Published online 23 October 2009 *Nucleic Acids Research, 2009, Vol. 37, No. 22* ***7429–7440***
doi:10.1093/nar/gkp694

Limited complementarity between U1 snRNA and a retroviral 5′ splice site permits its attenuation via RNA secondary structure

Daniela Zychlinski[1], Steffen Erkelenz[2], Vanessa Melhorn[1], Christopher Baum[3], Heiner Schaal[2] and Jens Bohne[1,*]

[1]Institute for Virology, Hannover Medical School, 30625 Hannover, [2]Institute for Virology, Heinrich-Heine-University, 40225 Düsseldorf and [3]Department of Experimental Hematology, Hannover Medical School, 30625 Hannover, Germany

Received April 1, 2009; Revised June 26, 2009; Accepted August 6, 2009

ABSTRACT

Multiple types of regulation are used by cells and viruses to control alternative splicing. In murine leukemia virus, accessibility of the 5′ splice site (ss) is regulated by an upstream region, which can fold into a complex RNA stem–loop structure. The underlying sequence of the structure itself is negligible, since most of it could be functionally replaced by a simple heterologous RNA stem–loop preserving the wild-type splicing pattern. Increasing the RNA duplex formation between U1 snRNA and the 5′ss by a compensatory mutation in position +6 led to enhanced splicing. Interestingly, this mutation affects splicing only in the context of the secondary structure, arguing for a dynamic interplay between structure and primary 5′ss sequence. The reduced 5′ss accessibility could also be counteracted by recruiting a splicing enhancer domain via a modified MS2 phage coat protein to a single binding site at the tip of the simple RNA stem–loop. The mechanism of 5′ss attenuation was revealed using hyperstable U1 snRNA mutants, showing that restricted U1 snRNP access is the cause of retroviral alternative splicing.

INTRODUCTION

Alternative splicing extensively expands the human transcriptome (1) and proteome, giving rise to roughly 100 000 protein isoforms from only 25 000 genes (2). In retroviruses a single pre-mRNA corresponding to the complete genome also undergoes alternative splicing to express all viral genes (3).

The splicing reaction is executed by the spliceosome (4). The core of the spliceosome consists of five small nuclear RNPs (U snRNPs; 5), some of which participate in splice site recognition via RNA:RNA interactions (6,7). The first step towards mRNA splicing is the recognition of the 5′ss by the free, complementary 5′ end of U1 snRNA (8,9). Therefore, the hydrogen bonding pattern between all 11 nt of the 5′ss and U1 snRNA determines the intrinsic strength of the 5′ss and thus contributes to its recognition and frequency of usage, creating a first layer of regulation (10,11).

In contrast to yeast, where almost all splice sites match the consensus sequence (12), splice sites in retroviral and mammalian genomes are much more degenerated and recognition is frequently assisted by a number of splicing regulatory proteins (13). Accordingly, regions in proximity to splice sites often represent exonic or intronic splicing enhancers (ESE, ISE; 14), or silencers (ESS, ISS; 15,16). These elements modulate the intrinsic strength of splice sites mostly via recruitment of splicing factors like SR proteins or hnRNPs (17). The efficiency of splice sites can also be modulated by RNA structure (18,19). Either folding of the structure competes directly with formation of the U1 snRNA:5′ss RNA duplex or indirectly by masking binding sites of splicing regulatory proteins (20). Finally, transcriptional elongation can also regulate alternative splicing, illustrating the close connection between splicing and transcription (21).

Retroviruses represent very valuable model systems for studying alternative splicing (22). While they synthesize only one polycistronic primary transcript, which undergoes alternative splicing for full viral gene expression (Figure 1A; 23), retroviruses also need to tightly control the use of their splice sites to ensure optimal levels of unspliced versus spliced RNAs (3). The unspliced or genomic RNA is packaged into progeny virus

*To whom correspondence should be addressed. Email: bohne.jens@mh-hannover.de

The authors wish it to be known that, in their opinion, the second and third authors should be regarded as joint Second Authors.

and also serves as a translation template for the structural and enzymatic proteins Gag and Pol, whereas the spliced RNA encodes the envelope protein (Env; Figure 1A). The correct ratio of Gag and Env partly determines viral infectivity. HIV splicing is highly regulated by ESEs or ESSs [(24–26); reviewed in (27)] and by weak polypyrimidine tracts (PPTs), which are interrupted by weakening purines (28). In one case reported, a PPT is additionally attenuated by a secondary structure (29). For simple retroviruses such as Rous sarcoma virus (RSV) or murine leukemia virus (MLV), alternative splicing has also been attributed to weak 3′ss (30–32). In addition, RSV harbors a decoy 5′ss, which redirects splicing activity from the authentic 5′ss to a nonproductive one (33).

We could previously show that in MLV the 5′ss instead of the 3′ss is negatively regulated via upstream sequences, which can form a secondary structure. Moreover, the stability and integrity of this structure correlates with 5′ss attenuation (34). We now demonstrate that the restriction exerted by this upstream RNA secondary structure depends on limited complementarity between U1 snRNA and the 5′ss at position +6. We show that the RNA secondary structure-mediated U1 snRNA restriction could be counterbalanced by either increasing complementarity to U1 snRNA or SR protein-mediated stabilization of the 5′ss:U1 snRNA duplex. The latter was accomplished by an improved MS2-tethering system, which exerts a high affinity to a single stem–loop binding site. A heterologous RNA stem loop of comparable thermodynamic stability could replace the wild-type structure and preserve the splicing pattern. Interestingly, low complementarity to U1 snRNA affects splicing only in the context of the secondary structure, arguing for a dynamic interplay between structure and intrinsic strength of the 5′ss.

MATERIALS AND METHODS

Plasmids

All retroviral vector plasmids were derived from pSF91 (35). The mutants sm3 (stem–loop mutant 3), compensatory mutants thereof (sm3comp) and the deletion of the primer binding site (dPBS) were described previously (34). All overlap PCRs used the outer primers XbaI and ApaI (34) to generate the final PCR fragment, which was cloned into the XbaI/ApaI sites of SF91. Vectors carrying the C6U 5′ss mutation were cloned by overlapping PCRs using SF91, SF91sm3, SF91dPBS and SF91stem–loop as a template (fw: 5′-TAA GTT GGC CAG CGG TCG TTT CG-3′; rv: 5′-GGC CAA CTT ACC TCC CGG C-3′). For insertion of the heterologous stem–loop (SF91stem–loop) into the leader region of SF91, the KpnI/MscI fragment was replaced by a PCR product including the heterologous stem–loop sequences (fw: 5′-GTA CGG TAC CGT ATT CCC AAT AAA GCC TCT TGC TGT TTG CAT CCG AAT CGT GGA GGT CAA GAA TTC GCG GAC ACC ATC-3′; rv: 5′-GCA TCC TGG CCA GCT TAC CTC CCG GCG GAG GTC AAG AAT TCG CGG ACC CTG ATG GTG TCC GCG AAT TCT TGA CCT CCA CGA TTC-3′). The extended stem–loop was generated by introducing 20 additional base pairs into the EcoRI sites of SF91stem–loop (fw: 5′-[Phos] AAT TCG ATA TCC CGT GCG GAC ACC ATC AGG GTC CGC ACG GGA TAT CG-3′; rv: 5′-[Phos] CGA TAT CCC GTG CGG ACC CTG ATG GTG TCC GCA CGG GAT ATC GAA TT-3′) by ligation. By using the SF91stem–loop as a template, the SF91stem–loop antisense mutant (fw: 5′-CGC CTC CAG TTC TTA AGC GCC TCG GGA GGT AAG CTG GCC AGC GGT CGT TTC G-3′; rv: 5′-AGG CGC TTA AGA ACT GGA GGC GCC CTG ATG GTG TCC GCG AAT TCT TGA C-3′) was generated by overlapping PCRs and the final PCR product was cloned into the XbaI/ApaI sites of SF91. The SCSenv HIV/MLV hybrid constructs were also cloned by overlapping PCRs. The 5′PCR product was generated by using SCS11 (36) as a template (fw: 5′-GCG GTA TAC GCT AGC TTA AGT AG-3′; rv: 5′-TAC TTA CTG CCC GGC GGG GGG GTC GGT G-3′). To generate the 3′PCR product containing the HIV 5′ss, NLenv was used as a template (fw: 5′-CCC CCC CGC CGG GCA GTA AGT AGT ACA TGT AAT GC-3′; rv: 5′-GGT TGC TTC CTT CCA CAC AGG TAC-3′). Using the outer primers, the final PCR product was cloned into the Bst1107I/BstEII sites of the NLCenv backbone (37). In order to transfer the sm3 and the C6U mutation into an MLV provirus, an AflII/PstI fragment of the respective SF91 plasmids was generated and ligated into MOVSFGFP (34,38).

U1 snRNA mutants were cloned by PCR using a reverse primer (rv: 5′-CGC GGA TCC TCC ACT GTA GGA TTA AC-3′) including a BamHI restriction site and different forward primers containing the U1 mutations flanked by a BglII restriction site (fw U1G11C: 5′-GCC CGA AGA TCT CAT ACT TAC CTC GCA G-3′; fw U1 perfect: 5′-GCC CGA AGA TCT CCA GCT TAC CTC GCA G-3′). The PCR products were cloned as a BamHI/BglII fragment into the pUC19 U1wt plasmid (kind gift from A. Weiner, Seattle, WA, USA).

For construction of plasmid SV SD4/SA7 NLS-MS2 SRp55pA, the BamHI/XhoI fragment of SV NLS-MS2 9G8 (39) was replaced by a PCR product generated with primers (fw: 5′-GGT GGA TCC CGC ACA AGC CAT AGG CGA TC-3′; rv: 5′-AGA CTC GAG TTA ATC TCT GGA ACT CGA CCT GG-3′) and CMV myc SRp55 (kindly provided by A. Cochrane, Toronto, Canada) as a template. Plasmid SV NLS-MS2-ΔSR has been described previously (39). SV SD4/SA7 NLS-MS2ΔFG-ΔSR and SVSD4/SA7 NLS-MS2 ΔFG-SRp55 were generated by substitution of the XmaI/EcoRI fragment by a PCR-amplified fragment with primers (fw: 5′-GAC CCC GGG ATG GGG CCG CAA AAA ACG CCG C-3′ and rv: 5′-TCG GAA TTC GTA GCG AAA ATT GGA ATG GTT AGT TCC ATA TTT AAG TAC GAA CGC CAG GCG CCT-3′), containing a deletion of the sequence encoding the flexible α-helical loop connecting the two β-strands F and G (FG loop) within the MS2 coat protein, and SV SD4/SA7 NLS-MS2-ΔSR as a template.

Cells, transfections and virus titer

293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, 1 mM sodium pyruvate and 1% penicillin/streptomycin. The day before transfection, 4×10^6 293T cells were seeded in a 10-cm plate. Transfections were performed using the calcium phosphate precipitation method with 5 µg of the SF91 plasmid (10 µg for proviral constructs). Medium was changed 8 h post-transfection, and the cells were harvested 48 h later. Transfection efficiency was measured by FACS analysis and ranged between 40% and 60% (FACSCalibur; Becton-Dickinson, Heidelberg, Germany). The co-transfection assays (MS2 and U1) were performed using 5 µg SF91 plasmid and 10 µg U1 plasmid or 5 µg MS2-plasmid. HeLaP4 cells (40) were cultured under conditions identical to those for 293T cells. The day before transfection, 6×10^5 HeLaP4 cells were seeded in a 6-cm plate. Cells were transfected using the ICAFectin^{TM} 441 DNA transfection reagent (Eurogentec, Brussels, Belgium) and 2.75 µg plasmid DNA. RNA was harvested 36 h later.

For virus production, 293T cells were transfected with 10 µg of proviral plasmids. Supernatants were collected 36 h and 48 h after transfection and passed through a 0.22 µm filter (Millipore, Schwalbach, Germany). Initial titers were determined by transduction of 1×10^5 SC-1 cells (murine fibroblasts) using serial dilutions in the presence of 4 µg/ml protamine sulfate. Cells were centrifuged at 950 g and 25°C for 60 min and incubated for not more than 24 h to avoid re-infection. Cells were harvested and GFP-positive cells were counted by flow cytometry. GFP is encoded in frame with the MLV env ORF (38). SC-1 cells were then infected with a multiplicity of infection of 0.3. The spreading infection was monitored using GFP fluorescence and supernatants were collected at 90% GFP-positive cells. The supernatants were re-titrated on SC-1 cells to determine the viral titers after replication in murine cells.

RNA preparation and analysis

Preparation of total RNA, gel electrophoresis, blotting and detection with a radiolabeled probe were performed as described previously (41). The eGFP-specific probe corresponds to the eGFP cDNA and was generated by digestion of the SF91 plasmid with AgeI and EcoRI. To detect 18S rRNA, a genomic fragment was PCR-amplified and subcloned into pCR2.1 (Invitrogen, Karlsruhe, Germany). The probes were radiolabeled using the DecaLabel Kit (Fermentas, St. Leon-Rot, Germany). RNA was quantified photometrically and 10 µg were used for northern blot analysis, if not stated otherwise.

Phosphoimager analysis

The different RNA species were quantified by phosphoimager analysis (Storm 820; GE Healthcare, Chalfont St. Giles, UK) using a short and a longer exposure. Only experiments where the fold enhancement or splicing ratios remained constant over both exposure times were processed. The percentage of unspliced RNA was calculated using the following formula: [unspliced RNA/(spliced RNA + unspliced RNA)] × 100 = % unspliced RNA.

Reverse transcription and quantitative real-time PCR analysis

For reverse transcription, 5 µg total RNA were digested using the Ambion TURBO^{TM} DNase (Austin, TX, USA). 500–800 ng RNA were reverse transcribed using the QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany). Quantitative PCR was performed on a Stratagene Mx 3000P (La Jolla, CA, USA) using the QuantiFast SYBR Green (Qiagen) PCR Kit. To detect the unspliced RNA, fw primer 5′-GAG GGT CTC CTC AGA TTG ATT GAC-3′ and rv primer 5′-GAC AGA CAC GAA ACG ACC GC-3′ were used. To detect the spliced RNA, the fw primer was combined with an exon junction primer: 5′-TGT AAG TGA GCT CCC GGC-3′. As a standard we used a U1 snRNA primer set (fw: 5′-CTT ACC TGG CAG GGG AGA TAC-3′ and rv: 5′-GAA AGC GCG AAC GCA GTC-3′).

Electromobility shift assay

Using SF91stem–loop and stem–loop antisense templates (Figure 3A), PCR products were generated carrying a T7 promoter. The products were ligated into pCR2.1. T7 transcripts correspond exactly to the sequence depicted in Figure 3A. The resulting plasmids were linearized and approximately 500 ng were used for *in vitro* transcription using Ambion's Maxiscript Kit in the presence of 50 µCi of [^{32}P]UTP (400 Ci/mmol, Hartmann Analytic, Braunschweig, Germany) and unlabeled UTP to a final concentration of 40 µM. Free nucleotides were removed by MoBiTec S300 columns and probes were purified using denaturing PAGE. The radiolabeled RNA probe (25 000 c.p.m.) was incubated in 25 mM Tris–HCl, pH 7.9, 5 mM $MgCl_2$, 10% (v/v) glycerol, 0.4 mM dithiothreitol, 0.5 mM EDTA, 10 U RNAsin (Promega), 4 µg BSA and 1 µg of yeast tRNA in a total volume of 15 µl for 30 min at 20°C to allow folding of the RNA structure followed by denaturing for 2 min at 90°C. Purified U1 snRNP was purchased from Phadia (Freiburg, Germany). The purification procedure was performed using the original protocol from the Lührmann laboratory (42). U1 snRNP was added 10 min prior to loading on a 6% (60:1) polyacrylamide gel run in Tris–borate–EDTA buffer. The presence of U1 snRNA in the U1 snRNP fraction was assessed using RT-PCR along with a U1 plasmid as a control (Supplementary Figure 2).

Western blotting

Cells were resuspended in lysis buffer (final concentration, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 100 mM NaF, 10% glycerol, 10 mM EDTA, 50 mM Tris pH 8, 1 mM PMSF, Protease Inhibitor Cocktail Tablet; Roche, Mannheim, Germany) and incubated on ice for 20 min. The lysates were cleared by centrifugation at 15 000*g* for 10 min. Twenty micrograms protein were run on SDS–PAGE (10%), and transferred

to nitrocellulose membranes (Whatman, Dassel, Germany). MLV capsid was detected by a polyclonal rabbit anti-MLV p30 serum (kindly provided by C. Stocking, Hamburg, Germany) diluted 1:5000 in PBS-Tween, 0.1% and 5% dry milk. The membranes were re-probed using a mouse anti-actin antibody (Millipore, Schwalbach, Germany). Detection was carried out by chemi-luminescence (West Pico, Pierce, Bonn, Germany).

Computer software tools

RNA folding analysis of the leader region and of mutants thereof was performed using the mFOLD software (43; web access: http://mfold.bioinfo.rpi.edu). RNA structures were assembled using the XRNA software (http://rna.ucsc.edu/rnacenter/xrna/xrna.html). The intrinsic 5′ss strength was determined using the HBond-score algorithm (http://www.uni-duesseldorf.de/rna/html/hbond_score.php). Statistical analysis was performed using Graph Pad Prism 4 software.

RESULTS

Increasing 5′ splice site strength partially relieves the restriction exerted by the secondary structure

In order to reveal the molecular requirements for the previously described RNA secondary structure-mediated restriction of the murine leukemia virus (MLV) 5′ss, we used our previously described splicing reporter SF91 (34). This retroviral vector is derived from an MLV provirus and contains the packaging signal (ψ) embedded in an intron flanked by the authentic MLV 5′ and 3′ splice sites followed by eGFP as a marker gene (Figure 1A; 35). All reporter constructs were transiently transfected into 293T cells and total RNA was analyzed by northern blotting and quantified using a Phosphoimager. In addition, we independently assessed the splicing ratios by quantitative RT-PCR using the same RNAs.

We now considered whether the intrinsic strength of the 5′ss or its complementarity to U1 snRNA might contribute to the splicing regulation exerted by the secondary structure. Therefore, we converted position +6 of the MLV 5′ss from C to U (Figure 1B), thereby increasing the complementarity from an HBond score of 17.1 to 19.1. The C6U mutant displayed a 2.5-fold enhancement of splicing in the context of the secondary structure (Figure 1C and D; lanes 1 and 2). In comparison, mutations which exclusively prevent stem formation led to an 8-fold enhancement of splicing (sm3, Figure 1B; Figure 1C and D; compare lanes 2 and 3). These results confirm that, in a less structured region, fewer complementary nucleotides to U1 snRNA are sufficient to constitute an efficient 5′ss. A combination of the two mutations led to an almost complete lack of unspliced RNA (Figure 1C, lane 4). As shown in Figure 1D, the phosphoimager data are highly comparable to the qRT-PCR, although the latter tend to yield slightly higher values for the unspliced RNA in general (Figure 1D, black bars). Therefore, we continued to use northern blotting in the subsequent experiments to be able to detect cryptic splicing events as a possible result of the introduced mutations.

To investigate whether increasing the 5′ss:U1 snRNA complementarity could also relieve the restriction of a thermodynamically even more stable RNA secondary structure, we used a deletion of the primer binding site (Figure 1B), which folds into a more stable stem–loop ($\Delta G = -68.5$ kcal/mol in comparison to –61.4 kcal/mol for SF91 wt) and shows a stronger attenuation of the 5′ss (dPBS; Figure 1C, lane 5). Also in this context, the C6U mutation enhanced splicing (Figure 1C, lane 6; and Figure 1D). Both the C6U and dPBS mutations seem to influence the transcript levels of the SF91 RNA (Figures 1C and 3).

In summary, regulation at the 5′ss occurs on two layers: the primary splice site sequence and an upstream secondary structure.

The regulation of splicing is conserved in a replication-competent provirus

In order to confirm that splicing in the provirus is also regulated on two layers, secondary structure and primary 5′ss sequence, we cloned two key splicing mutations, namely sm3 (structural mutant; Figure 1C, lane 3) and C6U (splice site mutant; Figure 1C, lane 2), into an infectious MLV provirus (Figure 1A). Both types of mutations yielded an enhancement of splicing identical to that seen in the splicing reporter (Figure 2A). Even though the provirus splices less efficiently than the reporter, the magnitude of splicing enhancement is maintained (Figure 2C). Oversplicing of the genomic RNA should lead to reduced levels of full-length RNA to be packaged into viral progeny and to lower amounts of structural proteins since they are translated only from the unspliced RNA. Western blot analysis of cell lysates harvested from 293T cells revealed a correlation between the splicing ratio and the amount of capsid (p30) protein (Figure 2B). We also tested whether these mutations hamper viral replication. Indeed, replication of the sm3 provirus in murine fibroblast cells (SC-1) led to an almost 4-fold decrease in viral titer compared to wild-type MLV (Figure 2D). The C6U mutant showed only a modest titer reduction in agreement with the viral RNA and protein analysis. Thus, there is a correlation between the extent of splicing and viral fitness.

To sum up, the regulation of alternative splicing is conserved in the context of the complete provirus and mutations that severely affect splicing ratios impede viral replication.

The attenuating effect of the secondary structure is transferable to a heterologous 5′ splice site

Next we asked: did the inhibitory secondary structure evolve and function only in murine leukemia virus? In addition, we wanted to rule out the possibility that sequences far downstream of the 5′ss are involved in its attenuation. For this purpose, we used the previously described HIV-NLenv system (Figure 3A; 44). Briefly, this subviral envelope expression system is based on the HIV-1 proviral clone NL4-3. Using the Tat-independent

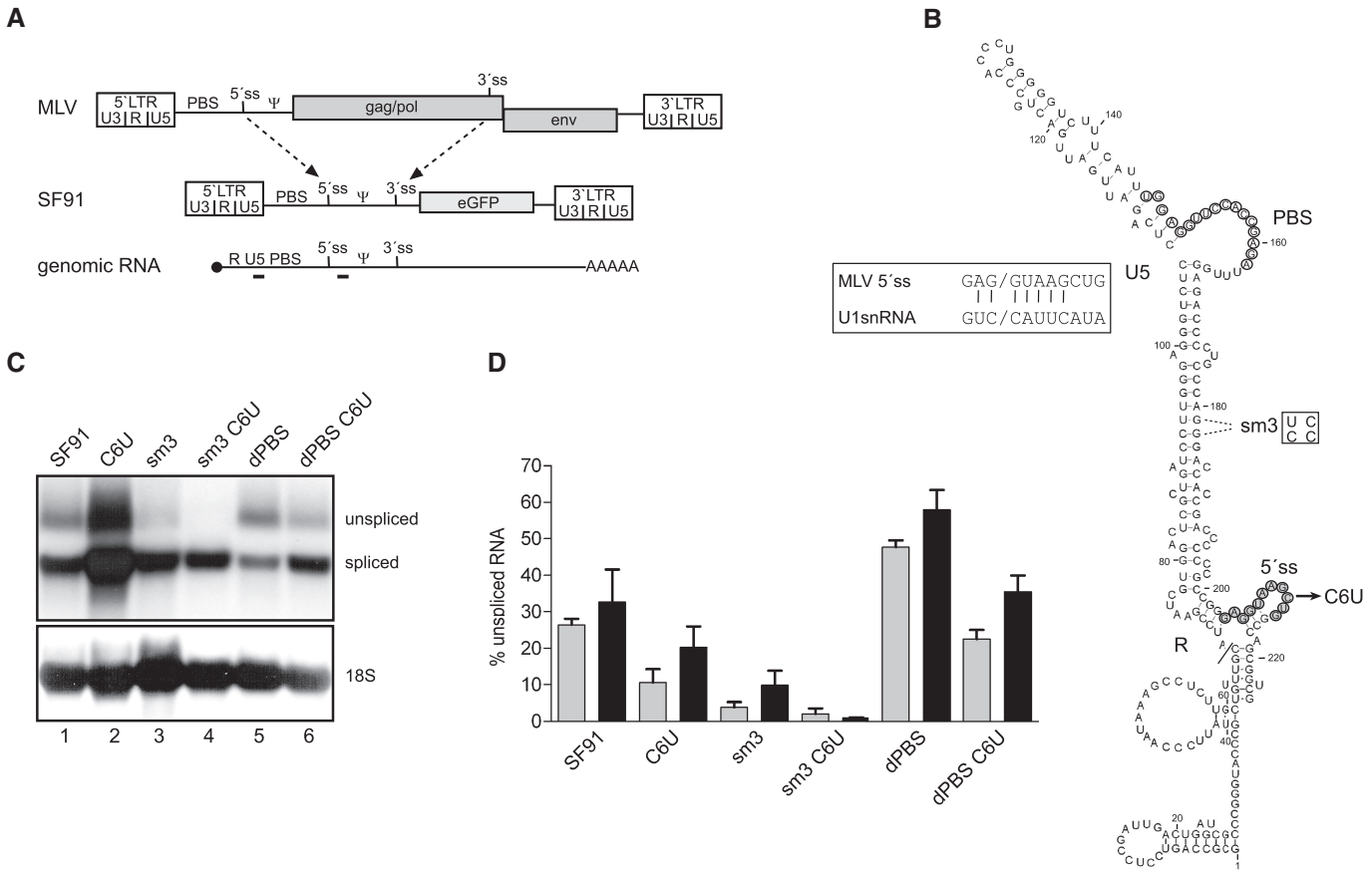

Figure 1. Dual attenuation of a 5′ss via the structure and a mismatch at position +6. (**A**) Schematic drawing of the murine leukemia provirus. The retrovirus is characterized by the long terminal repeats (LTRs, open boxes), the presence of the genes *gag*, *pol* and *env* (grey boxes) and the primer binding site (PBS). The packaging signal (ψ) and the *gag*/*pol* reading frame are flanked by 5′ and 3′ss. Below this, the MLV-derived splicing reporter SF91 is depicted. The proviral 3′ss was cloned downstream of the packaging signal and upstream of the eGFP start codon (grey box). Therefore, SF91 encodes a genomic RNA (black line), which contains an intron constituted by the authentic MLV splice sites. (**B**) RNA secondary structure of the region upstream of the packaging signal. The structure is adapted from the chemically validated structure (77). The regions R and U5 are part of the LTR. The nucleotides representing the primer binding site and the 5′ss are highlighted by open and grey circles, respectively. At the 5′ss, the +6C to U mutation is marked on the right. In the central region of the stem, the structural mutant sm3 including the nucleotide exchanges is depicted. The inset shows the MLV 5′ss pairing to U1 snRNA. Base pairs are indicated by vertical lines. (**C**) Northern blot using 10 µg total RNA from 293T cells transiently transfected with 5 µg of the indicated constructs and harvested 48 h post-transfection. The blot was hybridized with an eGFP-specific probe. The identity of the RNA species is stated on the right. As a loading control, the blot was re-hybridized with a probe corresponding to 18S rRNA. (**D**) Phosphoimager analysis of northern blot as shown in (C). The extent of alternative splicing is given as a percentage of unspliced RNA (grey bars). The black bars represent quantitative RT-PCR analysis of the same RNA as in C. The primers to detect the unspliced RNA are shown as thick black lines in (A). For detection of the spliced RNA, the same forward primer was used in combination with an exon junction primer (primer position not shown). The mean values and standard deviations of three independent experiments are given.

CMV promoter instead of the viral LTR (NLCenv), the construct expresses a sequence-identical HIV *env* mRNA, which can be alternatively spliced to yield *nef* mRNA (Figure 3A; 37). RNA analysis showed that NLCenv displays alternative splicing due to the weak 3′ss upstream of *nef* (Figure 3C, lane 2; 20,44).

In order to transfer the splicing regulation from MLV to a heterologous 5′ss, we inserted the 5′ leader sequence of MLV (bases 1–203; Figure 1B) immediately upstream of the HIV 5′ss, thereby creating SCSenv (light grey box, Figure 3A). SCSenv expresses an MLV/HIV fusion transcript containing the MLV-derived secondary structure (Figure 1B) followed by the HIV 5′ss and downstream *env* mRNA sequences. The proposed secondary structure of the fusion transcript is depicted in Figure 3B. The MLV RNA stem–loop may also force the HIV 5′ss into an inhibitory conformation. In addition, we transferred two structural mutations (sm3 and dPBS; Figure 1C) to SCSenv. For this experiment, we used HeLaP4 cells, in which the NLenv system was established (44), instead of 293T cells. Transfection into HeLaP4 cells revealed that the MLV-derived sequences can also strongly attenuate the HIV 5′ss (Figure 3C, compare lanes 2 and 3). The level of unspliced RNA was enhanced 2.5-fold (Figure 3D) and exceeded the level of 5′ss attenuation observed for the basic construct SF91 containing the identical leader sequence (Figure 1C and D). Comparison of the two 5′ss using the HBond

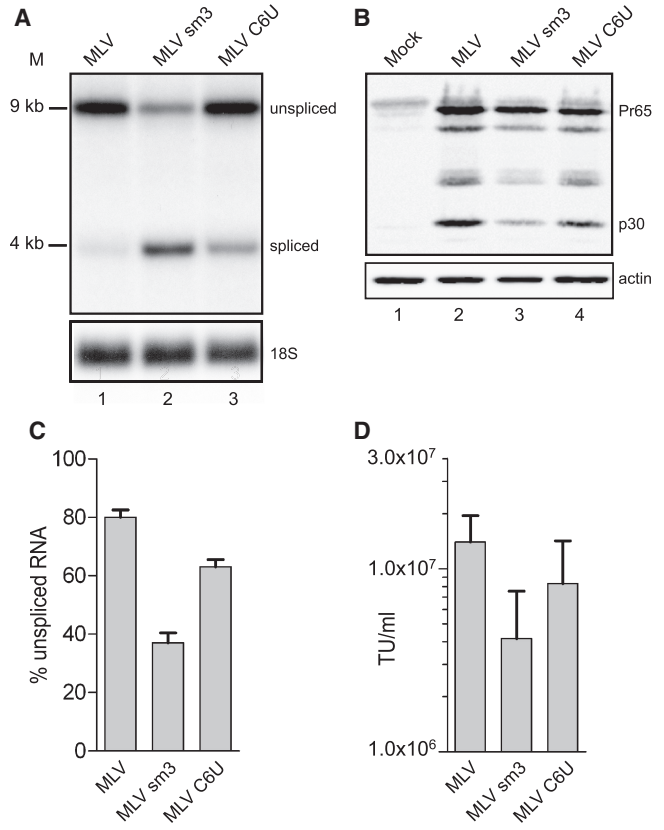

Figure 2. Splicing regulation is transferable to a complete provirus. (**A**) Northern blot of 10 µg of total RNA from 293T cells transfected with 10 µg of the proviral constructs (Figure 1A) harboring two splicing mutants (sm3 and C6U; Figure 1B and C). Size standard is given on the left and the RNA species are named on the right. (**B**) Western blot analysis of the same transfection as in (A). Twenty micrograms of protein were separated on a 10% SDS–PAGE and MLV-capsid protein was detected using a polyclonal anti-capsid serum. The polyprotein (Pr65) and the mature capsid protein (p30) are labeled on the right. Re-probing with actin-specific antibody served as loading control. (**C**) Phosphoimager analysis. The extent of splicing is given as the percentage of unspliced RNA. The mean values and standard deviations of three independent experiments are shown. (**D**) Supernatants from transient transfections were adjusted in titer and used to infect murine SC-1 cells. At the peak of infection, supernatants were collected and re-titrated on SC-1 cells. The titer is expressed as transducing units/ml in log scale. Three independent infections were performed for each construct.

score algorithm (45) revealed that the HIV 5′ss is intrinsically weaker than MLV (HIV HBond score 15.7; MLV HBond score 17.1). This is in agreement with our hypothesis that a 5′ss with lower complementarity should be even more prone to attenuation via the RNA secondary structure. Moreover, the mode of splicing regulation could be transferred onto a heterologous 5′ss since the sm3 mutant is spliced more efficiently and a deletion of the PBS leads to more unspliced RNA (Figure 3D, compare columns 3 and 4). Furthermore, no downstream MLV-derived sequences were involved in 5′ss attenuation. We noted that the sm3 mutant did not splice to the same extent as NLCenv, although the 5′ss should be accessible to the spliceosome (Figure 3C and D). However, this can be explained by a purine-rich splicing enhancer present only in the HIV leader sequence upstream of the 5′ss (39,46). In addition, the degree of splice site attenuation influenced the total amount of RNA (Figure 3C). The dPBS mutant in particular displayed the highest level of unspliced RNA, but the lowest amount of RNA (Figure 3C, lanes 5 and 6). In order to visualize the ratio of unspliced vs. spliced RNA, this part of the northern blot had to be overexposed (Figure 3C, lane 6).

The experiments using the HIV *env* expression system demonstrate that the splicing regulation observed is not restricted to MLV, but can be transferred to a heterologous 5′ss.

Replacement of the RNA structure with random sequences capable of stem formation results in proper splicing regulation

In order to differentiate whether mere structural requirements of the leader region or specific sequences harboring splicing regulatory protein binding sites cause 5′ss attenuation, we replaced the upper part of the RNA structure with a heterologous stem–loop harboring the ability to form a stem with free energy similar to that of the wild type (Figures 1B and 4A; SF91: $\Delta G = -49.6$ kcal/mol; heterologous stem–loop: $\Delta G = -47.2$ kcal/mol). RNA analysis of transient transfection of the stem–loop construct revealed a splicing pattern almost identical to that of the wild-type reporter (Figure 4B, lanes 1 and 2). Thus, 5′ss attenuation can be attributed to RNA stem–loop formation and not to the primary sequence, which forms the structure. As a control, we reversed the descending part of the stem, resulting in a stem–loop antisense (as) construct, which is unable to form the secondary structure upstream of the 5′ss. This construct displayed complete splicing (Figure 4B, lane 3). In contrast, strengthening the stem by a 20-bp extension on either side ($\Delta G = -68.8$ kcal/mol) led to more unspliced RNA (Figure 4B and C; lane 4), reminiscent of the deletion of the PBS (Figure 1C, lane 5). Furthermore, the C6U mutation enhances splicing as in the wild-type context, suggesting that this interplay does not require specific cellular proteins binding to the MLV stem loop (Figure 4B, lane 5). In addition, we noted that, similar to the SCSenv plasmid (Figure 3C), the splicing efficiency correlated with the overall transcript levels (Figure 4B, compare lanes 3 and 4). Therefore, the lane containing the stem–loop antisense (as) was underloaded for proper visualization (Figure 4B, lane 3). The antisense mutant showed that the 5′ss could function highly efficiently in the absence of the stem–loop despite the mismatch at position +6. This position became only critical in conjunction with the secondary structure, pointing to a novel dynamic interplay between structure and primary splice site sequence.

SR protein domains targeted to the heterologous stem–loop partially overcome 5′ss attenuation

It has been shown in yeast that 5′ and 3′ss mutations can be rescued by SR proteins even though *Saccharomyces cerevisiae* does not code for such proteins (47). This hints at a mechanism where SR proteins might stabilize

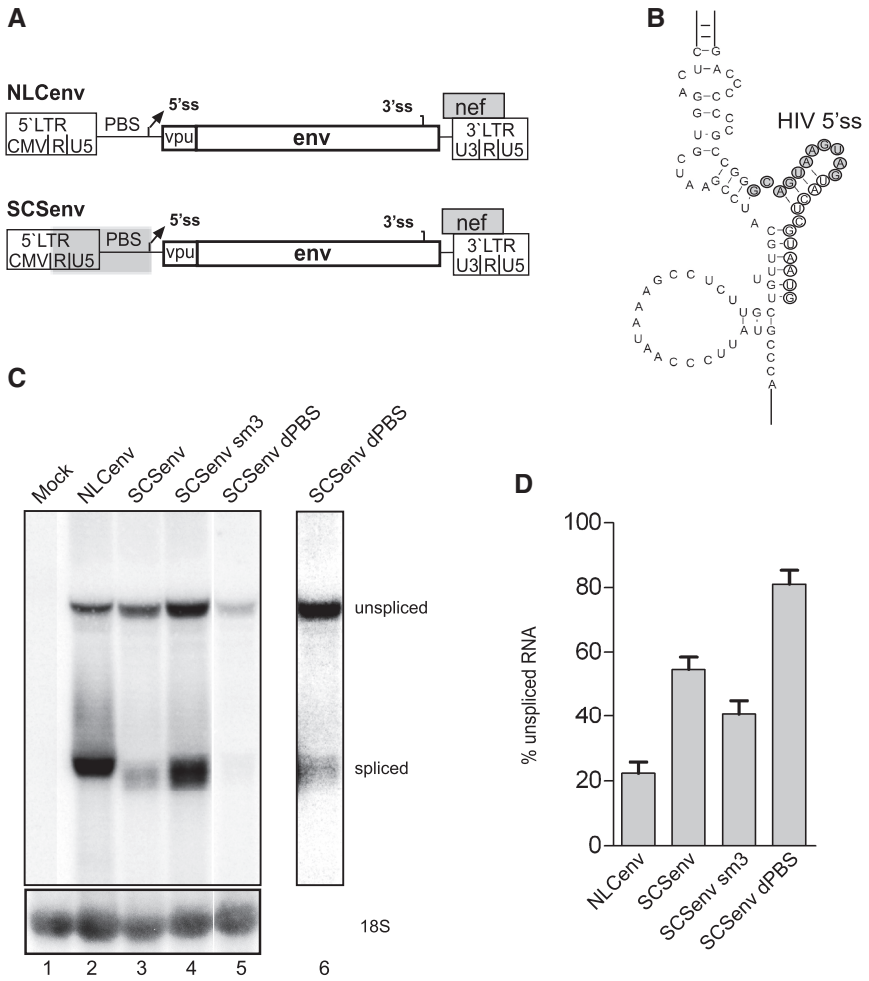

Figure 3. The MLV secondary structure is able to repress an HIV 5′ss. (**A**) Schematic drawing of the NLCenv plasmid. NLCenv is derived from the HIV molecular clone NL4-3. Expression is directed by the authentic LTR sequences except that the HIV promoter was replaced by CMV. ORFs are shown as boxes. The splice sites are indicated as well as the primer binding site. Below this, the hybrid construct SCSenv is shown. The grey region highlights the SF91-derived sequence in the HIV backbone. (**B**) The lower part of the secondary structure of MLV/HIV fusion transcript is shown. The HIV 5′ss is highlighted by grey circles and additional HIV-derived sequences as open circles. The HIV 5′ss is forced into an inhibitory confirmation due to the upstream MLV stem loop. (**C**) Northern blot of 10 µg of total RNA from HeLaP4 cells transiently transfected with 2.75 µg of the indicated constructs. Two splicing mutations were transferred to SCSenv (sm3 and dPBS; Figure 1B). The RNA species are named on the right. The lane containing SCSenv dPBS was overexposed to display the ratio between unspliced and spliced RNA (lane 6). (**D**) Phosphoimager analysis as described in Figure 2C. The mean values and standard deviations of five independent experiments are shown.

RNA:RNA interactions at weak splice sites. Since the secondary structure may restrict access to U1 snRNP, we anticipated that targeting an SR protein into the vicinity of the 5′ss would overcome this attenuation by supporting 5′ss:U1 snRNA duplex formation. Tethering of protein domains to an RNA of interest became possible with the MS2-fusion tethering system (48,49). The coat protein of the MS2 phage binds with high affinity to its target sequence, which is a short RNA stem–loop (Figure 5A). However, due to structural constraints, our heterologous RNA stem–loop accommodates only one MS2 binding site. We therefore used the ΔFG variant of the MS2 coat protein, which leads to an increase in RNA binding affinity at the expense of dimer:dimer formation (50). This modification should theoretically result in enhanced binding to a single site *in vivo* (Figure 5A). In a co-transfection assay, we tested activation domains (RS domains) of various SR proteins fused to MS2ΔFG along with the SF91 construct and two RNA stem–loop variants harboring a single MS2 binding site (Figure 5B). Transfection of SF91 and a plasmid encoding the RS domain of SRp55 fused to MS2ΔFG is not neutral, as observed for SF91, and leads to a slight statistically significant enhancement of splicing (Figure 5B, lanes 1 and 2 and Figure 5C, *P*-value = 0.03). However, in the context of the RNA stem–loop and the extended stem–loop construct, co-transfection of the MS2-SRp55 plasmid led to a highly significant enhancement of splicing (Figure 5B and C, compare lanes 3–6; *P*-value = 0.009 and 0.002).

In general, these experiments proved for the first time *in vivo* that MS2ΔFG mutants can be targeted to a single binding site with reasonable efficiency. Moreover, tethering of an RS domain to the heterologous RNA

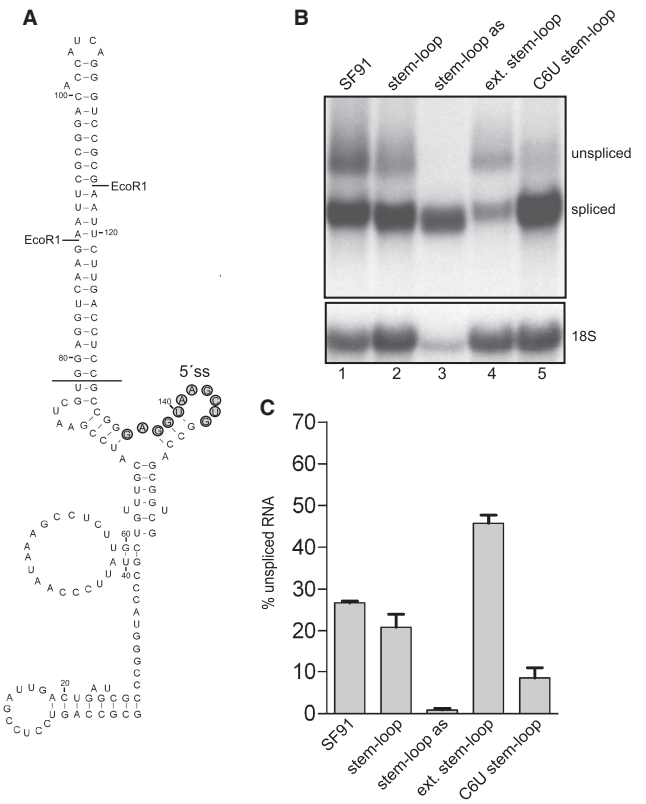

Figure 4. A heterologous stem–loop substitutes for the MLV secondary structure. (**A**) The secondary structure is depicted as in Figure 1B. The horizontal black line marks the insertion of heterologous sequences forming the synthetic stem–loop. On each side of the stem an EcoRI restriction site is indicated. The 5′ss is highlighted by grey circles. (**B**) Northern blot analysis of 10 µg total RNA from transfected 293T cells. The RNA species are marked on the right. Note that only 5 µg RNA were loaded in the stem–loop antisense (as) lane due to the high expression of this construct. (**C**) Phosphoimager analysis. Mean values and standard deviation represent three independent experiments.

stem–loop partially overcomes 5′ss attenuation, in agreement with previous results that the RNA structure may restrict access to U1 snRNP.

5′ splice site attenuation can be rescued by hyperstable U1 snRNA suppressor mutants

In addition to protein:protein contacts of U1 snRNP with exonic or intronic sequences (51,52), recognition of 5′ss is initiated by RNA:RNA interactions (9,53). Therefore, the hydrogen bonding pattern between U1 snRNA and the 5′ss is critical and can either be enhanced by mutations within the 5′ss (C6U mutation, Figure 1C, lane 2) or by overexpression of U1 snRNA suppressor mutants that increase the complementarity to a given 5′ss (9,54). In yeast, hyperstable U1 snRNA mutants cannot be displaced by U6 snRNA and therefore splicing is inhibited (55). In contrast, in mammalian cells, an extended U1 snRNA/5′ss interaction does not decrease splicing efficiency, but rather increases 5′ss recognition (56,57). We constructed two U1 snRNA mutants containing one substitution (U1 G11C) leading to eight complementary base pairs with the MLV 5′ss (Figure 6A; HBond score 18.8) or a perfect match of U1 snRNA leading to 11 continuous base pairs (U1 perfect; Figure 6A; HBond score 23.8). Overexpression of these suppressor mutants led to an increase in splicing depending on the intrinsic strength of the RNA duplex (Figure 6B and C). An increase in complementarity of one base pair (U1 G11C mutant, Figure 6A, lower panel) already led to an enhancement of splicing (Figure 6C). This argues for a dynamic balance between secondary structure and accessibility of the 5′ss to U1 snRNA. As a control, we co-transfected the structural sm3 mutant along with the U1 perfect suppressor snRNA and observed no change in the enhancement of splicing (Supplementary Figure 1).

In addition, we looked at direct interaction between U1 snRNP and the 5′ss *in vitro* by EMSA. In order to minimize unspecific protein:RNA interactions, we used the heterologous stem–loop and the antisense mutant thereof as depicted in Figure 4A. We first established conditions that permit folding of the secondary structure, but not of the antisense mutant (Figure 6D, lanes 1 and 2). Under denaturing conditions, both RNAs showed a similar migration behavior (data not shown). Purified U1 snRNP was incubated with the *in vitro* transcribed, ^{32}P-labeled and folded RNA. We observed an enhanced binding of U1 snRNP to the 5′ss in the absence of the RNA secondary structure (i.e. the antisense mutant, Figure 6D, lanes 4, 6 and 8) and reduced binding upon formation of the structure (Figure 6D, lanes 3, 5 and 7).

These experiments demonstrate that splicing regulation in MLV uses restricted access of U1 snRNP to the 5′ss exerted by the upstream secondary structure and limited complementarity of the primary 5′ss sequence to U1 snRNA.

DISCUSSION

As presented here, murine leukemia virus uses a dynamic interplay between RNA secondary structure and the intrinsic strength of the 5′ss to restrict access of U1 snRNP to its 5′ss, which ultimately results in alternative splicing and full viral gene expression.

Secondary structure has been implicated in alternative splicing early on (18,58–60). Cellular examples of attenuated 5′ss, which are part of a secondary structure, were discovered in association with different genetic diseases (56,61,62). Modulation of splicing efficiency by RNA secondary structures has recently also been described for the HIV-1 leader RNA structure, where the major 5′ss is embedded in a semi-stable hairpin (63). However, contrary to HIV-1, splicing regulation at the MLV 5′ss seems to be much more complex, since stem mutations even 25 nt upstream of the 5′ss already provoke an 8-fold enhancement of splicing (Figure 1). Certainly, secondary structures cannot only sequester 5′ss, but also modulate the binding efficiency of hnRNPs or SR proteins. On a global scale, it was shown that splicing enhancers and silencers are present mostly in single-stranded regions (64). There is also a particular

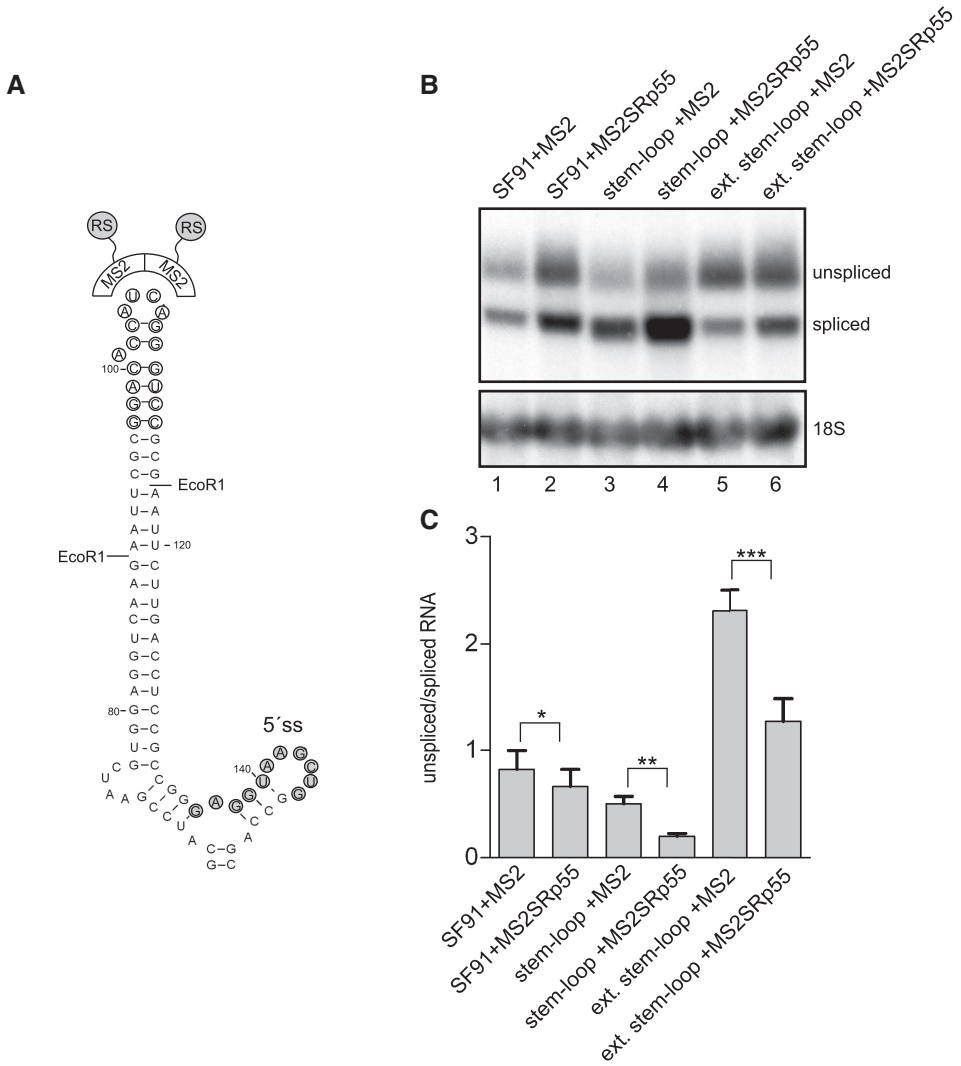

Figure 5. An RS domain can partially relieve 5′ss attenuation. (**A**) Depiction of the heterologous stem–loop structure as in Figure 4A. In addition to the 5′ss, the MS2 binding site is highlighted by open circles. MS2-RS fusion proteins recognize the MS2 binding site as dimers. (**B**) Northern blot performed as in Figure 1C. As a control, a plasmid encoding only the MS2 coat protein was co-transfected as indicated. The RS domain of SRp55 was fused to the modified MS2 protein. The RNA species are marked on the right. (**C**) Phosphoimager analysis. Splicing efficiency is displayed as unspliced/spliced RNA ratio. Student's *t*-test was performed using mean values from four independent experiments. $^{*}P = 0.03$; $^{**}P = 0.009$; $^{***}P = 0.002$.

example in HIV, where a change in secondary structure allows hnRNP H to bind and influence splicing (65). RNA secondary structures are also statistically associated with alternative 5′ss (66), allowing splice site selection via conformational variability.

An exchange of the upper part of the structure with a heterologous stem–loop proved that the main function of the stem is to force the 5′ss into an inhibitory conformation and the extent of splicing inhibition correlates closely with the free energy of the structure (Figure 4). Thus, the secondary structure can be viewed as a silencer element, which converts a strong 5′ss into a weak one. Not surprisingly, the complex RNA stem–loop possesses additional functions in the viral life cycle. The structure allows looping of the primer binding site, which binds a cellular tRNA as a primer to initiate reverse transcription (3). Yet, splicing regulation could be transferred to a complete provirus and to a heterologous HIV 5′ss (Figures 2 and 3). Using the MLV/HIV hybrid plasmids, it seemed that the degree of 5′ss inhibition correlates inversely with the overall RNA amount (Figure 3C). Similar effects have been observed in other studies, where splice sites are able to enhance transcriptional elongation (67) and gene expression in general (68). It appears that the CMV promoter is highly dependent on this positive feedback exerted by the interaction of U1 snRNP with a proximal 5′ss (69,70). This suggested to us that the secondary structure restricts access of U1 snRNP. Additional evidence was obtained by tethering RS domains to the heterologous stem–loop, which enhanced splicing (Figure 5). The relatively modest effects are most probably due to the insertion of only a single MS2 binding site (Figure 5A). The MS2ΔFG mutants used here can partly compensate

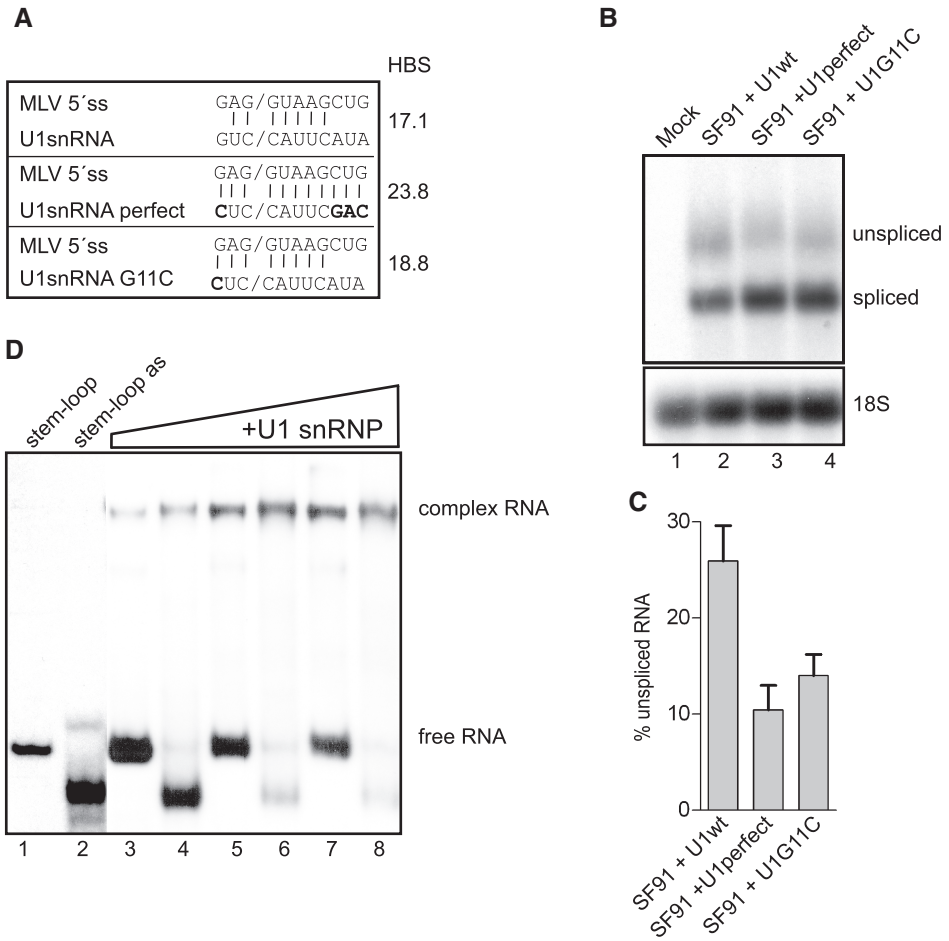

Figure 6. Hyperstable U1 snRNA suppressor mutants enhance splicing. (**A**) The top line shows the MLV 5′ss. Possible base pairs to U1 snRNA are indicated by vertical lines. In the middle and lower panel, pairing of the MLV 5′ss with the two U1 snRNA mutants is depicted. Mutated nucleotides in U1 snRNA are in bold. On the right-hand side, the HBond score (HBS) is given. (**B**) Northern blot using 10 µg of total RNA from transiently transfected 293T cells. Transfections were performed using 5 µg of SF91 plasmid and 10 µg of the respective U1 snRNA plasmid. The RNA species are named on the right. (**C**) Phosphoimager analysis. The efficiency of splicing is shown as the relative level of unspliced RNA. SF91 in the presence of co-transfected U1 wt plasmid was set to 1. Mean values and standard deviations represent five independent experiments. (**D**) Mobility shift assay using *in vitro* transcribed, ^{32}P-labeled and folded MLV RNA depicted in Figure 4A. The RNAs were incubated with increasing amounts of purified U1 snRNP (50 ng, 125 ng and 250 ng).

for the lack of multiple binding sites. RS domains may play a dual role in splice site selection. They can enhance RNA:RNA interactions at degenerated splice sites (47,71) or engage in protein:protein interactions, since an excess of SR proteins can select 5′ss in the absence of U1 snRNP (72). In our case, the RS domain may directly assist U1 snRNA binding to the attenuated 5′ss by interacting with U1-70K (73).

In addition, the retroviral 5′ss is characterized by a mismatch at position +6. Reversion of this position into a complementary nucleotide to U1 snRNA enhances splicing 2.5-fold. Although this position is less conserved on a genome scale, it turns into a preserved nucleotide if the 5′ss is degenerated or weakened by surrounding elements (10). For example, a U6C mutation in the 5′ss flanking exon 20 of the IKBKAP gene causes exon skipping, resulting in familial dysautonomia (74). In this case, pre-existing attenuation is due to a weak upstream 3′ss and possible splicing silencers (75). An additional weakening of the 5′ss finally results in the disease-causing splicing phenotype. In this line, position +6 of the MLV 5′ss regulates efficiency only in the context of the secondary structure. Therefore, there is a dynamic interplay between structure and the 5′ss sequence, where the structure makes the 5′ss susceptible to silencing as it was observed for cellular silencer motifs (76). The same result was obtained when we fused the complete second intron from β-globin to the MLV secondary structure. Alternative splicing occurred only after lowering the complementarity at positions +6 or −3 of the globin 5′ss (Zychlinski,D. and Bohne,J. unpublished data). We also mutated positions +7 and +8 of the MLV 5′ss back to the consensus and observed an enhancement of splicing (data not shown). However, the enhancement was not as strong as observed for position +6. One may speculate that not solely complementarity to U1 snRNA, but also the length and neighborhood of the 5′ss:U1 snRNA duplex determine the strength of a splice site.

This novel dynamic interplay of RNA secondary structure and low complementarity to U1 snRNA of the

5′ss demonstrates the fine-tuning of alternative splicing in retroviruses and mammalian cells and illustrates the sophisticated organization of the splicing code.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank M. Garcia-Blanco for critically reading the manuscript and for his helpful suggestions during the course of this work, M. Galla, J. Kraunus, A. Schambach and J. Meyer for discussion and experimental advice and T. Schulz for his generous support.

FUNDING

DFG grant BO 2512/2-1 (to J.B.); SCHA 909/2-2 (to H.S.); the Stiftung für AIDS-Forschung, Düsseldorf to H. Schaal; the DFG Cluster of Excellence REBIRTH; and SFB 738 project C4 to (C.B.). Funding for open access charge: DFG grant BO 2512/2-1 (to J.B.).

Conflict of interest statement. None declared.

REFERENCES

1. Wang,E.T., Sandberg,R., Lou,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P. and Burge,C.B. (2008) Alternative isoform regulation in the human transcriptomes. *Nature*, **456**, 471–476.
2. Maniatis,T. and Tasic,B. (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature*, **418**, 236–243.
3. Goff,S.P. (2001) Retroviridae: the retroviruses and their replication. In Knipe,D.M. and Howley,P.M. (eds), *Fields Virology*, Vol. 2, 4th edn. Lippincott Williams and Wilkins, Philadelphia, PA, pp. 1871–1940.
4. Wahl,M.C., Will,C.L. and Lührmann,R. (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell*, **136**, 701–718.
5. Will,C.L. and Lührmann,R. (2006) *The RNA World.,* 3rd edn. CSHL Press.
6. Valadkhan,S. (2007) The spliceosome: a ribozyme at heart? *Biol. Chem*., **388**, 693–697.
7. Konarska,M.M. and Query,C.C. (2005) Insights into the mechanisms of splicing: more lessons from the ribosome. *Genes Dev*., **19**, 2255–2260.
8. Mount,S.M., Pettersson,I., Hinterberger,M., Karmas,A. and Steitz,J.A. (1983) The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site in vitro. *Cell*, **33**, 509–518.
9. Zhuang,Y. and Weiner,A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell*, **46**, 827–835.
10. Hartmann,L., Theiss,S., Niederacher,D. and Schaal,H. (2008) Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? *Front. Biosci*., **13**, 3252–3272.
11. Roca,X., Sachidanandam,R. and Krainer,A.R. (2005) Determinants of the inherent strength of human 5′ splice sites. *RNA*, **11**, 683–698.
12. Burge,C.B., Tuschl,T. and Sharp,P.A. (1999) In Cech,T.R. (ed.), *The RNA World*, 2 edn. Cold Spring Harbor, NY, pp. 525–560.
13. Hastings,M.L. and Krainer,A.R. (2001) Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol*., **13**, 302–309.
14. Black,D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem*., **72**, 291–336.
15. Tange,T.O., Damgaard,C.K., Guth,S., Valcarcel,J. and Kjems,J. (2001) The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J*., **20**, 5748–5758.
16. Wang,Z. and Burge,C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*, **14**, 802–813.
17. Hertel,K.J. (2008) Combinatorial control of exon recognition. *J. Biol. Chem*., **283**, 1211–1215.
18. Solnick,D. (1985) Alternative splicing caused by RNA secondary structure. *Cell*, **43**, 667–676.
19. Buratti,E. and Baralle,F.E. (2004) Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol. Cell Biol*., **24**, 10505–10514.
20. Marchand,V., Mereau,A., Jacquenet,S., Thomas,D., Mougin,A., Gattoni,R., Stevenin,J. and Branlant,C. (2002) A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J. Mol. Biol*., **323**, 629–652.
21. Kornblihtt,A.R., de la Mata,M., Fededa,J.P., Munoz,M.J. and Nogues,G. (2004) Multiple links between transcription and splicing. *RNA*, **10**, 1489–1498.
22. Cochrane,A.W., McNally,M.T. and Mouland,A.J. (2006) The retrovirus RNA trafficking granule: from birth to maturity. *Retrovirology*, **3**, 18.
23. Purcell,D.F. and Martin,M.A. (1993) Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J. Virol*., **67**, 6365–6378.
24. Amendt,B.A., Si,Z.H. and Stoltzfus,C.M. (1995) 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**, 4606–4615.
25. Caputi,M., Mayeda,A., Krainer,A.R. and Zahler,A.M. (1999) hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J*., **18**, 4060–4067.
26. Kammler,S., Otte,M., Hauber,I., Kjems,J., Hauber,J. and Schaal,H. (2006) The strength of the HIV-1 3′ splice sites affects Rev function. *Retrovirology*, **3**, 89.
27. Stoltzfus,C.M. and Madsen,J.M. (2006) Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. *Curr. HIV Res*., **4**, 43–55.
28. O'Reilly,M.M., McNally,M.T. and Beemon,K.L. (1995) Two strong 5′ splice sites and competing, suboptimal 3′ splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology*, **213**, 373–385.
29. Jacquenet,S., Ropers,D., Bilodeau,P.S., Damier,L., Mougin,A., Stoltzfus,C.M. and Branlant,C. (2001) 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**, 464–478.
30. McNally,M.T. and Beemon,K. (1992) Intronic sequences and 3′ splice sites control Rous sarcoma virus RNA splicing. *J. Virol*., **66**, 6–11.
31. Lee,J.T., Yu,S.S., Han,E. and Kim,S. (2004) Engineering the splice acceptor for improved gene expression and viral titer in an MLV-based retroviral vector. *Gene Ther*., **11**, 94–99.
32. Logg,C.R., Baranick,B.T., Lemp,N.A. and Kasahara,N. (2007) Adaptive evolution of a tagged chimeric gammaretrovirus: identification of novel cis-acting elements that modulate splicing. *J. Mol. Biol*., **369**, 1214–1229.
33. McNally,L.M. and McNally,M.T. (1999) U1 small nuclear ribonucleoprotein and splicing inhibition by the rous sarcoma virus negative regulator of splicing element. *J. Virol*., **73**, 2385–2393.
34. Kraunus,J., Zychlinski,D., Heise,T., Galla,M., Bohne,J. and Baum,C. (2006) Murine leukemia virus regulates alternative splicing through sequences upstream of the 5′ splice site. *J. Biol. Chem*., **281**, 37381–37390.
35. Hildinger,M., Abel,K.L., Ostertag,W. and Baum,C. (1999) Design of 5′ untranslated sequences in retroviral vectors developed for medical use. *J. Virol*., **73**, 4083–4089.
36. Schambach,A., Mueller,D., Galla,M., Verstegen,M.M., Wagemaker,G., Loew,R., Baum,C. and Bohne,J. (2006) Overcoming promoter competition in packaging cells improves production of self-inactivating retroviral vectors. *Gene Ther*., **13**, 1524–1533.
37. Bohne,J., Schambach,A. and Zychlinski,D. (2007) New way of regulating alternative splicing in retroviruses: the promoter makes a difference. *J. Virol*., **81**, 3652–3656.

38. Erlwein,O., Buchholz,C.J. and Schnierle,B.S. (2003) The proline-rich region of the ecotropic Moloney murine leukaemia virus envelope protein tolerates the insertion of the green fluorescent protein and allows the generation of replication-competent virus. *J. Gen. Virol.*, **84**, 369–373.
39. Caputi,M., Freund,M., Kammler,S., Asang,C. and Schaal,H. (2004) A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *J. Virol.*, **78**, 6517–6526.
40. Kimpton,J. and Emerman,M. (1992) Detection of replication-competent and pseudotyped human immunodeficiency virus with a sensitive cell line on the basis of activation of an integrated beta-galactosidase gene. *J. Virol.*, **66**, 2232–2239.
41. Wodrich,H., Bohne,J., Gumz,E., Welker,R. and Krausslich,H.G. (2001) A new RNA element located in the coding region of a murine endogenous retrovirus can functionally replace the Rev/Rev-responsive element system in human immunodeficiency virus type 1 Gag expression. *J. Virol.*, **75**, 10670–10682.
42. Bach,M., Krol,A. and Lührmann,R. (1990) Structure-probing of U1 snRNPs gradually depleted of the U1-specific proteins A, C and 70k. Evidence that A interacts differentially with developmentally regulated mouse U1 snRNA variants. *Nucleic Acids Res.*, **18**, 449–457.
43. Zuker,M. (1989) Computer prediction of RNA structure. *Methods Enzymol.*, **180**, 262–288.
44. Bohne,J., Wodrich,H. and Krausslich,H.G. (2005) Splicing of human immunodeficiency virus RNA is position-dependent suggesting sequential removal of introns from the 5′ end. *Nucleic Acids Res.*, **33**, 825–837.
45. Freund,M., Asang,C., Kammler,S., Konermann,C., Krummheuer,J., Hipp,M., Meyer,I., Gierling,W., Theiss,S., Preuss,T. *et al.* (2003) A novel approach to describe a U1 snRNA binding site. *Nucleic Acids Res.*, **31**, 6963–6975.
46. Kammler,S., Leurs,C., Freund,M., Krummheuer,J., Seidel,K., Tange,T.O., Lund,M.K., Kjems,J., Scheid,A. and Schaal,H. (2001) 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**, 421–434.
47. Shen,H. and Green,M.R. (2006) RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes Dev.*, **20**, 1755–1765.
48. Selby,M.J. and Peterlin,B.M. (1990) Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. *Cell*, **62**, 769–776.
49. Peabody,D.S. (1993) The RNA binding site of bacteriophage MS2 coat protein. *EMBO J.*, **12**, 595–600.
50. Chao,J.A., Patskovsky,Y., Almo,S.C. and Singer,R.H. (2008) Structural basis for the coevolution of a viral RNA-protein complex. *Nat. Struct. Mol. Biol.*, **15**, 103–105.
51. Lund,M. and Kjems,J. (2002) Defining a 5′ splice site by functional selection in the presence and absence of U1 snRNA 5′ end. *RNA*, **8**, 166–179.
52. Du,H. and Rosbash,M. (2002) The U1 snRNP protein U1C recognizes the 5′ splice site in the absence of base pairing. *Nature*, **419**, 86–90.
53. Seraphin,B., Kretzner,L. and Rosbash,M. (1988) A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site. *EMBO J.*, **7**, 2533–2538.
54. Nandabalan,K., Price,L. and Roeder,G.S. (1993) Mutations in U1 snRNA bypass the requirement for a cell type-specific RNA splicing factor. *Cell*, **73**, 407–415.
55. Staley,J.P. and Guthrie,C. (1999) An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell*, **3**, 55–64.
56. Singh,N.N., Singh,R.N. and Androphy,E.J. (2007) Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Res.*, **35**, 371–389.
57. Freund,M., Hicks,M.J., Konermann,C., Otte,M., Hertel,K.J. and Schaal,H. (2005) Extended base pair complementarity between U1 snRNA and the 5′ splice site does not inhibit splicing in higher eukaryotes, but rather increases 5′ splice site recognition. *Nucleic Acids Res.*, **33**, 5112–5119.
58. Eperon,L.P., Graham,I.R., Griffiths,A.D. and Eperon,I.C. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell*, **54**, 393–401.
59. Fu,X.Y. and Manley,J.L. (1987) Factors influencing alternative splice site utilization in vivo. *Mol. Cell Biol.*, **7**, 738–748.
60. Goguel,V., Wang,Y. and Rosbash,M. (1993) Short artificial hairpins sequester splicing signals and inhibit yeast pre-mRNA splicing. *Mol. Cell Biol.*, **13**, 6841–6848.
61. Bratt,E. and Ohman,M. (2003) Coordination of editing and splicing of glutamate receptor pre-mRNA. *RNA*, **9**, 309–318.
62. Grover,A., Houlden,H., Baker,M., Adamson,J., Lewis,J., Prihar,G., Pickering-Brown,S., Duff,K. and Hutton,M. (1999) 5′ splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *J. Biol. Chem.*, **274**, 15134–15143.
63. Abbink,T.E. and Berkhout,B. (2008) RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor. *J. Virol.*, **82**, 3090–3098.
64. Hiller,M., Zhang,Z., Backofen,R. and Stamm,S. (2007) Pre-mRNA secondary structures influence exon recognition. *PLoS Genet.*, **3**, e204.
65. Jablonski,J.A., Buratti,E., Stuani,C. and Caputi,M. (2008) The secondary structure of the human immunodeficiency virus type 1 transcript modulates viral splicing and infectivity. *J. Virol.*, **82**, 8038–8050.
66. Shepard,P.J. and Hertel,K.J. (2008) Conserved RNA secondary structures promote alternative splicing. *RNA*, **14**, 1463–1469.
67. Fong,Y.W. and Zhou,Q. (2001) Stimulatory effect of splicing factors on transcriptional elongation. *Nature*, **414**, 929–933.
68. Furger,A., O'Sullivan,J.M., Binnie,A., Lee,B.A. and Proudfoot,N.J. (2002) Promoter proximal splice sites enhance transcription. *Genes Dev.*, **16**, 2792–2799.
69. Damgaard,C.K., Kahns,S., Lykke-Andersen,S., Nielsen,A.L., Jensen,T.H. and Kjems,J. (2008) A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell*, **29**, 271–278.
70. Bohne,J. and Krausslich,H.G. (2004) Mutation of the major 5′ splice site renders a CMV-driven HIV-1 proviral clone Tat-dependent: connections between transcription and splicing. *FEBS Lett.*, **563**, 113–118.
71. Valcarcel,J., Gaur,R.K., Singh,R. and Green,M.R. (1996) Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA. *Science*, **273**, 1706–1709.
72. Crispino,J.D., Blencowe,B.J. and Sharp,P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866–1869.
73. Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) Protein-protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. *Nature*, **368**, 119–124.
74. Slaugenhaupt,S.A., Blumenfeld,A., Gill,S.P., Leyne,M., Mull,J., Cuajungco,M.P., Liebert,C.B., Chadwick,B., Idelson,M., Reznik,L. *et al.* (2001) Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. *Am. J. Hum. Genet.*, **68**, 598–605.
75. Ibrahim el,C., Hims,M.M., Shomron,N., Burge,C.B., Slaugenhaupt,S.A. and Reed,R. (2007) Weak definition of IKBKAP exon 20 leads to aberrant splicing in familial dysautonomia. *Hum. Mutat.*, **28**, 41–53.
76. Yu,Y., Maroney,P.A., Denker,J.A., Zhang,X.H., Dybkov,O., Lührmann,R., Jankowsky,E., Chasin,L.A. and Nilsen,T.W. (2008) Dynamic regulation of alternative splicing by silencers that modulate 5′ splice site competition. *Cell*, **135**, 1224–1236.
77. Mougel,M., Tounekti,N., Darlix,J.L., Paoletti,J., Ehresmann,B. and Ehresmann,C. (1993) Conformational analysis of the 5′ leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. *Nucleic Acids Res.*, **21**, 4677–4684.