is the most efficient of all the HIV-1 5' ss. All the downstream splicing events depend on its presence since no splicing occurs when it is mutated (21). It is located in an RNA region that contains several important functional elements and has a highly conserved 2D RNA structure in all HIV-1 strains (22,23). A dependency of site D1 efficiency on this peculiar 2D structure was recently shown (24). Recent studies have shown that the low strength of the D2 and D3 5' ss is important to limit splicing at the upstream 3' ss A1 and A2, respectively (25,26) and therefore maintains the required level of unspliced RNA. Interestingly, unusual functions, not related to splicing, have been attributed to some of the HIV-1 5' ss. For instance, site D1 is responsible for preventing premature cleavage and polyadenylation at a 3' processing site located 70 nts downstream from the cap structure (27,28). Another example is the role of site D4 in stabilization of the *env* mRNA, which relies on the base-pairing of U1 snRNA with the 5' ss and on the stabilization

of U1 snRNP by ASF/SF2 bound to the upstream GAR element (19). A recently identified additional D1A site is also proposed to participate to the stability of unspliced RNA (11). In addition, U1 snRNA complementary sequences in HIV-1 RNA were also recently found to facilitate the recruitment of some components of the transcription machinery (29), reinforcing the idea of cross-talk between transcription and splicing complexes (see also Section 9).

All the HIV-1 3' ss are suboptimal (Figure 2). They include suboptimal polypyrimidine tracts (PPT) (30-32) and non-canonical branch-site sequences (33-35). For some 3' ss, two or more branch points can be used, including nucleotides other than adenosine residues and they exhibit moderate or low intrinsic efficiency (31,33,36,37). In addition, 3' ss are frequently included within highly stable 2D RNA structures, as shown for sites A3, A4b,a,c, A5 and A7 (32,38,39). Their utilization is down regulated by the presence of numerous *cis* splicing-silencer elements (ESS) that bind hnRNP proteins, most frequently hnRNP A1/B (38,40-45). By counteracting all these negative characteristic features of HIV-1 3' ss, SR proteins (ASF/SF2, SC35, SRp40 or 9G8) allow viral RNA splicing (35,44,46-48). However, this is a slow process as compared to some efficiently spliced cellular pre-mRNAs, so that RNA folding is likely to take place prior to splicing. In addition, HIV-1 RNA splicing is an ordered process, as introns are eliminated according to their order of appearance, starting from the 5' extremity of the RNA molecule (21). Furthermore, the equilibrium between spliced and unspliced RNAs depends also on the presence of the suboptimal D2 and D3 5' ss (25,26).

Studying the complex regulation of HIV-1 RNA alternative splicing is very important to fully understand the biology of this virus and to define new therapeutic strategies to limit its multiplication. It also represents a good model to study the mechanisms of alternative splicing in vertebrates. Three of the eight 3' ss (A2, A3 and A7), which are conserved in all HIV strains, are subjected to very tight regulation. They have been intensively studied, by us and others: sites A2 and A3 are used for *vpr* and *tat* mRNA production, respectively. Site A7 is used for both *tat* and *rev* mRNA production (Figure 1). A competition between hnRNP A1/B and SR proteins was found to occur at these three sites. However, the mechanisms involved at each of these sites display specific characteristics, due to a different organization of the protein binding sites and to the influence of the RNA 2D structure. In the next part of the review, we will focus on these 3 splice sites and will highlight the role of the HIV-1 RNA secondary structure that is known to be crucial for the synthesis of full-length RNA (TAR element) and for the export of unspliced RNA (RRE). When appropriate, we will also complete our review with some additional unpublished data. For complementary information see the review "Modulating HIV-1 RNA processing and utilization" by McLaren *et al.* in this series and a previously published review by Stoltzfus and Madsen (49).

4. A SMALL REGULATORY ELEMENT THAT FOLDS INTO A STEM-LOOP STRUCTURE CONTROLS SITE A2 UTILIZATION

MT Stoltzfus and colleagues first identified a 24-nt sequence downstream from site A2 in the HIV-1 RNA. This sequence, containing three PyUAG motifs (numbered 1, 2 and 3 in Figure 3) was called ESSV (V for Vpr) because its binding to hnRNP A1 was found to have a strong negative effect on site A2 utilization (41) (Figure 3). They also showed that this negative effect on site A2 results from the limited association of U2AF65 to the PPT (50). In parallel, we found that the SR protein ASF/SF2 strongly activates site A2 utilization, both *in vitro* and *in cellulo* (48). Among the four SR proteins that we tested *in cellulo* (SC35, SRp40, 9G8 and ASF/SF2), only ASF/SF2 strongly favoured site A2 utilization at the expense of sites A1, A3, A4b,a,c and A5 (48) and this was confirmed by the Darlix team, using an HIV-1 pNL4-3 clone (51). In agreement with these data, the IDC16 indole derivative that specifically suppresses splicing activation by ASF/SF2 is a strong inhibitor of site A2 utilization *in vivo* and *in vitro,* and the addition of a large excess of recombinant ASF/SF2 *in vitro* can compensate partially the inhibitory effect of IDC16 (52). More recently, Madsen and Stoltzfus refined the ESSV limits by generation of dinucleotide substitutions in the 24-nt sequence of ESSV within plasmid pΔPSP (25). This plasmid contains an HIV-1 pNL4-3 cDNA truncated within the D1-A1 intron, but it retains all the HIV-1 splice sites which have the same relative efficiency as in the entire HIV-1 RNA (32,41). ESSV was redefined as a 16-nt element containing PyUAG motifs 2 and 3, and an important AUAG element (motif 4 in Figure 3). Mutations in the (Py/A)UAG elements 2-4 have a strong positive effect on splicing at site A2. In contrast, mutations in the sequence between elements 1 and 4 have a negative effect on splicing, suggesting that they contain an enhancer element (25). Importantly, when some of the mutations in the restricted ESSV were generated in the entire pNL4-3 HIV-1 virus, replication was reduced by 95% and exon 3 inclusion was increased as expected from the increased A2 activity. In addition, the levels of unspliced RNA and p55-Gag were strongly reduced (25).

We recently obtained additional structural information which explains the competitive action of hnRNP A1 and ASF/SF2 at site A2 (submitted data). First, the 2D structure of the HIV-1 RNA region, which contains exon A2-D3 and both the functional and regulatory elements of site A2, was established *in vitro,* using chemical and enzymatic probes in conditions of *in vitro* splicing assays (Figure 3). As shown in figure 3, exon A2-D3 folds into a stem-loop structure (SLS 4-A2). The possibility to form this 2D structure is highly conserved in HIV-1 strains. In the established structure, the identified A2 branch-point sequence (BP) is involved in a stable base-pair interaction. In contrast, the polypyrimidine tract (PPT) is only partially base-paired and site A2 is located at the junction between two helices. The 4 (Py/A)UAG elements of the ESSV element are located in the apical part of SLS 4-A2 (Figure 3).

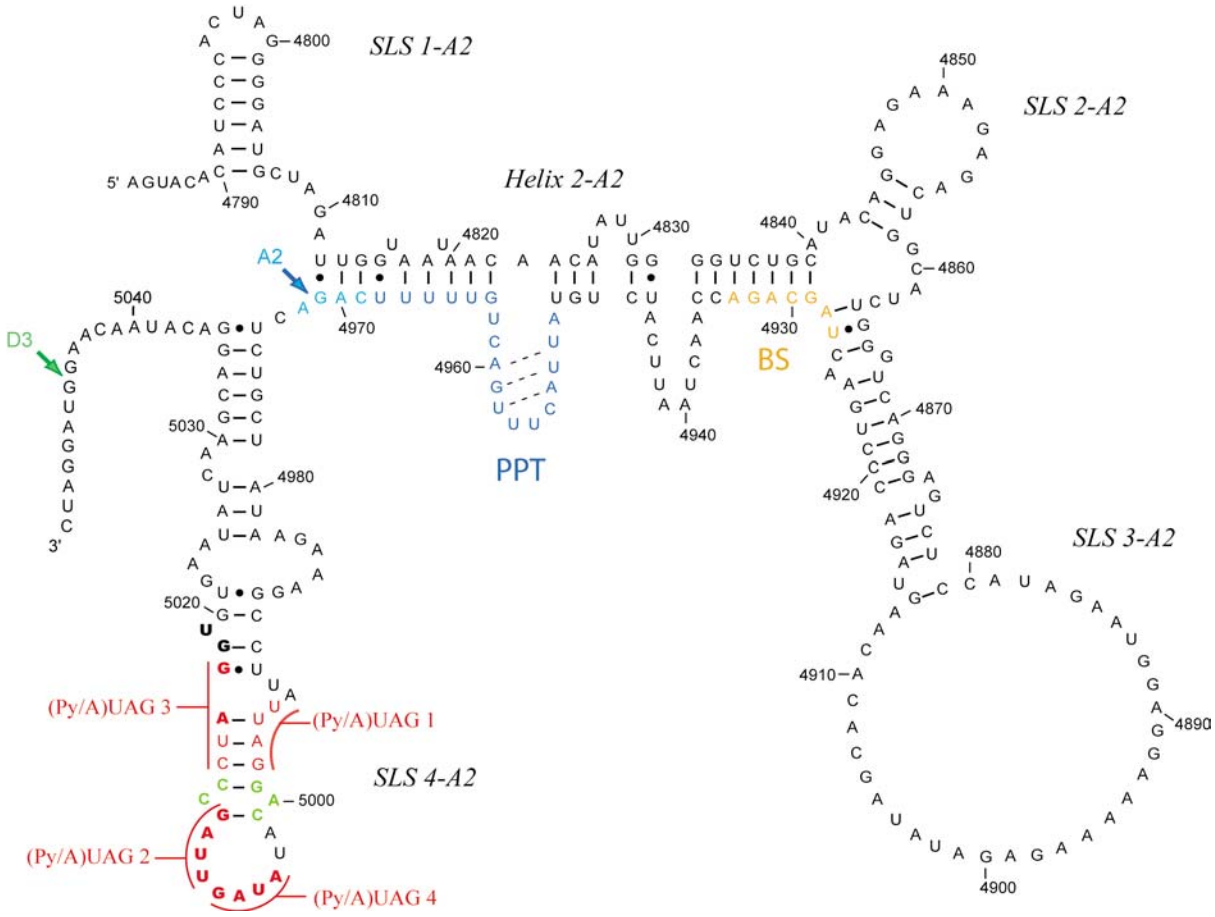

Figure 3. Proposed secondary structure for the HIV-1 Bru region containing the A2 3' ss and its functional and regulatory element. The structure of stem-loop sequences SLS 1,2,3,4-A2 and helix 2-A2 was defined by enzymatic and chemical probing (submitted data). The branch-site sequence (BS) is in orange. The polypyrimidine tract (PPT) and the A2 3' ss are in blue, site D3 is in green and the intron-exon and exon-intron junctions are shown by arrows. Sequences in red correspond to the three PyUAG elements numbered 1, 2 and 3 initially identified as putative hnRNP A1 binding sites in ESSV (41) and the AUAG motif ((Py/A)UAG 4) that is important for splicing inhibition *in cellulo* together with elements 2 and 3 (25). Residues required for ESSV activity according to Madsen and Stoltzfus are in bold (black or red), and residues which may correspond to an enhancer element are in green.

A precise delineation of the hnRNP A1 binding sites on SLS 4-A2 was obtained by footprinting analysis of complexes formed between a HIV-1 Bru RNA fragment (4807 to 5039) and recombinant hnRNP A1 (data submitted). As expected, protections were detected in the apical part of SLS 4-A2 that contains ESSV and the strong protection of the G residue 5007 in the AUAG motif confirmed hnRNP A1 binding to this motif. In addition, the 5' strand of the internal loop I in SLS 4-A2 (4983 to 4989), as well as site A2, were also protected (Figure 4A). In contrast, the entire PPT sequence became highly accessible in the presence of hnRNP A1, suggesting that the previously demonstrated inhibition of U2AF65 binding to PTT (50) does not result from a simple masking of the PTT by hnRNP A1. A complete protection of the SLS 4-A2 region from position 4970 to 5018, including the intron-exon junction, was also observed when RNA probing was performed using HeLa cell nuclear extract rather than pure hnRNP A1 (data submitted). Once again, only a very limited protection of the PTT sequence was obtained (Figure 4A) in agreement with Stoltzfus's data showing a weak binding of U2AF65 (50). These data demonstrate that competition between hnRNP A1 and splicing factors takes place at the intron-exon junction and not at the PPT.

Several polypurine-rich sequences known to be recognized by ASF/SF2 (53,54) are present in the intron sequence located upstream of site A2. However, their mutation or deletion had no effect on splicing efficiency at site A2, or on its activation by ASF/SF2 (submitted data). In contrast, AG to CU substitutions in the (Py/A)UAG 2, 3 and 4 of ESSV increased the splicing activity of the resulting substrate *in vitro* and *in cellulo* (25). For a more precise delineation of the ASF/SF2 binding sites, the complexes formed between an HIV-1 Bru RNA fragment (4986 to 5010) and recombinant ASF/SF2 were analysed

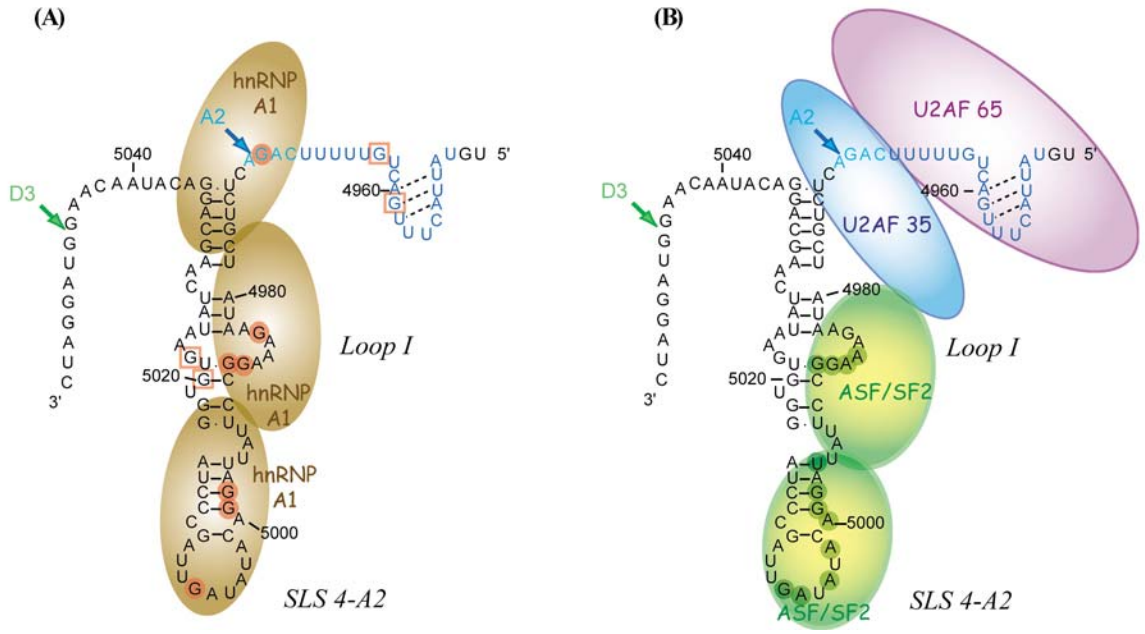

Figure 4. Proposed model for hnRNP A1 and ASF/SF2 binding to the HIV-1 RNA region containing the 3' ss A2. (A) Schematic model of binding and silencing activity of hnRNP A1 based on footprinting data on complexes formed with recombinant hnRNP A1 (submitted data). (B) Schematic model of binding and enhancing activity of ASF/SF2 based on footprinting data on complexes formed with recombinant ASF/SF2 (submitted data). Nucleotides in full circles correspond to nucleotides which are protected by hnRNP A1 (T1 RNase, panel A) or ASF/SF2 (chemical modifications and RNAse digestions, panel B). In panel A the nucleotides in squares show an increased reactivity to T1 RNase in the presence of hnRNP A1. The polypyrimidine tract (PPT) and 3' ss A2 are in blue.

by enzymatic footprinting assays (data submitted). Strong protections were found to extend from position 4986 to 5008 including the AUAG motif 4 of ESSV, suggesting the binding of at least 2 molecules of ASF/SF2 (Figure 4B). In agreement with Madsen and Stoltzfus's data, ASF/SF2 interacts with the segment linking motifs 1 and 4 that was shown to decrease splicing efficiency upon mutation in plasmid pΔPSP (25). Within the ASF/SF2 protected sequence, only the AAAGGC sequence (4985 to 4990) shows a good homology with the high affinity ASF/SF2 binding sites defined by Tacke and Manley, but we did not identify any sequence matching to the C/G-rich functional ASF/SF2 binding sites as defined by Smith *et al.* (55,56).

In conclusion, altogether the data show that the 50-nt long segment located downstream from site A2 contains the overlapping silencer and enhancer acting at site A2. HnRNP A1 is able to cover the entire segment, but it does not mask the PPT sequence. As this sequence is suboptimal, its stable recognition by U2AF65 is expected to require the binding of U2AF35 to the intron-exon junction. However, binding of hnRNPA1 to this junction might preclude U2AF35 association and therefore impair the stable binding of U2AF65 to the PTT. In contrast, when ASF/SF2 binds downstream of site A2, it may on the one hand compete out hnRNP A1 binding (RRM-dependent effect). On the other hand, through an interaction of its RS domain with those of both U2AF subunits, it may favour the binding of U2AF35 and U2AF65 to the intron-exon junction and to the PPT, respectively, and activate splicing at site A2. An involvement of both the RRM and the RS domain of ASF/SF2 is in agreement with our previous *in cellulo* data showing that over-expression of an ASF/SF2 protein truncated of its RS domain, in the presence of plasmid pΔPSP, leads to an increase of *vpr* mRNA production in HeLa cells, but at a level lower than that obtained when the full-length protein is over-expressed (48).

5. A COMPLEX EXTENDED EXON REGION REGULATES SITE A3 UTILIZATION

Two splicing silencer elements have been found to limit site A3 utilization: the ESS2p element, located very close to site A3, that binds protein hnRNP H (32), and the ESS2 element, located farther downstream (70 nts downstream of site A3), that binds hnRNP A1 (40,57). Our previous structural analysis demonstrated that site A3 is located in the terminal loop of a stem-loop structure called SLS 2-A3 (32) (Figure 5). Both the PPT and ESS2p element are involved in the stem (5' and 3' strands, respectively). The PPT is intrinsically suboptimal and its base-pair interaction with the exon sequence likely limits its accessibility (32) (Figure 5). Hence, as seen for site A2, the efficient binding of U2AF65 to the PPT of site A3 should depend on the association of U2AF35 with the intron-exon junction. We previously proposed that the binding of hnRNP H to this element might limit U2AF35 association with site A3 by steric hindrance, explaining how ESS2p decreases site A3 activity (32).

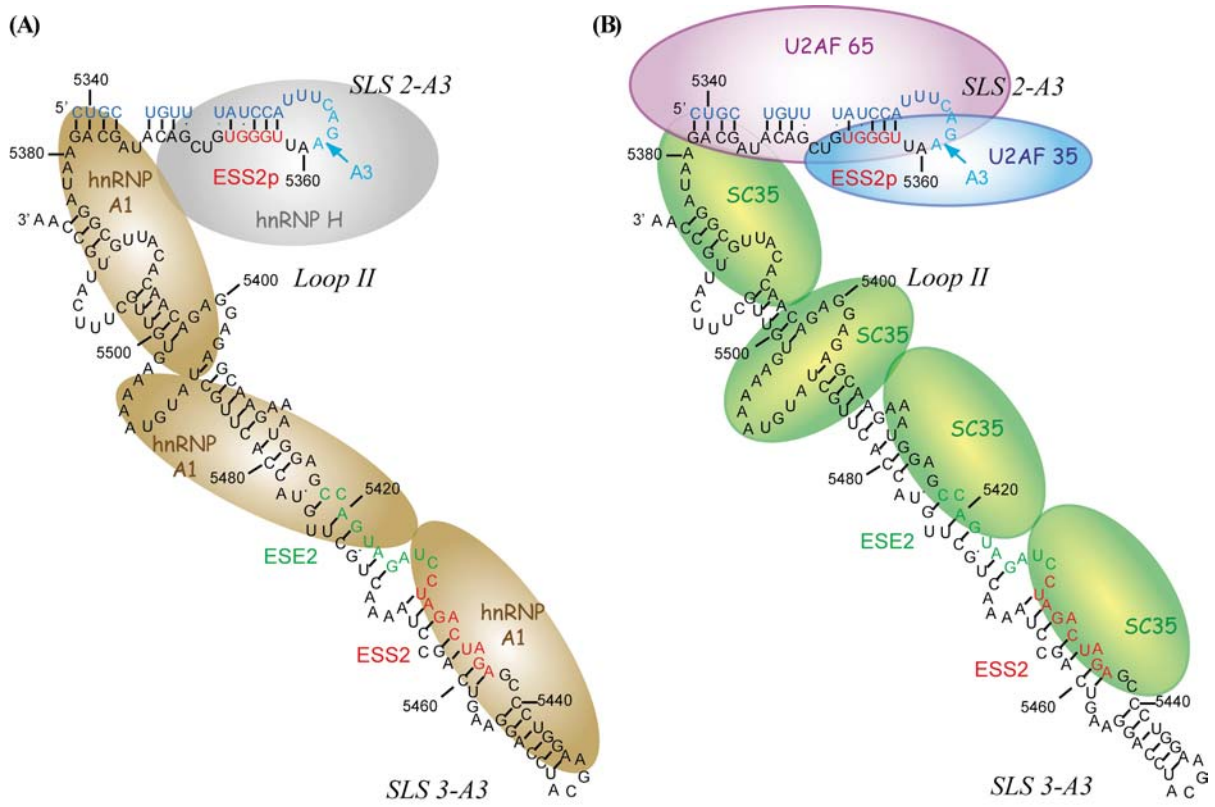

Figure 5. Two dimensional structure of the HIV-1 RNA region containing the 3' ss A3 and model for site A3 silencing by hnRNP proteins and enhancing by protein SC35 (46,47). Red and green sequences represent the splicing silencer (ESS2p and ESS2) and the splicing enhancer (ESE2), respectively. The polypyrimidine tract (PPT) and splice site A3 are in blue. (**A)** Model of site A3 inhibition by hnRNP A1 and H. (**B)** Model of site A3 activation by SC35.

Using several molecular approaches as well as NMR analyses, we recently succeeded in explaining how hnRNP A1 binding to ESS2 strongly limits site A3 activity, despite its far downstream location (47). According to the 2D structure that we previously established (32), ESS2 is located in the apical part of a long irregular stem-loop structure (SLS 3-A3) (Figure 5). By site-directed mutagenesis of SLS 3-A3, footprinting assays of complexes formed between hnRNP A1 and wild type or mutated RNAs, as well as gel-shift analyses, we showed that ESS2 plays a key role in the association of hnRNP A1 with the entire SLS 3-A3 region and with the junction between SLS 2-A3 and SLS 3-A3 (47). We also provided detailed information on how hnRNP A1 associates with this HIV-1 RNA region.

HnRNP A1 contains two RNA Recognition Motifs (RRMs) and a Gly-rich domain that allows its multimerization (58,59). The latter domain was also proposed to stabilize the interaction of hnRNP A1 with nucleic acids (60). By NMR analysis, we demonstrated the direct binding of the two RRMs of hnRNP A1 (UP1 protein) to ESS2 and to the terminal stem-loop of SLS 3-A3 but surprisingly, this interaction did not disturb the RNA 2D structure (47). By enzymatic footprinting experiments using complexes formed between an RNA containing SLS 2-A3 and SLS 3-A3 (SLS 2,3-A3 RNA) and increasing amounts of either hnRNP A1 or UP1, we showed that in both cases, two molecules of protein bind simultaneously to the SLS 3-A3 structure: one interacts with the two CUAG motifs of ESS2 and the terminal stem-loop, while the second one binds to an upstream UAG motif contained in the recently identified ESE2 enhancer (46) and to the 3' strand of the internal loop II in SLS 3-A3 (Figure 5A). We propose that this simultaneous binding takes place through an interaction of the RRM domains of UP1 or full-length hnRNP A1 (47). However, the stability of the assembled complex is higher for hnRNP A1 than for UP1. This means that, as previously suggested (58,59), the Gly-rich domain of hnRNP A1 can stabilize its interaction with RNA, either by electrostatic interaction with RNA or by establishing interactions with another nearby Gly-rich domain.

In addition, we observed that mutating one of the two UAG motifs in ESS2 is sufficient to strongly reduce hnRNP A1 binding to the entire SLS 2,3-A3 RNA (47). Taken together, our data support the following model for hnRNP A1 interaction at site A3: hnRNP A1 uses an entry site with highly characteristic features (ESS2) that allows, together with the surrounding sequences, the establishment of a stable RNA-protein interaction involving two hnRNP A1 molecules. Then, other molecules of hnRNP A1 can interact with sites of lower affinity, the interaction being

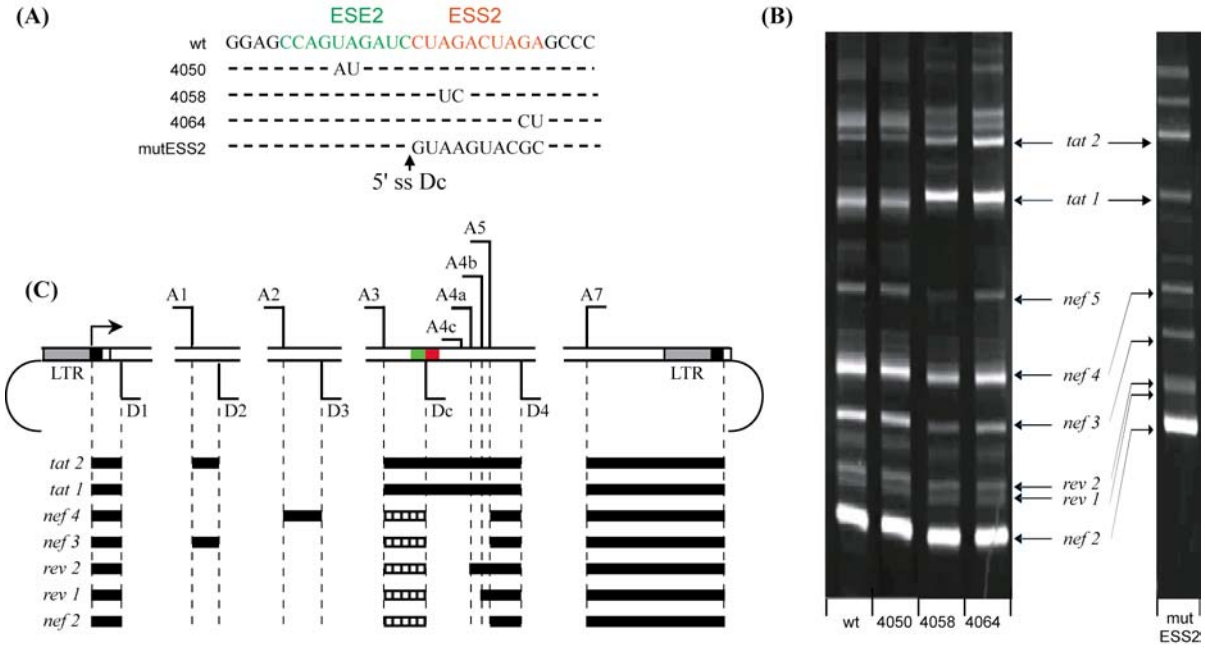

Figure 6. Effect of mutations in ESE2 and ESS2 on the splicing pattern of HIV-1 RNA. (**A)** Mutations introduced in the pΔPSP plasmid used in this study (42). Sequences of ESE2 and ESS2 that are located downstream from site A3 are in green and red, respectively. The mutant with substitution of most residues in the ESS2 element (mutESS2) was previously analyzed (40,46,57). (**B)** RT-PCR analysis of HIV-1 spliced products. HeLa cells were transfected with the pΔPSP plasmid. For each transfection assay, equal amounts of total RNA were subjected to RT-PCR analysis as previously described (48). The identity of the different mRNAs was confirmed by sequencing when necessary. (**C)** Schematic representation of HIV-1 genome and exon composition of the different HIV-1 mRNAs. For each mRNA, exons are represented by black boxes. Hashed boxes represent the ectopic exon which is included in the *nef* and *rev* mRNAs when the donor splice site Dc is introduced in the mutESS2 construct.

stabilized by protein-protein interactions through the Gly domains. This results in the wrapping of SLS 3-A3 and the SLS 2-A3/SLS 3-A3 junction with hnRNP A1, probably occluding directly the access to the PPT (Figure 5A).

In spite of the strong negative effect of hnRNP A1, site A3 is used in infected cells (8). We previously showed that its activity is highly increased when the SR proteins SC35 or SRp40 are over-expressed in HeLa cells or when recombinant SC35 or SRp40 protein is added in *in vitro* splicing assays (48). This was in agreement with the description of an SC35-dependent splicing enhancer element (ESE2) adjacent to ESS2 (46) and with results obtained in infected cells (51). By exhaustive footprinting analysis of complexes formed between the SLS 2,3-A3 RNA region and recombinant SC35 or SRp40 protein, we showed that the splicing activation can be explained by binding of these two proteins on several sites which overlap the hnRNP A1 binding sites (47). Interestingly, we found that the strongest SC35 binding site corresponds to the 5' strand of internal loop II in SLS3-A3 (Figure 5B). However, its deletion only reduced the splicing efficiency by a factor of 2.5 (47). Although the affinity of SC35 for ESE2 and ESS2 is lower than for loop II, the strong activation observed upon SC35 and SRp40 binding to ESE2 and ESS2 is due to a direct competition with hnRNP A1 binding to its key entry site in the SLS 2,3-A3 region. These data illustrate the fact that strong SR protein binding sites located downstream from splice sites may have limited functions in splicing regulation. However, they may have other functions, since SR proteins are known to be involved in exon junction complex (EJC) formation, in mRNA transport and in activation of translation (61,62). Therefore, when analysing splicing regulation, it is important to keep in mind that strong SR protein binding sites in exons may not be sufficient *per se* to define splicing enhancers.

In conclusion, the strong inhibitory activity of ESS2 is due to the capacity of this sequence to generate a complex array of hnRNP A1 interactions on the entire RNA region that extends up to the PPT. Accordingly, mutation in each of the UAGA motifs within ESS2 resulted in the disruption of hnRNP A1 binding to the entire region and to an improved splicing efficiency at site A3 (47). The data previously obtained *in vitro* using mini HIV-1 RNA substrates (47) or *in cellulo* using chimeric RNAs (63) were completed here by introducing the mutations in the pΔPSP plasmid (42) to evaluate their effects in an RNA containing all HIV-1 splice sites. Whereas the UA to AU dinucleotide substitution in ESE2 (variant 4050) had no detectable effect on the relative levels of *tat*, *rev* and *nef* mRNAs, dinucleotide mutations in the two UAG motifs of ESS2 (variants 4058 and 4064) resulted in a strong increase of the steady state levels of *tat1* and *tat2* mRNAs, at the expense of *nef3* to *nef5* mRNAs (Figure 6B). These data highlight the central role of the two UAG motifs of ESS2 in site A3 regulation. Accordingly, over-expression of SC35 or

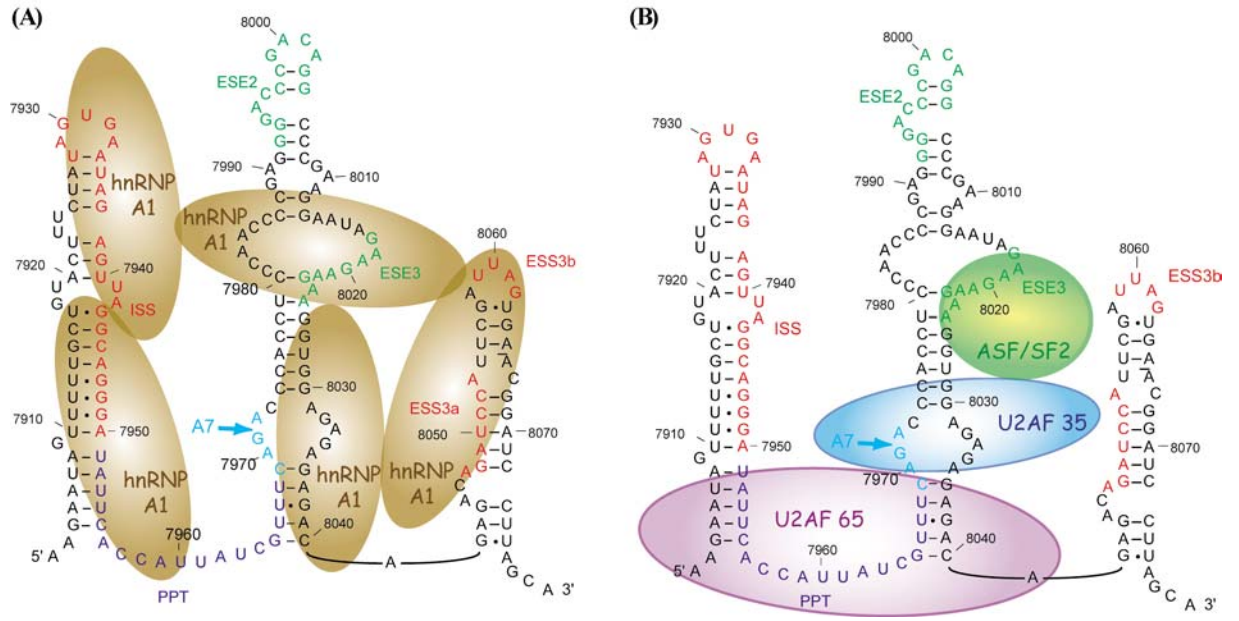

Figure 7. Two dimensional structure established for the HIV-1 region containing the 3' ss A7 and a model of silencing by hnRNP A1 and enhancing by ASF/SF2. Splicing silencers ISS, ESS3a and ESSb are represented in red (31,43,44). Splicing enhancers ESE2 and ESE3 are in green (35,39,64,96). ESE3 is both an enhancer and a silencer element according to footprinting and splicing data (38). The polypyrimidine tract (PPT) and splice site A7 are in blue. (**A)** Model of site A7 inhibition by hnRNP A1. (**B)** Model of site A7 activation by ASF/SF2.

SRp40 in HeLa cells transfected with the mutated constructs 4058 and 4064 did not result in a further increase of *tat* mRNAs (unpublished results). This finding reinforced the idea that the main function of SC35 is to antagonize the action of hnRNP A1 initiated by its binding to ESS2. In previous studies, the function of ESS2 was also tested on chimeric RNAs in which this entire element had been mutated (40,46,57). Hence, we also transferred this ESS2 mutation in the pΔPSP plasmid and looked at the splicing pattern in HeLa cells (Figure 6). Surprisingly, the splicing pattern was found to be profoundly modified (Figure 6B). By cDNA sequencing, we found that mutation of the original CUAGACUAGA ESS2 sequence to the variant GUAAGUACGC sequence created a new strong 5' ss in ESS2 that we called Dc. As Dc is located between the acceptor sites A3 and A4c, it is used in all the transcripts spliced at sites A4c, A4a, A4b and A5 (Figure 6C). Therefore, the results obtained for site A3 using this mutant might be biased by the generation of this new 5' ss and should be reanalysed carefully.

6. SITE A7 IS REGULATED BY DOWSTREAM AND UPSTREAM ELEMENTS

The numerous studies performed on site A7 revealed the presence of an intronic silencer sequence (ISS) that overlaps the branch site sequences (43), and a bipartite exonic silencer (ESS3), which binds hnRNP A1 (30,31,44,45,64). One exonic enhancer (ESE3) was proposed to bind ASF/SF2 (30,64). Our group and the Kjems group established the secondary structure of the A7 region of HIV-1 RNA (38,39). In this structure, ISS, ESE3 and ESS3 are located in three distinct stem-loop structures, SLS 1-A7, SLS 2-A7 and SLS 3-A7, respectively (Figure 7). In addition, our two groups showed that ESE3 constitutes a nucleation site for binding of hnRNP A1 to the entire HIV-1 region including the 3 SLSs (38,39). In agreement with previous data (44), footprinting assays revealed a multimerization of hnRNP A1 on the segment joining ESE3 to ESS3 (38,39) (Figure 7A). In this case, hnRNP A1 binds cooperatively to both exonic and intronic sequences and allows the propagation of its binding to the entire A7 region. Therefore, site A7 and the branch-site sequences are poorly accessible and splicing is highly reduced. In contrast, ASF/SF2 association with ESE3 is expected to destabilize the entire hnRNP A1 association to this region, allowing the assembly of splicing complexes. We consequently described ESE3 as a Janus element since it has at the same time a negative and a positive activity (38). The apical part of SLS 2-A7 was proposed to have an ASF/SF2-dependent enhancer activity on site A7 (35). However, our recent purification of complexes assembled onto SLS 2-A7 in a nuclear extract suggests the binding of other nuclear proteins at this site (unpublished data).

Finally, in addition to be involved in the modulation of site A7 activity, the ESS3a motif of the bipartite ESS3 silencer element was proposed to limit the Rev-mediated RNA transport via the binding of a cellular factor, although this factor was not identified (45,65). Recent analyses identified hnRNP E1 as a potential ESS3a-binding factor. However, hnRNP E1 was not found to modulate site A7 utilization, but rather to block translation of viral RNAs (66).

7. REGULATORY ELEMENTS LOCATED IN OTHER HIV-1 REGIONS

In addition to the RRE, whose secondary structure has been solved, splicing regulatory elements were detected in two other regions of HIV-1 RNA whose 2D structure has not yet been studied. First, one bidirectional element that binds both ASF/SF2 and SRp40 was found between the A5 3' ss and the D4 5' ss (19,20). SR protein binding to this enhancer is required for efficient utilization of site A5 and for association of the U1 snRNP to site D4. As mentioned above, the ASF/SF2-mediated binding of U1 snRNP to site D4 is required to stabilize the *env* mRNA (20). However, it should be noted that upon ASF/SF2 over-expression in HeLa cells transfected with the pΔPSP plasmid or in infected cells, no increase of the utilization of site A5 was detected. Only site A2 utilization was increased with the appearance of a larger amount of *vpr* mRNAs at the expense of *nef* mRNAs and, accordingly, production of Nef protein was reduced (48,51). Similarly, upon SRp40 over-expression, site A3 was predominantly activated at the expense of the other sites in the A3 to A5 cluster of 3' ss. Here again, this resulted in the reduced production of Nef (48,51). Therefore, among the A1 to A5 3' ss, sites A2 and A3 are the most strongly activated by ASF/SF2 and SRp40, respectively. Either the ESE element located downstream from site A5 is just used to maintain a basal utilization of site A5 upon ASF/SF2 increase, or the simultaneous binding of the ASF/SF2 and SRp40 proteins to this site is required for activation of site A5. To address this question, it will be interesting to cotransfect HeLa cells with plasmids over-expressing both ASF/SF2 and SRp40, in the presence of plasmid pΔPSP. Altogether, these data illustrate how results obtained using a pre-mRNA containing a few splice sites have to be compared with data obtained in the full genomic context. The possible concerted action of some SR proteins at given sites should also be considered.

As mentioned above, the low efficiency of the downstream 5' ss D2 and D3 is also involved in the low utilization of the A1 and A2 3' ss, respectively (25,26). The Stoltzfus team recently showed that both a GGGG silencer located immediately downstream from site D2 and an enhancer element found downstream from site A1, which binds the SR protein SRp75, modulate site D2 utilization and consequently site A1 utilization (67). The partner of the GGGG silencer has not yet been identified.

8. HNRNP A1 IS A GENERAL CELLULAR DOWN-REGULATOR OF HIV-1 RNA SPLICING

Taking into account its negative action at most of the HIV-1 3' ss and at some of the 5' ss, hnRNP A1 protein can be considered as one of the most essential cellular factors for limiting HIV-1 RNA splicing, and thus facilitating the Rev-dependent RNA transport of unspliced or partly spliced RNAs. As shown previously for regulation of cellular 5' ss (68), hnRNP A1 regulates HIV-1 3' ss by interacting with an entry site which contains at least one UAG motif. This is followed by a multimerization of hnRNP A1 along the RNA region. Interestingly, the UAG motifs in the entry site may be partially double stranded, like in ESSV and ESS2, but such folding does not prevent hnRNP A1 binding, as shown by our NMR analysis (see Section 5) (47). According to HIV-1 probing data, the presence of polypurine sequences may favour the propagation of hnRNP A1 along the RNA. Such polypurine sequences, especially when they are single-stranded, can also be recognized by SR proteins ASF/SF2, SC35 and SRp40, limiting hnRNP A1 propagation. Noticeably, whereas several of the HIV-1 splicing sites are activated by ASF/SF2, activation by SC35 is mostly restricted to *tat* mRNA production through the activation of the A3 3'ss (48,51). In line with this observation, the SC35 concentration in macrophages is increased during the first two weeks after infection, whereas the hnRNP A/B and H concentrations are decreased (69). Around the peak of virus production the opposite situation is observed. Concentration of hnRNP A/B proteins increases, while that of SC35 decreases. Therefore, soon after infection, the virus could alter the expression of host genes so that the level of splicing inhibitory proteins is reduced and that conditions are optimized to activate the synthesis of the tat mRNAs, allowing virus replication. Note that an increased expression of SC35 was also demonstrated two days after HIV-1 infection of the H9 lymphocyte cells (70), whereas expression of 9G8 was downregulated after infection of T cells (71). Interestingly, a severe dephosphorylation of some SR proteins, which is known to be detrimental for their function was observed in the course of HIV-1 infection (72). A similar phenomenon was also observed after infection by adenovirus or herpes virus (73,74, Akusjarvi in this series).

In contrast to hnRNP A1, hnRNP H was only found to play a negative role at the A3 3' ss (42). Moreover, intronic hnRNP H binding sites that have a positive role on splicing were recently identified upstream of site D3 (75), and it was proposed that proteins of the hnRNP F/H family may loop out the RNA and therefore favour some splicing events, by a mechanism similar to that found for proteins of the hnRNP A1/B family (76). Similarly, it should also be noted that in the strains from the IIIB family, activation of the additional A6D site depends on an enhancer that is bound by both hnRNP H and SC35 (77). Finally, another cellular protein, p32, also plays an essential role in HIV-1 splicing inhibition, through its interaction with Tat (see below).

9. VIRAL PROTEINS MAY LIMIT HIV-1 RNA SPLICING

By enabling the nuclear export of unspliced or partly spliced HIV-1 RNAs, Rev can be considered as a factor which allows the viral pre-mRNA to escape from the splicing process. However, how Rev interferes with spliceosome assembly remains unclear (49). Besides Rev, other viral proteins have the potential to influence splicing. Indeed, it was recently shown that Vpr, through its interaction with the SAP145 splicing factor, may limit cellular pre-mRNA splicing by interfering with the formation of the SAP145-SAP49 complex (78,79). Therefore, Vpr might also modulate viral RNA splicing.

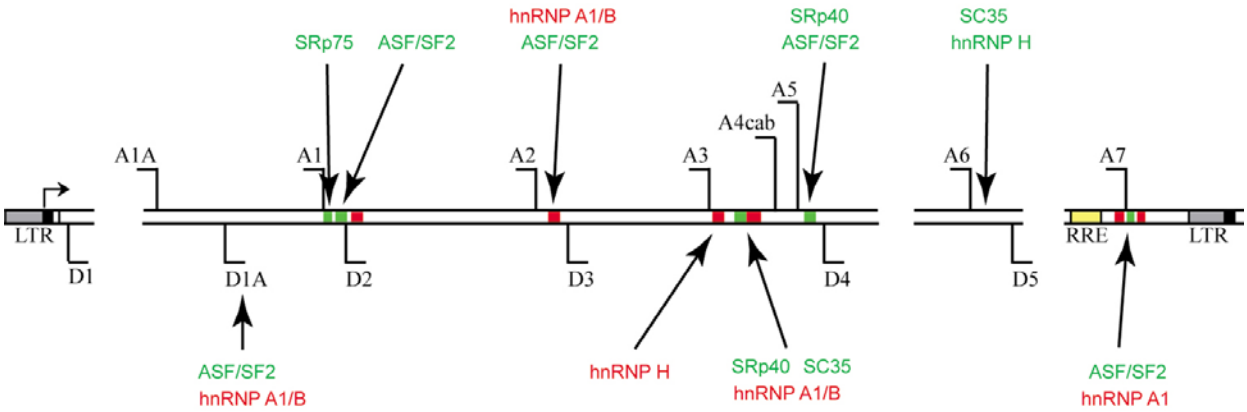

Figure 8. General overview of the regulation of HIV-1 RNA splicing by negative and positive nuclear factors. *Cis*- and *trans*-acting splicing activators and repressors are in green and red, respectively. Little red boxes represent (from left to right) GGGG (downstream from D2), ESSV, ESS2p, ESS2, ISS and the ESS3 silencers. Little green boxes represent (from left to right) the SRp75-binding ESEVif and the ESE-M1/M2 bound by ASF/SF2 (upstream of site D2), ESE2, GAR ESE and ESE3.

The Tat protein can modulate HIV-1 RNA splicing through various mechanisms that largely depend on the coupling between transcription and splicing. For instance, a recent study showed that replacing the HIV U3 region of the 5' LTR by a CMV promoter, in a simplified *env*/*nef* reporter (extending from D4 5' ss to the 3' LTR, making Tat synthesis impossible) led to more efficient splicing, and this effect was reversed upon co-expression of Tat (80). In line with this, another report showed that Tat can induce exon skipping in a non-HIV reporter, a result that was attributed to an increased transcription efficiency (81). This effect of Tat on splicing could also be indirect and mediated by proteins such as Tat-SF1 or SKIP, two components of the splicing machinery that can associate with pTEFb and stimulate Tat-dependent transcription elongation (82,83).

Another likely effector of Tat-mediated HIV-1 splicing inhibition is the cellular p32 protein. The first hint of its role came from the comparison of HIV-1 infection in human or murine cells. It is known for many years that murine cells do not support efficient HIV-1 multiplication, which has impeded scientists working on HIV from using the mouse as an animal model (84). This is due at least to specific amino acid changes between mouse and human in 2 proteins: cyclin T1 and p32, which respectively explain a lack of pTEFb/Tat/TAR-mediated activation of transcription and overly efficient HIV-1 RNA splicing in mouse cells (85,86). Therefore the p32 protein appears to be critical to explain the natural inefficiency of HIV-1 splicing in human cells, which is necessary to ensure the production of unspliced and partly spliced mRNA. Interestingly, while described a few years ago as an inhibitor of the SR protein ASF/SF2 (87), p32 is also known to have functions during transcription, especially in virus-infected cells. For example, p32 induces an hyperphosphorylation of the RNA Pol II CTD during adenovirus infection and blocks the expression of the adenovirus major late unit (88). Recently, Berro and colleagues showed that the acetylated form of HIV-1 Tat interacts efficiently and specifically with p32 protein (89). Interestingly, treatment of HIV-1 infected cells with the deacetylase inhibitor trichostatin A induced the recruitment of p32 to the LTR promoter and a shift of splicing from multiply to singly spliced isoforms. Tat acetylation is essential for HIV-1 multiplication since its effects on the Tat/pTEFb complex lead to the establishment and maintenance of an active transcription elongation complex (90,91). The strong processivity of this complex should not by itself favour splicing, especially at the suboptimal HIV-1 splicing sites (92). We can therefore anticipate that the recruitment of p32 by acetylated Tat helps to further inhibit splicing during the late phase of viral infection through a molecular mechanism which has yet to be determined, but which could involve a sequestration of ASF/SF2, one of the major activators of HIV-1 pre-mRNA splicing.

10. CONCLUSIONS: LESSONS FROM HIV-1 SPLICING FOR THE DESIGN OF FUTURE THERAPIES

Overall modulation of viral splicing during HIV-1 infection may be divided into 2 major phases. Early, during infection, the cellular splicing machinery is fully operational and the activity of SR proteins secures the production of *tat* and *rev* mRNAs, in some kind of "default" pathway. When a threshold concentration of Tat protein is reached in a later phase, the tremendous increase of RNA Pol II processivity might by itself alter the splicing pattern by preventing the utilization of the weakest splice sites. Altogether, the concerted action of Tat and cellular factors such as hnRNP A1 or p32, the action of Rev and a possible modulation of the activity of SR proteins, lead to a strong inhibition of HIV-1 splicing (especially the D4-A7 splicing) and to the synthesis of the proteins required for the assembly of new viruses. On the basis of this dual phase process, novel therapeutic strategies targeting splicing can be envisaged in two directions. The first approach would aim at decreasing splicing efficiency in the early phase to prevent Tat and Rev production. An alternative approach would be to mimic the situation that prevails in mouse cells, and improve the splicing of viral pre-mRNA during

the late phase of infection, preventing the production of the structural proteins.

So far, most efforts have been concentrated on targeting SR proteins, building on the work that has already been done in this field by several groups. For example, Fukuhara and colleagues have analysed the effect of a specific inhibitor of the SR protein kinases SRPK1 and SRPK2 (72). Unfortunately, even though it downregulated SRp75, that chemical compound had no significant effect on HIV-1 replication during a cell line infection assay. Another approach using an indole derivative that selectively inhibits the ESE-dependent splicing activity of ASF/SF2 has been developed recently (93). The results obtained with this kind of compound is more promising, since the drug IDC16 strongly decreased the production of multiply spliced HIV-1 transcripts, including *tat*, *rev* and *nef* mRNAs (52,93). Furthermore, the production of new viruses (measured by the quantification of the viral capsid protein p24) was reduced by at least 90% in macrophages or in activated peripheral blood mononuclear cells. This makes this drug a very promising candidate for anti-viral therapy, since it has a weaker effect than AZT on the expression of a panel of cellular genes (52). Finally, another strategy consists in targeting directly the HIV-1 RNA using U7 snRNA-coupled antisense sequences (94). These molecules, which are designed to mask splice sites and/or ESEs selectively, induce the skipping of specific exons and a strong decrease of the amounts of the targeted mRNA (in this case the *tat/rev* mRNAs). The consequence of this treatment was an inhibition, albeit incomplete, of HIV-1 production in human T lymphocytes. We can foresee that these different strategies will soon be developed further and that they will allow the emergence of new therapies to complement or improve the current medical treatment of AIDS.

11. ACKNOWLEDGEMENTS

JMS and CFB contributed equally to this review. We thank the « Centre National de la Recherche Scientifique » (CNRS), the « Institut National de la Sante et de la Recherche Medicale » (INSERM), the french « Ministere de l'Enseignement Superieur et de la Recherche », the « Agence Nationale de la Recherche sur le Sida et les hepatites virales » (ANRS 05195/06194), the « Agence Nationale pour la Recherche » (ANR-05-BLAN-0261-01) and the European Alternative Splicing Network (EURASNET, 6th Framework Program) for financial support. JMS was supported by a PhD studentship from ANRS. We are grateful to Chris Browning for his proofreading of the manuscript and to A. Cochrane for sending his review prior to publication in the same series.

12. REFERENCES

1. Johnson, J. M., J. Castle, P. Garrett-Engele, Z. Kan, P. M. Loerch, C. D. Armour, R. Santos, E. E. Schadt, R. Stoughton and D. D. Shoemaker: Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, 302, 2141-4 (2003)

2. Wang, Z. and C. B. Burge: Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA (*2008)

3. Matlin, A. J., F. Clark and C. W. Smith: Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol*, 6, 386-98 (2005)

4. Maniatis, T. and R. Reed: An extensive network of coupling among gene expression machines. *Nature*, 416, 499-506 (2002)

5. Buratti, E., A. F. Muro, M. Giombi, D. Gherbassi, A. Iaconcig and F. E. Baralle: RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. *Mol Cell Biol*, 24, 1387-400 (2004)

6. Karn, J.: Tackling Tat. *J Mol Biol*, 293, 235-54. (1999)

7.Peterlin, B. M. and D. H. Price: Controlling the elongation phase of transcription with P-TEFb. *Mol Cell*, 23, 297-305 (2006)

8. Purcell, D. F. and M. A. Martin: Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol.*, 67, 6365-78 (1993)

9. Benko, D. M., S. Schwartz, G. N. Pavlakis and B. K. Felber: A novel human immunodeficiency virus type 1 protein, tev, shares sequences with tat, env, and rev proteins. *J Virol*, 64, 2505-18 (1990)

10. Salemi, M., K. Strimmer, W. W. Hall, M. Duffy, E. Delaporte, S. Mboup, M. Peeters and A. M. Vandamme: Dating the common ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes using a new method to uncover clock-like molecular evolution. *FASEB J*, 15, 276-8 (2001)

11. Lutzelberger, M., L. S. Reinert, A. T. Das, B. Berkhout and J. Kjems: A novel splice donor site in the gag-pol gene is required for HIV-1 RNA stability. *J Biol Chem*, 281, 18644-51 (2006)

12. Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo and G. N. Pavlakis: Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J Virol*, 64, 2519-29 (1990)

13. Kestler, H. W., 3rd, D. J. Ringler, K. Mori, D. L. Panicali, P. K. Sehgal, M. D. Daniel and R. C. Desrosiers: Importance of the nef gene for maintenance of high virus loads and for development of AIDS. *Cell*, 65, 651-62 (1991)

14. Hope, T. J.: The ins and outs of HIV Rev. *Arch Biochem Biophys*, 365, 186-91 (1999)

15. Popov, S., M. Rexach, G. Zybarth, N. Reiling, M. A. Lee, L. Ratner, C. M. Lane, M. S. Moore, G. Blobel and M.

Bukrinsky: Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. *EMBO J*, 17, 909-17 (1998)

16. Ohagen, A. and D. Gabuzda: Role of Vif in stability of the human immunodeficiency virus type 1 core. *J Virol*, 74, 11055-66 (2000)

17. O'Reilly, M. M., M. T. McNally and K. L. Beemon: Two strong 5' splice sites and competing, suboptimal 3' splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology*, 213, 373-85 (1995)

18. Damier, L.: Etude de la régulation de l'épissage du transcrit primaire du virus de l'immunodéficience humaine de type 1. Université de Nancy 1 (1997).

19. Kammler, S., C. Leurs, M. Freund, J. Krummheuer, K. Seidel, T. O. Tange, M. K. Lund, J. Kjems, A. Scheid and H. Schaal: The sequence complementarity between HIV-1 5' splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. *RNA*, 7, 421-34 (2001)

20. Caputi, M., M. Freund, S. Kammler, C. Asang and H. Schaal: A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *J Virol*, 78, 6517-26 (2004)

21. Bohne, J., H. Wodrich and H. G. Krausslich: Splicing of human immunodeficiency virus RNA is position-dependent suggesting sequential removal of introns from the 5' end. *Nucleic Acids Res*, 33, 825-37 (2005)

22. Baudin, F., R. Marquet, C. Isel, J. L. Darlix, B. Ehresmann and C. Ehresmann: Functional sites in the 5' region of human immunodeficiency virus type 1 RNA form defined structural domains. *J Mol Biol*, 229, 382-97 (1993)

23. Abbink, T. E., M. Ooms, P. C. Haasnoot and B. Berkhout: The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. *Biochemistry*, 44, 9058-66 (2005)

24. Abbink, T. E. and B. Berkhout: RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor. *J Virol*, 82, 3090-8 (2008)

25. Madsen, J. M. and C. M. Stoltzfus: An exonic splicing silencer downstream of the 3' splice site A2 is required for efficient human immunodeficiency virus type 1 replication. *J Virol*, 79, 10478-86 (2005)

26. Madsen, J. M. and C. M. Stoltzfus: A suboptimal 5' splice site downstream of HIV-1 splice site A1 is required for unspliced viral mRNA accumulation and efficient virus replication. *Retrovirology*, 3, 10 (2006)

27. Ashe, M. P., L. H. Pearson and N. J. Proudfoot: The HIV-1 5' LTR poly (A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *EMBO J*, 16, 5752-63 (1997)

28. Ashe, M. P., A. Furger and N. J. Proudfoot: Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. *RNA*, 6, 170-7 (2000)

29. Damgaard, C. K., S. Kahns, S. Lykke-Andersen, A. L. Nielsen, T. H. Jensen and J. Kjems: A 5' splice site enhances the recruitment of basal transcription initiation factors *in vivo*. *Mol Cell*, 29, 271-8 (2008)

30. Staffa, A. and A. Cochrane: The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3' splice site. *J Virol*, 68, 3071-9 (1994)

31. Si, Z., B. A. Amendt and C. M. Stoltzfus: Splicing efficiency of human immunodeficiency virus type 1 tat RNA is determined by both a suboptimal 3' splice site and a 10 nucleotide exon splicing silencer element located within tat exon 2. *Nucleic Acids Res*, 25, 861-7 (1997)

32. Jacquenet, S., D. Ropers, P. S. Bilodeau, L. Damier, A. Mougin, C. M. Stoltzfus and C. Branlant: Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3' splice site and its cis-regulatory element: possible involvement in RNA splicing. *Nucleic Acids Res*, 29, 464-78 (2001)

33. Damier, L., L. Domenjoud and C. Branlant: The D1-A2 and D2-A2 pairs of splice sites from human immunodeficiency virus type 1 are highly efficient *in vitro*, in spite of an unusual branch site. *Biochem Biophys Res Commun*, 237, 182-7 (1997)

34. Swanson, A. K. and C. M. Stoltzfus: Overlapping cis sites used for splicing of HIV-1 env/nef and rev mRNAs. *J Biol Chem*, 273, 34551-7 (1998)

35. Tange, T. O. and J. Kjems: SF2/ASF binds to a splicing enhancer in the third HIV-1 tat exon and stimulates U2AF binding independently of the RS domain. *J Mol Biol*, 312, 649-62 (2001)

36. Dyhr-Mikkelsen, H. and J. Kjems: Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript *in vitro*. *J Biol Chem*, 270, 24060-6 (1995)

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

38. Marchand, V., A. Mereau, S. Jacquenet, D. Thomas, A. Mougin, R. Gattoni, J. Stevenin and C. Branlant: A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J Mol Biol*, 323, 629-52 (2002)

39. Damgaard, C. K., T. O. Tange and J. Kjems: hnRNP A1 controls HIV-1 mRNA splicing through cooperative

binding to intron and exon splicing silencers in the context of a conserved secondary structure. *RNA*, 8, 1401-15 (2002)

40. Caputi, M., A. Mayeda, A. R. Krainer and A. M. Zahler: hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J*, 18, 4060-7 (1999)

41. Bilodeau, P. S., J. K. Domsic, A. Mayeda, A. R. Krainer and C. M. Stoltzfus: RNA splicing at human immunodeficiency virus type 1 3' splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element. *J Virol*, 75, 8487-97 (2001)

42. Jacquenet, S., A. Mereau, P. S. Bilodeau, L. Damier, C. M. Stoltzfus and C. Branlant: 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*, 276, 40464-75 (2001)

43. Tange, T. O., C. K. Damgaard, S. Guth, J. Valcarcel and J. Kjems: The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J*, 20, 5748-58 (2001)

44. Zhu, J., A. Mayeda and A. R. Krainer: Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell*, 8, 1351-61 (2001)

45. Asai, K., C. Platt and A. Cochrane: Control of HIV-1 env RNA splicing and transport: investigating the role of hnRNP A1 in exon splicing silencer (ESS3a) function. *Virology*, 314, 229-42 (2003)

46. Zahler, A. M., C. K. Damgaard, J. Kjems and M. Caputi: SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing. *J Biol Chem*, 279, 10077-84 (2004)

47. Hallay, H., N. Locker, L. Ayadi, D. Ropers, E. Guittet and C. Branlant: Biochemical and NMR study on the competition between proteins SC35, SRp40 and hnRNP A1 at the HIV-1 Tat exon 2 splicing site. *J Biol Chem* (2006)

48. Ropers, D., L. Ayadi, R. Gattoni, S. Jacquenet, L. Damier, C. Branlant and J. Stevenin: 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*, 279, 29963-73 (2004)

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

50. Domsic, J. K., Y. Wang, A. Mayeda, A. R. Krainer and C. M. Stoltzfus: Human immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV antagonizes binding of U2AF65 to viral polypyrimidine tracts. *Mol Cell Biol*, 23, 8762-72 (2003)

51. Jacquenet, S., D. Decimo, D. Muriaux and J. L. Darlix: Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. *Retrovirology*, 2, 33 (2005)

52. Bakkour, N., Y. L. Lin, S. Maire, L. Ayadi, F. Mahuteau-Betzer, C. H. Nguyen, C. Mettling, P. Portales, D. Grierson, B. Chabot, P. Jeanteur, C. Branlant, P. Corbeau and J. Tazi: Small-molecule inhibition of HIV pre-mRNA splicing as a novel antiretroviral therapy to overcome drug resistance. *PLoS Pathog*, 3, 1530-9 (2007)

53. Bourgeois, C. F., F. Lejeune and J. Stevenin: Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog Nucleic Acid Res Mol Biol*, 78, 37-88 (2004)

54. Zheng, Z. M.: Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *J Biomed Sci*, 11, 278-94 (2004)

55. Tacke, R. and J. L. Manley: The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *EMBO J.*, 14, 3540-51 (1995)

56. Smith, P. J., C. Zhang, J. Wang, S. L. Chew, M. Q. Zhang and A. R. Krainer: An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet*, 15, 2490-508 (2006)

57. Amendt, B. A., Z. H. Si and C. M. Stoltzfus: 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*, 15, 4606-15 (1995)

58. Dreyfuss, G., M. J. Matunis, S. Pinol-Roma and C. G. Burd: hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.*, 62, 289-321 (1993)

59. Cartegni, L., M. Maconi, E. Morandi, F. Cobianchi, S. Riva and G. Biamonti: hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins. *J Mol Biol*, 259, 337-48 (1996)

60. Casas-Finet, J. R., J. D. Smith, Jr., A. Kumar, J. G. Kim, S. H. Wilson and R. L. Karpel: Mammalian heterogeneous ribonucleoprotein A1 and its constituent domains. Nucleic acid interaction, structural stability and self-association. *J Mol Biol*, 229, 873-89 (1993)

61. Sanford, J. R., J. Ellis and J. F. Caceres: Multiple roles of arginine/serine-rich splicing factors in RNA processing. *Biochem Soc Trans*, 33, 443-6 (2005)

62. Huang, Y. and J. A. Steitz: SRprises along a messenger's journey. *Mol Cell*, 17, 613-5 (2005)

63. Amendt, B. A., D. Hesslein, L. J. Chang and C. M. Stoltzfus: Presence of negative and positive cis-acting RNA

splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1. *Mol Cell Biol*, 14, 3960-70 (1994)

64. Mayeda, A., G. R. Screaton, S. D. Chandler, X. D. Fu and A. R. Krainer: Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. *Mol. Cell. Biol.*, 19, 1853-63 (1999)

65. Pongoski, J., K. Asai and A. Cochrane: Positive and negative modulation of human immunodeficiency virus type 1 Rev function by cis and trans regulators of viral RNA splicing. *J Virol*, 76, 5108-20 (2002)

66. Woolaway, K., K. Asai, A. Emili and A. Cochrane: hnRNP E1 and E2 have distinct roles in modulating HIV-1 gene expression. *Retrovirology*, 4, 28 (2007)

67. Exline, C. M., Z. Feng and C. M. Stoltzfus: Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. *J Virol*, 82, 3921-31 (2008)

68. Eperon, I. C., O. V. Makarova, A. Mayeda, S. H. Munroe, J. F. Caceres, D. G. Hayward and A. R. Krainer: Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Mol Cell Biol*, 20, 8303-18 (2000)

69. Dowling, D., S. Nasr-Esfahani, C. H. Tan, K. O'Brien, J. L. Howard, D. A. Jans, D. F. Purcell, C. M. Stoltzfus and S. Sonza: HIV-1 infection induces changes in expression of cellular splicing factors that regulate alternative viral splicing and virus production in macrophages. *Retrovirology*, 5, 18 (2008)

70. Maldarelli, F., C. Xiang, G. Chamoun and S. L. Zeichner: The expression of the essential nuclear splicing factor SC35 is altered by human immunodeficiency virus infection. *Virus Res*, 53, 39-51. (1998)

71. Ryo, A., Y. Suzuki, M. Arai, N. Kondoh, T. Wakatsuki, A. Hada, M. Shuda, K. Tanaka, C. Sato, M. Yamamoto and N. Yamamoto: Identification and characterization of differentially expressed mRNAs in HIV type 1-infected human T cells. *AIDS Res Hum Retroviruses*, 16, 995-1005 (2000)

72. Fukuhara, T., T. Hosoya, S. Shimizu, K. Sumi, T. Oshiro, Y. Yoshinaka, M. Suzuki, N. Yamamoto, L. A. Herzenberg and M. Hagiwara: Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. *Proc Natl Acad Sci U S A*, 103, 11329-33 (2006)

73. Kanopka, A., O. Muhlemann, S. Petersen-Mahrt, C. Estmer, C. Ohrmalm and G. Akusjarvi: Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature*, 393, 185-7 (1998)

74. Sciabica, K. S., Q. J. Dai and R. M. Sandri-Goldin: ICP27 interacts with SRPK1 to mediate HSV splicing inhibition by altering SR protein phosphorylation. *EMBO J*, 22, 1608-19 (2003)

75. Schaub, M. C., S. R. Lopez and M. Caputi: Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. *J Biol Chem*, 282, 13617-26 (2007)

76. Martinez-Contreras, R., J. F. Fisette, F. U. Nasim, R. Madden, M. Cordeau and B. Chabot: Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol*, 4, e21 (2006)

77. Caputi, M. and A. M. Zahler: SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. *EMBO J*, 21, 845-55 (2002)

78. Hashizume, C., M. Kuramitsu, X. Zhang, T. Kurosawa, M. Kamata and Y. Aida: Human immunodeficiency virus type 1 Vpr interacts with spliceosomal protein SAP145 to mediate cellular pre-mRNA splicing inhibition. *Microbes Infect*, 9, 490-7 (2007)

79. Kuramitsu, M., C. Hashizume, N. Yamamoto, A. Azuma, M. Kamata, Y. Tanaka and Y. Aida: A novel role for Vpr of human immunodeficiency virus type 1 as a regulator of the splicing of cellular pre-mRNA. *Microbes Infect*, 7, 1150-60 (2005)

80. Bohne, J., A. Schambach and D. Zychlinski: New way of regulating alternative splicing in retroviruses: the promoter makes a difference. *J Virol*, 81, 3652-6 (2007)

81. Nogues, G., S. Kadener, P. Cramer, D. Bentley and A. R. Kornblihtt: Transcriptional activators differ in their abilities to control alternative splicing. *J Biol Chem*, 277, 43110-4 (2002)

82. Fong, Y. W. and Q. Zhou: Stimulatory effect of splicing factors on transcriptional elongation. *Nature*, 414, 929-33 (2001)

83. Bres, V., N. Gomes, L. Pickle and K. A. Jones: A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. *Genes Dev*, 19, 1211-26 (2005)

84. Bieniasz, P. D. and B. R. Cullen: Multiple blocks to human immunodeficiency virus type 1 replication in rodent cells. *J Virol*, 74, 9868-77 (2000)

85. Garber, M. E., P. Wei, V. N. KewalRamani, T. P. Mayall, C. H. Herrmann, A. P. Rice, D. R. Littman and K. A. Jones: The interaction between HIV-1 Tat and human cyclin T1 requires zinc and a critical cysteine residue that is not conserved in the murine CycT1 protein. *Genes Dev*, 12, 3512-27 (1998)

86. Zheng, Y. H., H. F. Yu and B. M. Peterlin: Human p32 protein relieves a post-transcriptional block to HIV replication in murine cells. *Nat Cell Biol*, 5, 611-8 (2003)

87. Petersen-Mahrt, S. K., C. Estmer, C. Ohrmalm, D. A. Matthews, W. C. Russell and G. Akusjarvi: The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J*, 18, 1014-24 (1999)

88. Ohrmalm, C. and G. Akusjarvi: Cellular splicing and transcription regulatory protein p32 represses adenovirus major late transcription and causes hyperphosphorylation of RNA polymerase II. *J Virol*, 80, 5010-20 (2006)

89. Berro, R., K. Kehn, C. de la Fuente, A. Pumfery, R. Adair, J. Wade, A. M. Colberg-Poley, J. Hiscott and F. Kashanchi: Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. *J Virol*, 80, 3189-204 (2006)

90. Kaehlcke, K., A. Dorr, C. Hetzer-Egger, V. Kiermer, P. Henklein, M. Schnoelzer, E. Loret, P. A. Cole, E. Verdin and M. Ott: Acetylation of Tat defines a cyclinT1-independent step in HIV transactivation. *Mol Cell*, 12, 167-76 (2003)

91. Ott, M., A. Dorr, C. Hetzer-Egger, K. Kaehlcke, M. Schnolzer, P. Henklein, P. Cole, M. M. Zhou and E. Verdin: Tat acetylation: a regulatory switch between early and late phases in HIV transcription elongation. *Novartis Found Symp*, 259, 182-93; discussion 193-6, 223-5 (2004)

92. Nogues, G., S. Kadener, P. Cramer, M. de la Mata, J. P. Fededa, M. Blaustein, A. Srebrow and A. R. Kornblihtt: Control of alternative pre-mRNA splicing by RNA Pol II elongation: faster is not always better. *IUBMB Life*, 55, 235-41 (2003)

93. Soret, J., N. Bakkour, S. Maire, S. Durand, L. Zekri, M. Gabut, W. Fic, G. Divita, C. Rivalle, D. Dauzonne, C. H. Nguyen, P. Jeanteur and J. Tazi: Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. *Proc Natl Acad Sci U S A*, 102, 8764-9 (2005)

94. Asparuhova, M. B., G. Marti, S. Liu, F. Serhan, D. Trono and D. Schumperli: Inhibition of HIV-1 multiplication by a modified U7 snRNA inducing Tat and Rev exon skipping. *J Gene Med*, 9, 323-34 (2007)

95. Gottlinger, H. G., T. Dorfman, E. A. Cohen and W. A. Haseltine: The role of the tnv protein and tnv RNA splicing signals in replication of HIV-1 IIIB isolates. *Virology*, 189, 618-28 (1992)

96. Staffa, A. and A. Cochrane: Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol Cell Biol*, 15, 4597-605 (1995)

Abbreviations: HIV-1: human immunodeficiency virus type 1; SR protein: serine/arginine protein; hnRNP: heterogeneous ribonucleoprotein; LTR: long terminal repeat; TAR: *trans*-activation response; pTEF-b: positive transcription elongation factor b; CDK9: cyclin dependant kinase 9; CTD: C-terminal domain; RRE; Rev responsive element; ESE: exonic splicing enhancer; ESS: exonic splicing silencer; ISS: intronic splicing silencer; PPT: polypyrimidine tract; SLS: stem-loop sequence; NMR: nuclear magnetic resonance; RRM: RNA recognition motif; RNA Pol II: RNA polymerase II.

Key words: HIV-1, alternative splicing, RNA structure, SR proteins, hnRNP proteins, splicing enhancer, splicing silencer

Send correspondence to: Christiane Branlant, MAEM UMR 7567, Universite Henri Poincare - Nancy I, BP 239, 54506 Vandoeuvre-les-Nancy, France, Tel: 33-383684303, Fax: 33-383684307, E-mail: Christiane.Branlant@maem.uhp-nancy.fr

http://www.bioscience.org/current/vol14.htm