

Regulation of human and HIV-1 splice sites

Die Regulation von humanen und HIV-1 Spleißstellen

Inaugural-Dissertation

zur Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Steffen Erkelenz

aus Krefeld

Düsseldorf, April 2012

aus dem Institut für Virologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. H. Schaal

Koreferent: Prof. Dr. D. Willbold
Prof. Dr. A. Bindereif

Tag der mündlichen Prüfung: 23.10.2012

FÜR MEINE ELTERN

Zusammenfassung

Unabhängig von der Anzahl der proteinkodierenden Gene wird durch das alternative Spleißen das humane Proteom vergrößert. Die Nutzung von Spleißstellen hängt im erheblichen Maße von spleißregulatorischen Elementen (SRE) in ihrer unmittelbaren Nachbarschaft ab. Diese werden von spleißregulatorischen Proteinen gebunden und können sowohl aktivatorisch als auch hemmend auf die Spleißstellennutzung wirken. Im ersten Teil dieser Doktorarbeit wurde gezeigt, dass im Gegensatz zur langläufigen Klassifizierung der SR Proteine als Aktivatoren und hnRNP Proteine als Repressoren, nahezu alle Spleißfaktoren die Eigenschaft besitzen, positionsabhängig sowohl positiv als auch negativ auf die Spleißstellennutzung wirken zu können. SR Proteine förderten das Spleißen ausschließlich von einer Position im 5'-wärts gelegenen Exon, während hnRNP und hnRNP-ähnliche Proteine hierfür eine Position im 3'-wärts gelegenen Intron benötigten. Erstaunlicherweise waren alle aktivatorischen Spleißproteine nach ihrer Positionierung an die jeweils gegenüberliegende Seite der 5'ss auch dazu in der Lage das Spleißen zu reprimieren. Die Spleißinhibition schien nicht durch eine fehlende Erkennung der 5'ss durch das U1 snRNP bedingt zu sein, sondern durch eine Unfähigkeit zur Bildung späterer Spleißosomintermediate.

Im zweiten Teil konnte ein HIV-1 exonischer Spleißenhancer – genannt ESE_{vpr} – identifiziert werden, der für die *vpr*-mRNA Expression und den Einschluss des nichtkodierenden Exons 3 in die verschiedenen viralen mRNAs erforderlich ist. Seine Inaktivierung führte zu einer dramatischen Verringerung der *vpr*-mRNA Mengen in Provirus-transfizierten Zellen. Die ESE_{vpr} Aktivität konnte, in Abwesenheit des 5'-wärts gelegenen spleißinhibitorischen ESSV, für einen exzessiven Exon 3 Spleißphänotyp sowie einen schweren Defekt in der viralen Replikation verantwortlich gemacht werden. Eine gleichzeitige Mutation des ESE_{vpr} konnte die Virusproduktion wiederherstellen. Demnach scheinen ESSV und ESE_{vpr} in kombinatorischer Weise die Aktivierung der Exon 3 Spleißstellen zu regulieren und für ein ausgeglichenes Exon 3 Spleißen essenziell zu sein.

Summary

Regardless of the number of protein-coding genes, alternative splicing significantly expands the human proteome. Utilization of splice sites critically depends on splicing regulatory elements (SREs) in their immediate neighbourhood. These are bound by nuclear RNA binding proteins and can act through either activating or repressing splice site usage.

In the first part of this thesis, it was shown that in contrast to long-held classifications of SR proteins as general enhancers and hnRNP proteins as general repressors of splicing, almost all splicing factors show the common peculiarity to positively and negatively act on splice site usage. SR proteins enhanced splicing only from the upstream exonic position, while hnRNP or hnRNP-like proteins needed a location within the downstream intron. Remarkably, all activating proteins were also capable of repressing splicing by simply relocating to the opposite side of the 5'ss. Splicing inhibition appeared to not be due to a failure of the U1 snRNP to recognize the 5'ss, but rather an inability to form later spliceosomal intermediates.

In the second part of this work an HIV-1 exonic splicing enhancer – termed ESE_{vpr} – was identified, which was required for *vpr*-mRNA expression and the inclusion of the *non*-coding exon 3 into distinct viral mRNA species. Inactivation of the ESE_{vpr} led to a dramatic decline in the amounts of *vpr*-mRNA within provirus-transfected cells. An excessive exon 3 splicing phenotype as well as a severe defect in viral replication in absence of the upstream splicing inhibitory ESSV could be attributed to the ESE_{vpr} activity. Simultaneous mutation of the ESE_{vpr} could restore virus particle production. Therefore, ESSV and ESE_{vpr} appear to combinatorially regulate the activation of the exon 3 splice sites and to be essential for a balanced exon 3 splicing.

Table of contents

ZUSAMMENFASSUNG	I
SUMMARY	II
I. INTRODUCTION	1
I.1 The Human Immunodeficiency Virus Type 1 (HIV-1)	1
I.2 Expression of the entire HIV-1 protein arsenal is controlled by alternative splicing	2
I.2.1 HIV-1 replication cycle	2
I.2.2 Alternative splicing of the HIV-1 pre-mRNA allows the expression of the entire viral proteome	6
I.3 Mechanism of pre-mRNA splicing	8
I.3.1 Spliceosome assembly and catalysis	8
I.3.2 Exon and intron definition complexes	14
I.3.3 Splice site recognition	16
I.3.4 Alternative splicing and <i>cis</i> -regulatory elements	25
I.3.5 Splicing regulatory proteins	27
I.4 Alternative splicing by HIV	45
I.4.1 <i>Cis</i> -regulatory elements control HIV-1 pre-mRNA splicing	47
I.4.2 Expression of intron-containing <i>vpr</i> -mRNAs and HIV-1 exon 3 splicing regulation	51
I.4.3 Manipulation of the host cell splicing machinery to optimize viral replication	53
I.5 Aim of this work	55
II. MATERIALS AND METHODS	56
II.1 Materials	56
II.1.1 Chemicals, culture media and solvents	56
II.1.2 Enzymes	58
II.1.3 Cells	59
II.1.4 Oligonucleotides	59
II.1.5 Recombinant plasmids	71
II.1.6 Antibodies	82
II.1.7 Algorithms and databases	83
II.2 Methods	85
II.2.1 Cloning	85

Table of contents

II.2.2 Eukaryotic cell culture _____	88
II.2.3 Flow cytometrical analysis of transiently transfected HeLa cells _____	90
II.2.4 Western Blot Analysis _____	90
II.2.5 Immunoprecipitation (IP) _____	92
II.2.6 Reverse transcriptase (RT)-PCR analysis _____	92
II.2.7 RNA affinity chromatography _____	97
II.2.8 <i>In vitro</i> splicing _____	102
 III. RESULTS _____	 103
III.1 “Don’t buy the house, buy the neighbourhood”: relative positioning determines the phenotype of a splicing regulatory protein _____	103
III.1.1 The position of splicing regulatory elements relative to a 5’ss determines the splicing outcome _____	104
III.1.2 The number of binding sites positively correlates with 5’ss use _____	114
III.1.3 Multiple enhancing PDEs compensate for weak intrinsic strength of a 5’ss _____	117
III.1.4 MS2-tethering of splicing regulatory proteins mimics their position-dependence for 5’ss activation _____	119
III.1.5 Splicing regulatory proteins modulate the competition of two 5’ss _____	122
III.1.6 Position-dependent enhancers (PDEs) can act as silencers of splicing _____	124
III.1.7 The progression into a functional spliceosome is determined by the side from which U1 snRNP is recruited to the 5’ss _____	131
III.1.8 SRSF7 PDE-depleted nuclear extracts inhibits proximal 5’ss use _____	138
III.1.9 Recruitment of U1 snRNP by hnRNP and hnRNP-like proteins might not exclusively be mediated through interactions with U1-C _____	140
 III.2 “S(p)lice up the defence”: How SREs within the HIV-1 exon 3 act in concert to regulate expression of viral infectivity factor Vpr _____	 142
III.2.1 An exonic splicing enhancer (ESE) acts positively on the inclusion of the exon 3 into reporter mRNAs _____	143
III.2.2 Inactivation of the exonic splicing enhancer restores viral replication in the context of an ESSV-negative provirus _____	153
III.2.3 Recognition of to the viral 5’ss D3 is critical for viral Vpr expression _____	163
III.2.4 Increased activation of 3’ss A3 reduces <i>vpr</i> -mRNA expression in the context of the HIV-1 based 4-exon minigene _____	178
III.2.5 TAR DNA binding protein-43 (TDP-43) is not involved in positive splicing regulation of HIV-1 exon 3 _____	180
 IV. DISCUSSION _____	 183
IV.1 Position-dependent splicing activation and repression by SR and hnRNP proteins _____	183

Table of contents

IV.1.1 Almost all common splicing regulatory proteins can enhance U1 snRNP recruitment from defined positions	184
IV.1.2 The overall ability of the spliceosome to recognize a 5'ss is the net result of U1 snRNA complementarity and enhancer strength	192
IV.1.3 Disclosing a "SRcret" in splice site regulation: Repression by SR proteins is coupled to "dead-end" U1 snRNP recruitment	193
IV.1.4 Inhibition of spliceosome assembly after 5'ss recognition might promote intron retention versus exon skipping	199
IV.1.5 Future prospects	203
IV.2 An exonic splicing enhancer (ESE _{vpr}) within HIV-1 exon 3 is required for exon 3 splicing and <i>vpr</i> -mRNA expression	204
IV.2.1 An exonic splicing enhancer (ESE _{vpr}) promotes exon 3 inclusion and formation of <i>vpr</i> -mRNA	204
IV.2.2 Intronic SR proteins might be implicated in specific retention of the <i>vpr</i> -intron	206
IV.2.3 HMGA1a might trap the U1 snRNP at 5'ss D3 to promote downstream <i>vpr</i> -intron retention	210
IV.2.4 Future prospects	212
V. CONCLUSIONS	213
VI. REFERENCES	214
VII. APPENDIX	276
VII.1 Experimental Data	276
VII.2 Abbreviations and units	286
VII.2.1 Abbreviations	286
VII.2.2 Units	288
VII.3 Publications	289
VII.4 Curriculum Vitae	291
VII.5 Erklärung	292
DANKSAGUNG	293

I. Introduction

I.1 The Human Immunodeficiency Virus Type 1 (HIV-1)

The history of the Human Immunodeficiency Virus Type 1 (HIV-1) started in the early 1980's as an increasing number of people with an undefined immunological dysfunction clinically appeared. At that time, an abnormally high incidence of rare types of cancer (like Kaposi Sarcoma) and opportunistic infections (such as *Pneumocystis pneumonia*) was reported, unusually amongst young men (181, 346, 496). This was in 1981, and it finally took three years until a retrovirus was identified to be the etiologic agent for these clinical manifestations, in 1982 named AIDS ("Acquired Immunodeficiency Syndrome"). Two research groups, those of Luc Montagnier and Robert Gallo, could isolate a cytopathic retrovirus from blood T-cells of AIDS patients (22, 168). Based on morphological and genetical similarities it was taxonomically placed into the lentivirus group and finally named human immunodeficiency virus type-1 (HIV-1) (430, 568). In 1986, a second type of HIV was discovered, termed HIV-2 (99). Infections of people with HIV-2 occur less frequently and are prevalently seen in western countries of Africa. Patients with HIV-2 show a slowed down pace of disease because of low viral loads and are mostly what is called a long term non progressor (LTNP). However, since HIV was identified it has been a subject of extensive research to open up novel therapeutic concepts. The disease is characterized by a decline of the T-cell number in the blood, which gradually weakens the immune system to defend the body against opportunistic pathogens and other malignancies (133, 182, 559). Finally, it is damaged in such seriousness, that AIDS is manifested in patients. Although there are therapies available that delay the progression of AIDS and improve the outcomes of the disease, at this point in time, a cure is still far away. In 2008 the World Health Organisation (WHO) together with the UN Programme on HIV/AIDS (UNAIDS) estimated that the epidemic includes 33 million people worldwide having HIV and that each year 2 million people die as a result of infection (551, 552). Current therapeutic strategies face various problems, e.g. cell reservoirs, in which HIV hides in a latent form (379, 405). Although highly active antiretroviral therapy (HAART) is able to deplete HIV below detectable levels, it cannot eliminate dormant virus that lies quietly within resting CD4⁺ memory T-cells (96, 98), macrophages (479) and

other cells types (172, 536). These sanctuaries can act as stepping stones for resurrection of virus particle production (97, 615). Furthermore, the high genetic variability of HIV promotes the induction of escape mutants, which successfully evade antiviral treatments (245, 311, 330, 395, 445, 578), revisiting the arduous challenge to develop new drugs.

I.2 Expression of the entire HIV-1 protein arsenal is controlled by alternative splicing

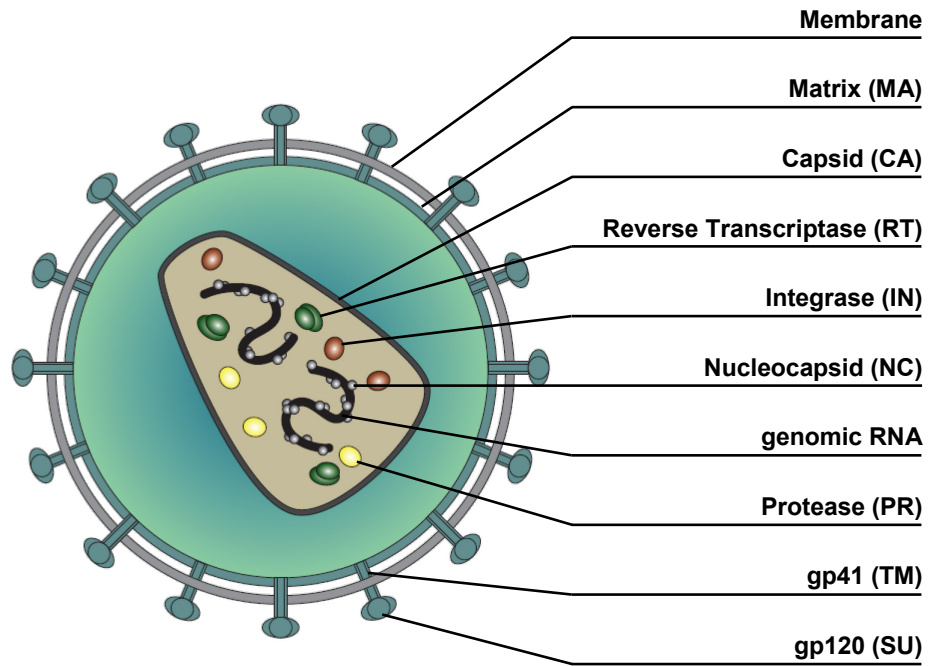
HIV causes a severe drop in the number of CD4⁺ T-lymphocytes within the body of an infected host. The chronic depletion of this immune relevant cell reservoir relentlessly affects the ability of the infected people to repel pathogenic attacks and certain cancers. The underlying mechanisms leading to programmed cell death of CD4⁺ T-cells are not well understood (112). However, it strikingly seems that most of those cells dying are not productively infected CD4⁺ T-cells but rather bystander cells, which commit suicide by sensing incomplete reverse transcripts (127) or soluble and membrane-bound viral proteins (112, 559). Therefore, viral proteins substantially contribute to the collapse of the immune system emphasizing the importance of a refined understanding of what are the principles beyond the regulation of their expression. Alternative splicing is the key mechanism ensuring (i) balanced expression of the entire viral proteome with accurately adjusted levels for each protein and (ii) a timely ordered viral gene expression guiding proper replication within the infected cell (267, 274, 365, 425).

I.2.1 HIV-1 replication cycle

Reverse transcription from a single stranded RNA genome back to cDNA typifies the class of retroviruses and accordingly, is a hallmark shared by HIV-1 (Fig.I-1A). Once the virus has entered the host cells cytoplasm, the viral RNA genome is converted into double-stranded DNA and released from the conical capsid throughout uncoating (Fig.I-1B). The resultant cDNA is associated with viral and cellular proteins to form a preintegration complex (PIC) that can actively be transported into the nucleus. This

also takes place in non-dividing cells (such as macrophages or microglial cells) in the presence of an intact nuclear envelope, which is a unique feature of the lentivirus subgroup. The imported DNA can be integrated as a provirus or circularized to episomes containing either one long terminal repeat (1-LTR circle) or two long terminal repeats (2-LTR circle). The extrachromosomal LTR circles are transcriptionally active, but cannot replicate and are therefore rapidly lost in dividing cells (63, 476, 502). Viral gene expression is under control of a promoter located within the 5'-long terminal repeat (5'-LTR). Independent on the integration site, the promoter is packaged into three nucleosomes, nuc-0,-1 and -2 (564). Nuc-1 is placed immediately downstream of the transcription start site and has to be removed for efficient transcription. However, cellular transcription factors can bind to a large nucleosome-free region between nuc-0 and nuc-1 and thereby facilitate low expression of Tat protein (246, 558, 564, 565). Tat binds to TAR (transactivation-responsive region), a stem loop formed at the 5'-end of all nascent HIV-1 transcripts. This results not only in the displacement of nuc-1 from the transcriptional start site (55) but also recruits pTEF-b (transcription elongation factor b), which hyperphosphorylates the C-terminal domain (CTD) of the RNA polymerase II (RNAPII) and increases transcriptional elongation (399). The RNAP II-mediated transcription coordinates the processing of the viral pre-mRNA, including 5'-end capping and polyadenylation. However, a single primary transcript is generated, which encodes the entire viral proteome. In order to allow the ordered and balanced expression of all HIV-1 proteins, the pre-mRNA undergoes extensive alternative splicing. The HIV-1 regulatory proteins Tat, Rev and Nef are expressed early in the virus life cycle and are required for efficient gene expression as well as infiltration of cellular pathways, which control innate immune defense (10, 119, 275, 420, 435, 587). After sufficient Rev protein has accumulated within the nucleus, the export of the intron-containing mRNAs is facilitated. The accessory proteins Vif, Vpr and Vpu and the glycoprotein precursor gp160 (Env) are translated from intron-containing mRNAs, while the unspliced mRNAs serve as a template for the expression of the structural and enzymatic precursor proteins Gag and Gag/Pol or as genomic RNA for progeny virus. Production of Gag and Gag/Pol drives the assembly of capsids at the plasma membrane each enclosing two copies of viral genomic RNA.

A



B

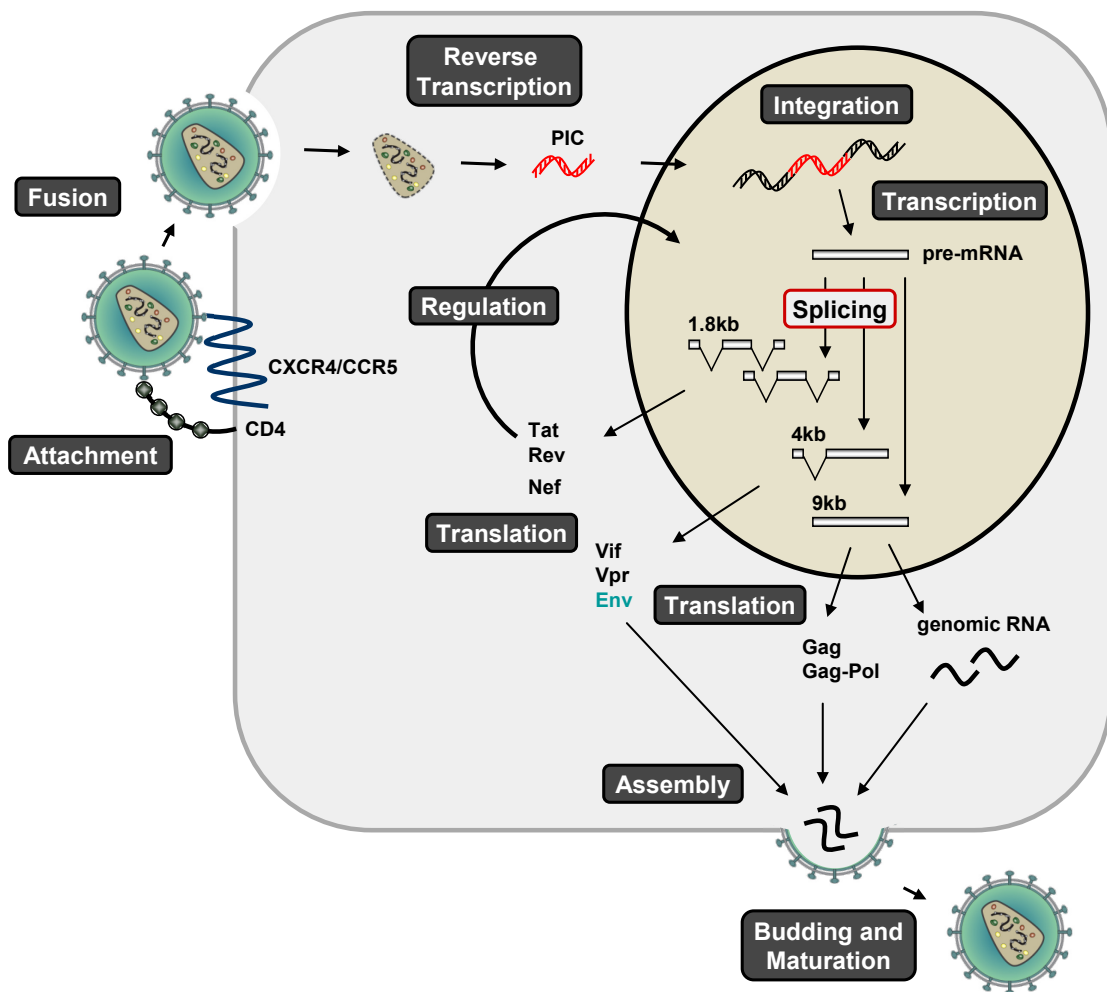


Fig.I-1: HIV-1 replication cycle [continued on next page]

During and after budding from the host cell the Gag and Gag/Pol precursors undergo proteolytic cleavage, which releases the individual structural and enzymatic proteins and matures the virion for a new round of infection.

Fig.I-1: continued

(A) HIV-1 virus particle. The virion is enveloped by a host-cell derived plasma membrane from which the viral glycoprotein sticks out. The viral matrix protein is attached to the plasma membrane by N-myristoylation. Two copies of the (+)-strand RNA genome are associated with nucleocapsid proteins and enclosed by the cone-shaped capsid. Further proteins (RT, IN, PR), which are necessary for viral infectivity, are embedded within the capsid.

(B) Infection starts with the attachment of the HIV-1 virion to the plasma membrane of a target cell by interaction of the viral surface glycoprotein gp120 with the CD4 receptor (114, 272, 358, 359). Changes in the conformation expose the V3 domain of gp120 for interaction with either chemokine receptor CXCR4 (150) or CCR5 (3, 124, 136). Binding to the coreceptor is accompanied by further conformational rearrangements, which releases the hydrophobic fusion peptide of the viral transmembrane glycoprotein gp41 and allow the fusion of the viral and the cellular membrane (82, 579). After entry of the viral capsid into the cytoplasm, the (+)-strand RNA genome is reverse transcribed and the preintegration complex (PIC) containing the proviral DNA actively enters the cells nucleus. The proviral DNA is then integrated into the cellular genome and transcribed into the viral pre-mRNA. Alternative splicing generates more than 40 different viral mRNA isoforms, which are separated into 3 major classes: multiply spliced 1.8kb-, intron-containing 4kb and unspliced 9kb sized mRNAs. The latter are translated into the structural and enzymatic proteins or serve as viral genome, which is enclosed into nascent capsids. After budding from the cellular plasma membrane the Gag and Gag-Pol precursor proteins are proteolytically processed into structural proteins and enzymes, releasing the mature virion.

I.2.2 Alternative splicing of the HIV-1 pre-mRNA allows the expression of the entire viral proteome

During LTR-driven transcription RNAPII synthesizes a single HIV-1 pre-mRNA. According to the rules of ribosomal scanning in which translation starts with the first AUG codon at the 5'-end of the mRNA, full-length transcripts result in the sole translation of the Gag and Gag/Pol open reading frames (ORFs) (282). However, viral replication critically depends on the expression of seven additional downstream located genes: Vif, Vpr, Tat, Rev, Nef, Vpu and Env (Fig.I-2A). To circumvent their translational inhibition, CAP-proximal AUGs are removed by alternative splicing from the primary transcript leading to different mRNAs, which are each specified (except for the bicistronic *vpu/env* mRNA) by the viral ORF nearest the 5'-end (Fig.I-2B). HIV-1 captures the cellular splicing machinery to generate more than 40 viral mRNA isoforms, which can be categorized into three distinct size classes: unspliced 9kb, intron-containing 4kb and intronless 1.8kb viral RNAs (167, 190, 425, 434, 467-469). The expression of the distinct viral mRNA classes takes place in a serially ordered, temporal fashion (267, 274, 365, 425). Upon the early phase of viral gene expression the HIV-1 pre-mRNA undergoes extensive splicing, which leads to the accumulation of multiply spliced 1.8kb mRNA species, encoding for the regulatory proteins Tat and Rev and the accessory proteins Nef and Vpr, respectively (425). The latter ones exploit cellular pathways to hide the infected cell from recognition by the host immune system and reprogram the environment for optimal virus replication (275, 294, 587), whereas Tat shuttles back into the nucleus to switch the RNAPII elongation complex into a processive one. However, Rev is essential for the onset of the late phase of viral gene expression (110, 218, 270, 335, 525, 586). An RNA element called Rev-responsive element (RRE) within the Env-coding sequence serves as a specific binding site for Rev (102, 113, 115, 206, 271, 611). Rev/RRE interactions target the intron-containing (4kb) and unspliced (9kb) viral mRNAs for CRM1 export receptor-mediated transport into the cytoplasm and thereby, bypasses cellular mechanisms, which normally retain these intron-containing transcripts within the nucleus (142, 149, 334). During the late phase of viral gene expression, the accessory and structural proteins Vif, Vpr, Vpu and gp160 (Env) are expressed from intron-containing viral mRNAs (4kb).

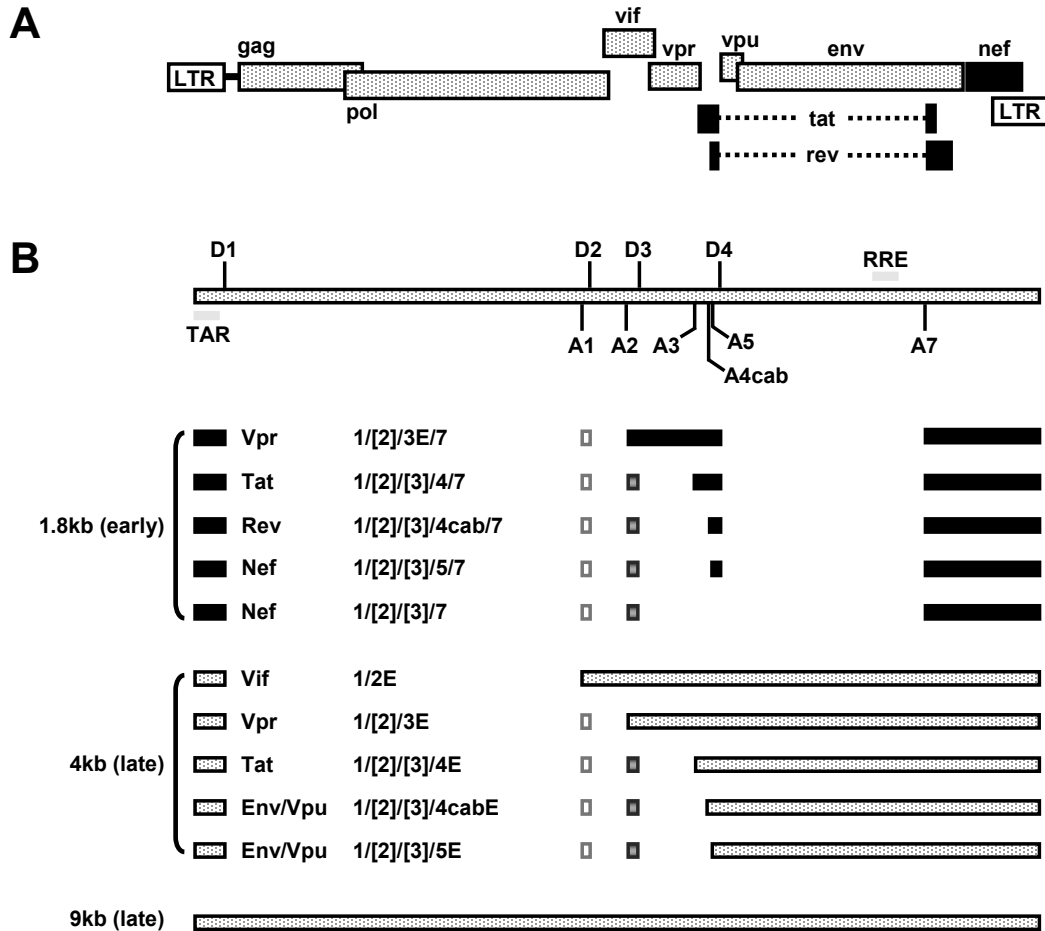


Fig.I-2: Alternative splicing of the HIV-1 pre-mRNA affords the expression of the entire viral proteome.

(A) Schematic diagram of the HIV-1 genome. One primary transcript is synthesized from the promotor located within the 5'-LTR. Nine open reading frames (ORFs) (Gag, Gag/Pol, Env, Vif, Vpr, Vpu, Tat, Rev and Nef) encode the 15 viral proteins. The HIV-1 genome is characterized by a compacted organisation in which the ORFs (with exception of Nef) overlap or punctuate each other. CAP-dependent ribosomal scanning along the unspliced 9kb pre-mRNA leads to exclusive expression of Gag and Gag/Pol protein precursors. ORFs are indicated by black (regulatory proteins; early expression) or scattered boxes (accessory, structural and enzymatic proteins; late expression). The long terminal repeats (LTRs) flanking the viral genome are shown as white boxes.

(B) The HIV-1 pre-mRNA undergoes extensive alternative splicing to unfurl the entire coding capacity and express the whole viral proteome. Regulated activation of the viral 5'splice sites (5'ss: D1 to D4) and 3'splice sites (3'ss: A1 to A7) produces an assortment of more than 40 mRNA isoforms from the HIV-1 primary transcript. The spliced mRNA isoforms are categorized due to their size as 1.8kb and 4kb mRNA class, respectively. The 4kb mRNA species are formed by splicing from the major 5'ss D1 to one out of the seven central 3'ss A1, A2, A3, A4cab or A5, and retention of the downstream intronic sequences. The 1.8kb mRNA species are additionally spliced from 5'ss D4 to 3'ss A7. Alternative inclusion of the non-coding leader exon 2 (white box) and exon 3 (dark grey box) generates respective isoforms of each of the mRNA species. [E: extended exons]

Furthermore, full-length viral mRNAs (9kb) are transported to the cytosol and are used for translation of the structural and enzymatic functions or as genomic RNA for progeny virions. Summarized, viral gene expression is defined by a Rev-dependent shift within the mRNA pool towards isoforms with increased intron content.

I.3 Mechanism of pre-mRNA splicing

Nearly all eukaryotic protein-coding genes display an interrupted structure of alternating exons and introns. The non-coding introns must be removed from the primary transcript (pre-mRNA) to generate a mature mRNA, which in turn can serve as template for translation. The process in which introns are excised and exons are joined together is named splicing. Splicing proceeds through two transesterification reactions, which are mediated by a macromolecular cell machinery termed spliceosome [reviewed in (567, 583)] (Fig.I-3). The spliceosome recognizes specific sequence elements, which define the exon/intron boundaries, the 5' splice site (5'ss; also referred to as splice donor, SD) and the 3' splice site (3'ss; also referred to as splice acceptor, SA).

I.3.1 Spliceosome assembly and catalysis

The major spliceosome is comprised of five uridine-rich small nuclear ribonucleoprotein particles (U snRNPs), i.e. U1, U2, U4/U6 and U5 snRNP and a broad inventory of non-snRNP proteins (59, 248). In addition to the U2-type spliceosome, which recognizes the vast majority of canonical GT-AG pre-mRNA introns, cells express a minor non-conventional spliceosome in which the small nuclear ribonucleoproteins (snRNPs) U1 and U2 are replaced by U11 and U12. Each snRNP is composed of one or two (U4/U6) U snRNA(s) associated with a ring of seven Sm proteins (referred to as B, D1, D2, D3, E, F and G) or like-Sm (Lsm) proteins in case of U6 snRNP, respectively, and a variable set of U snRNP particle specific proteins. Two models of pre-mRNA splicing have been proposed in the past: i) the spliceosome is a pre-assembled, “ready-to-go” complex, which is recruited to the pre-mRNA and has already acquired its splicing-competence and ii) the

spliceosome is sequentially assembled on the pre-mRNA for each splicing reaction through a series of intermediates (51, 228, 567, 583).

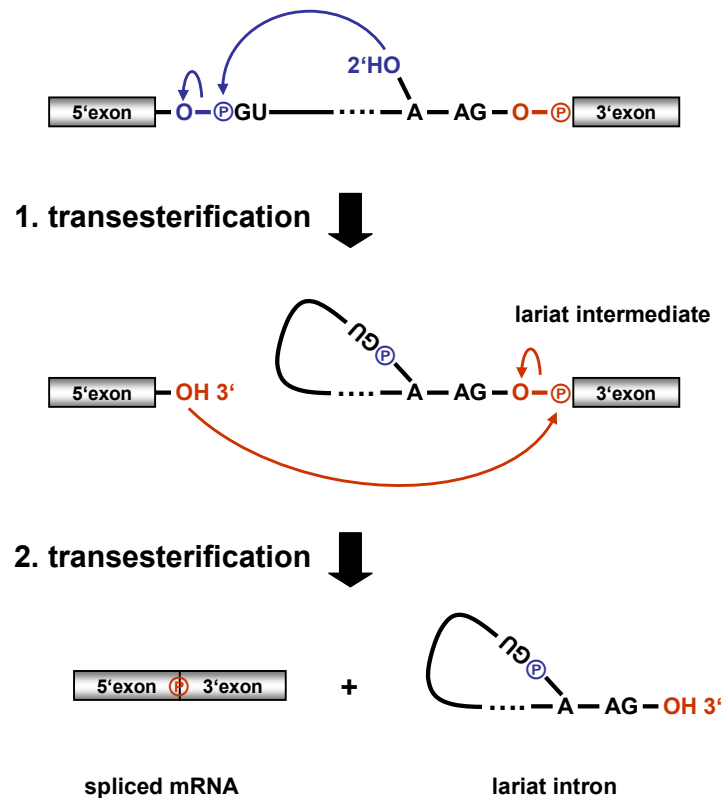


Fig.I-3: Splicing reaction occurs via two sequential transesterification steps.

In the first catalytic step, the oxygen of the 2'-hydroxyl group from the branch point adenosine (A) attacks the phosphorus atom, which attaches the intron to the 5'-exon (top). A nucleophilic substitution generates a lariat/3'-exon intermediate and the bond between the intron and the 5'-exon is dissolved. For the second transesterification reaction, the oxygen atom of the hydroxyl group of the 5'-exon forms a bond with the phosphorus atom that connects the intron and the 3'-exon (middle). This replaces the bond between the lariat intron and the 3'-exon and pieces both exons together to create the spliced mRNA (bottom).

Although a large 200S RNP particle named supraspliceosome could be isolated from HeLa cell extracts, which is able to catalyze RNA splicing *in vitro* (17, 515), it awaits further studies to elucidate whether *in vivo* the spliceosome may exist in a large pre-assembled complex (penta-snRNP or supraspliceosome) or assembles itself anew for each splicing reaction through a series of intermediates. However, since for a variety of reasons the latter is the more commonly accepted model, it will be described here in detail (Fig.I-4). Spliceosome formation is initiated by recognition of

RNA sequences at the exon/intron borders: the 5'ss, characterized by the canonical GU-dinucleotide and the 3'ss composed of the branch point sequence (BPS), a polypyrimidine tract of variable length (PPT), which resides between the branch point and the AG-dinucleotide. Spliceosome formation occurs in a sequential order and involves dynamic changes within the architecture of the complex (518, 567, 583). Early (E) complex formation initiates with the ATP-independent binding of the U1 snRNP to the pre-mRNA through base pairing between the 5'-end of the U1 snRNA and the 5'ss (366). In addition, the non-snRNP proteins SF1/mBBP (splicing factor 1/mammalian branch point binding protein) and both subunits of U2AF (U2 snRNP auxiliary factor) collectively bind to the BPS, PPT and the AG-dinucleotide of the 3'ss, respectively. Progression into pre-spliceosomal A complex requires ATP-dependent association of the U2 snRNP with the BPS. As a result of basepairing between the U2 snRNA and the BPS, the branch point adenosine is bulged out of the duplex and specified as nucleophile for the first transesterification (427). For chasing into B complex the preassembled U4/U6•U5 tri-snRNP joins the spliceosome. Theoretically, all components are now present to perform the first catalytic step. However, the B complex is still catalytically inactive and has to undergo ample conformational and compositional rearrangements to gain its splicing-competence (B* complex) (51, 567). During catalytic activation U1 and U4 are released from the spliceosome. Furthermore, U6 snRNA simultaneously interacts with the 5'ss and the U2 snRNA to bring both splice sites close together. The first catalytic step then converts B* into C complex and generates a lariat intermediate in which the first nucleotide of the 5'ss is attached to the branch point adenosine. Prior to the second transesterification and joining of both exons the catalytic center of the spliceosome is again extensively remodelled (278). As a consequence of the splicing reaction a protein complex is loaded 20-24nt upstream of the exon-exon junction onto the spliced mRNA (Exon Junction Complex; EJC), communicating with other cellular pathways to direct export, translation and quality control of the newly spliced mRNA (46, 85, 298-300, 497, 619). Once the second transesterification is completed the spliceosome falls apart, releasing the U snRNPs for participation in a new round of splicing.

The conservation of the splice sites in eukaryotic genes is relatively low and an aberrantly spliced mRNA with a +1 nucleotide shift in a splice site would suffice to generate a deleterious protein. However, the spliceosome removes even introns with a size up to 10^5 bases with a high degree of accuracy and especially with single

nucleotide precision (160). As a general principle, splice site nucleotides destined for the splicing reaction are repeatedly specified by alternate binding of spliceosomal components (505, 567). This recurrent handover from one splicing factor to the other over the course of spliceosome formation is predominantly controlled by ATPases and significantly contributes to splicing fidelity (60, 61, 350, 351, 597). Spliceosome assembly is essentially regulated by neighbouring sequences, so called splicing regulatory elements (SREs) (40, 347, 388, 574). Despite the outnumbering examples in which SREs control assembly of a functional spliceosome by enhancement or inhibition of initial splice site recognition, SREs were also shown to arrest transition into A complex (375, 477, 478), the spliceosomal assembly step, which coincides with functional splice site pairing (281, 308). A recent study gives new evidence for an interaction network across the intron, which coordinates functional communication between U1 and U2 snRNPs bound to the 5'ss and 3'ss, respectively (475). Herein, the SR-related proteins CAPER and Prp5 might attach to the U1 and U2 snRNPs, respectively, in order to bridge the splice sites by RS-domain mediated interactions. Therefore, commitment to the splicing pathway is determined by two serial steps, first recognition and then pairing of the splice sites.

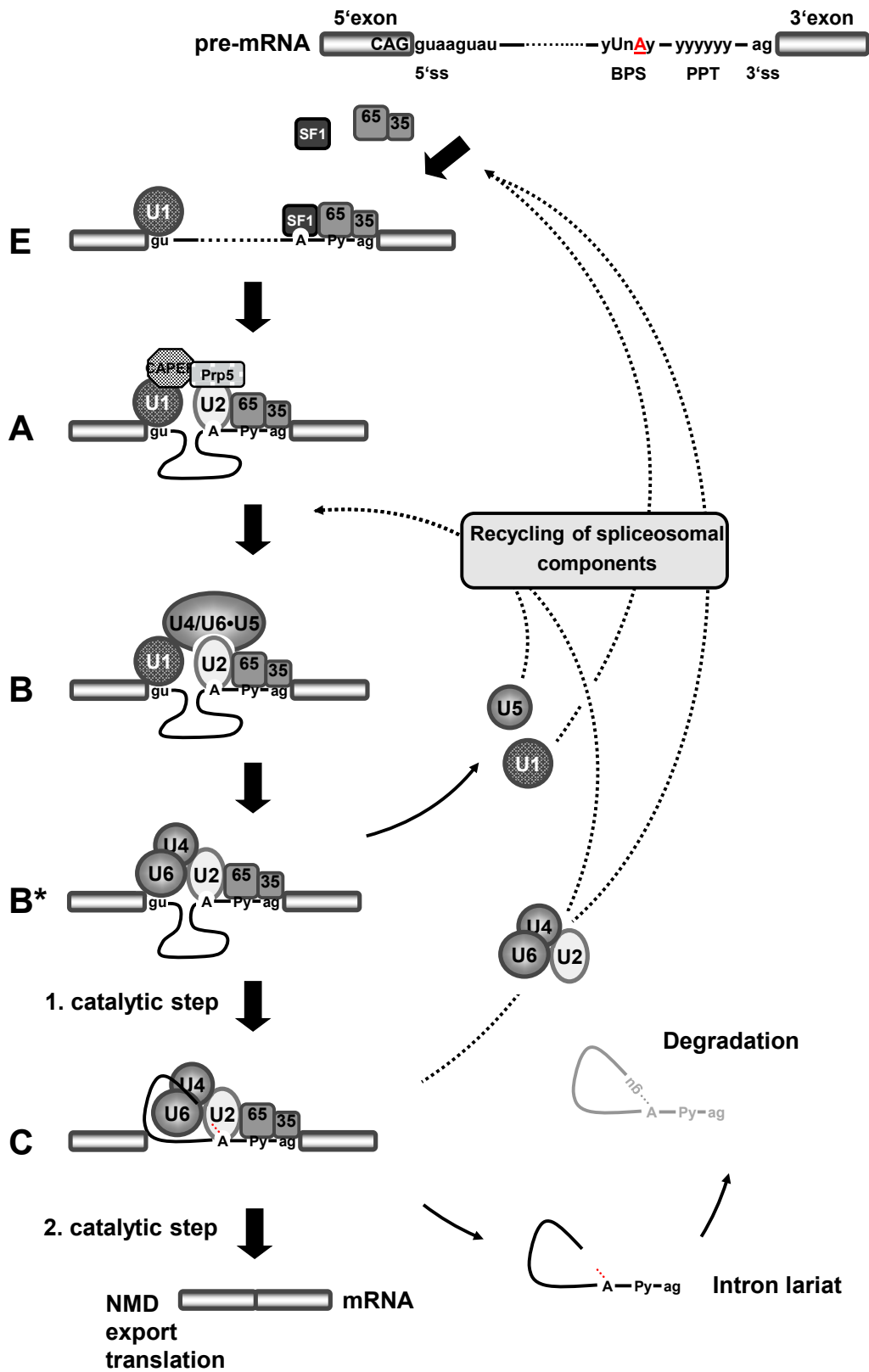


Fig.I-4: Spliceosome assembly [continued on next page]

Fig.I-4: continued

The exon/intron boundaries are delimited by the 5'ss and the 3'ss. The 5'ss is characterized by 11nt long sequence covering the 5'-exon/intron border with an invariant GU dinucleotide at position +1 and +2. The 3'ss contains three sequence elements; that are an invariant AG dinucleotide at positions -1 and -2, the preceding polypyrimidine tract (PPT) and the branch point sequence (BPS). Spliceosome assembly initiates with recognition of the 5'ss by the U1 snRNP (U1). Furthermore, the non-spliceosomal proteins SF1/mBBP and both subunits of U2AF (65 and 35) cooperatively bind to the BPS, the PPT and the AG-dinucleotide of the 3'ss, respectively; thereby leading to early (E) complex formation. The E complex is then chased into A complex by ATP-dependent binding of the U2 snRNP (U2), which releases SF1/mBBP from the BPS. Functional splice site pairing between the 5'ss and the 3'ss is established at this step, including interactions between CAPER and Prp5 (475). Addition of the pre-formed U4/U6•U5 tri-snRNP (U4/U6•U5) then progresses the A into B complex. Extensive structural remodelling releases the U1 and U4 snRNP from the spliceosome and renders the B complex active for the first transesterification reaction (B* complex). The first catalytic step then converts B* into C complex, thereby generating a lariat intermediate in which the first nucleotide of the 5'ss is attached to the branch point adenosine. Prior to the second transesterification and joining of both exons the catalytic center of the spliceosome is again extensively remodelled. After the second transesterification reaction the spliceosome falls apart and releases the spliced mRNA. The lariat intron is targeted for destruction by debranching enzymes and exonucleases, while the spliceosomal components are recycled to participate in a new round of splicing. [A: branch point adenosine, Y: pyrimidine; U1: U1 snRNP; 65: U2AF65; 35: U2AF35; U2: U2 snRNP; U4/U6•U5: U4/U6•U5 tri-snRNP; NMD: nonsense mediated decay]

I.3.2 Exon and intron definition complexes

Splice site pairing across the intron is an immanent feature of the splicing process. Most of the human genes are alternatively spliced, producing various mRNAs by selecting different combinations of exons from the same primary transcript (388, 406, 574). Therefore, it is fundamental that the exons are precisely recognized. In contrast to their human counterparts, yeast introns are much shorter and occur in comparatively low numbers. As a result, the recognition of exons is very likely established through interactions spanning the intron in a process called intron definition (159, 437). However, in vertebrates exons are relatively small (50-250bp) with an median size of 123nt and separated by long introns (1458nt) (464), so that exon definition seems to be the preferred pathway of splice site pairing (Fig.I-5). The exon definition model was first proposed following the observation that addition of a 5'ss to the second exon of a two exon *in vitro* splicing substrate positively influenced the rate of upstream intron removal (28, 433).

Further support came from studies, which reported entire exon-skipping for mutations of just one of the splice sites, reinforcing the notion of communication between the two ends of an exon. The molecular bridge initiates with the binding of the U1 snRNP to the 5'ss, which in turn directs the upstream assembly of U2AF and subsequently U2 snRNP at the 3'ss (215). Splicing regulatory elements (SREs) placed within the exon stabilise the cross-exon interactions and promote the connection of the U1 and the U2 snRNP at the opposite ends (49, 215). Short distances alleviate end-to-end bridging so that cross-exon interactions are the favourable mode of splice site pairing for vertebrates. However, because the splicing reaction occurs across an intron, the exon definition complex must be switched into a cross-intron complex where the 3'ss contacts the 5'ss of the upstream exon (45, 219, 477). *In vitro* purification experiments revealed that the U4/U6•U5 tri-snRNP is already attached to exon definition complexes. These can then be promptly converted into splicing-competent B complexes when a 5'ss-containing RNA oligo is added to the reaction, indicating a flip from exon- to intron-definition complex (465).

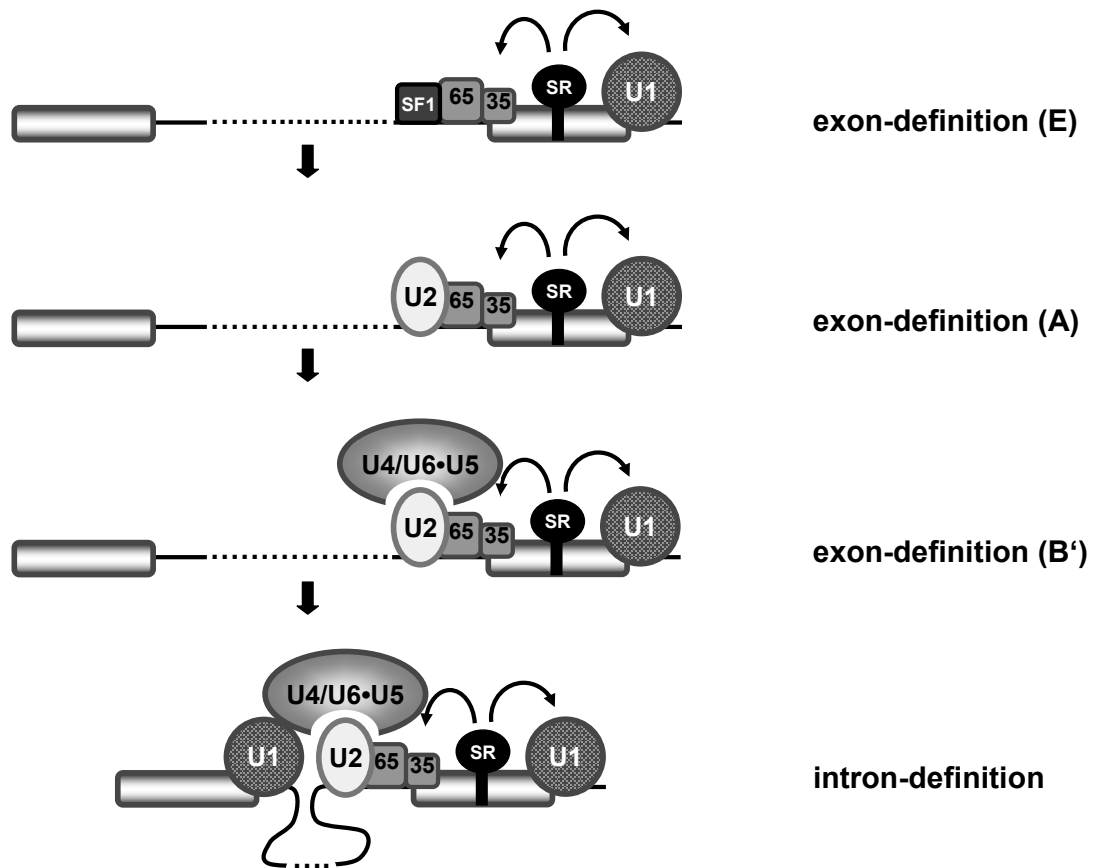


Fig.I-5: Exon- and intron-definition

Since mammalian exons are relatively small compared to their flanking introns, exon definition is supposed to precede splice site pairing across the intron. Exonic SR proteins stabilize general spliceosomal components at the flanking splice sites and thereby enhance communication between the 5'- and the 3'-end of the exon. Finally, exon definition complexes are chased into intron-definition complexes, which elicit the splicing reaction. [U1: U1 snRNP; 65: U2AF65; 35: U2AF35; U2: U2 snRNP; U4/U6•U5: U4/U6•U5 tri-snRNP]

I.3.3 Splice site recognition

Whereas yeast splice sites are well conserved, those found in metazoans are highly degenerated not rarely found to significantly deviate from the consensus motifs (59). However, there are many cases of metazoan introns with weak splice sites, which are spliced efficiently or even more efficiently than those containing splice sites with stronger matches to the consensus sequences (160). In fact, a bona-fide splice site is not only defined by its conservation and intrinsic strength alone but also by its sequence environment (574). Neighbouring SREs can promote the use of a splice site with a poor match to the consensus motif by the splicing apparatus or – in contrast – inhibit the recognition of a splice site, which at the first sight would be predicted to be used more efficiently. Accordingly, the recognition of a splice site can be considered as the net result of its intrinsic strength and the sequence neighbourhood in which it is embedded (21, 574).

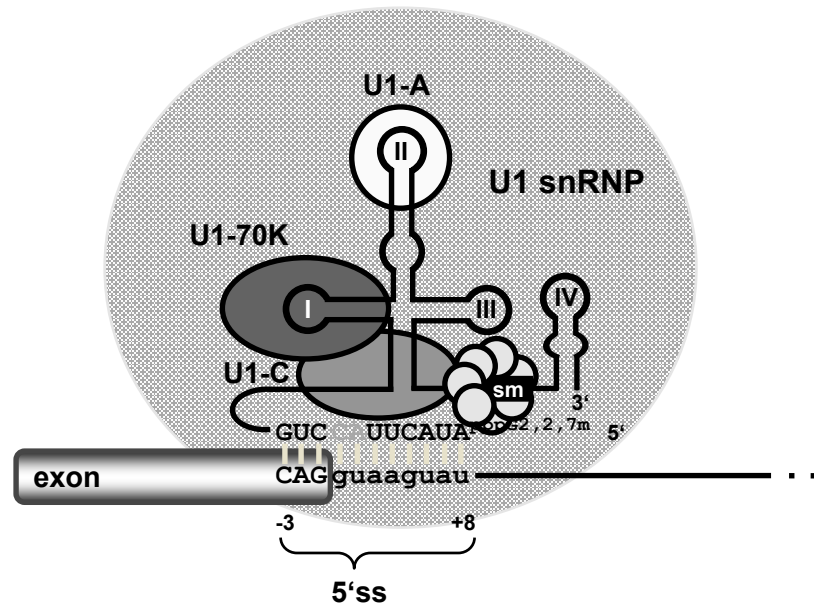
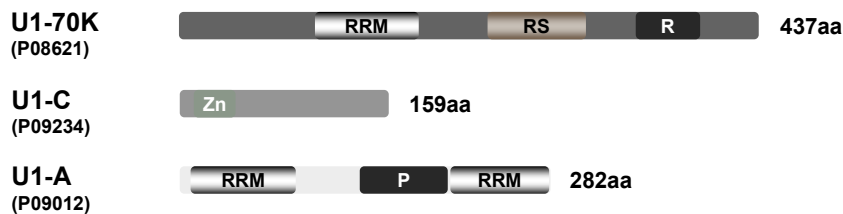
I.3.3.1 5'ss recognition

The splicing apparatus precisely identifies 5'ss among a pool of pseudo-splice sites to define the correct exon/intron junctions. True 5'ss consist of an 11nt long sequence marked by an invariant GU-dinucleotide at the position +1 and +2 with regard to the exon/intron border. Recognition of the 5'ss by the U1 snRNP is the very first step of spliceosome assembly and significantly contributes to the commitment of the pre-mRNA to the splicing pathway (Fig.I-6A). The canonical U1 snRNP is composed of the U1 snRNA (164nt) and ten proteins among, which are seven Sm proteins (E, F, G, B, D1, D2 and D3) and three U1-specific proteins (U1-70K, U1-A and U1-C) (50, 385) (Fig.I-6B). The free 5'-end of the U1 snRNA undergoes basepairing with the 5'ss sequence to form an RNA duplex, which theoretically can occlude the positions -3 to -8 relative to the exon/intron junction (626). As a result, the complementarity to the U1 snRNA constitutes the intrinsic strength of a 5'ss. Although sequence comparison of human canonical 5'ss revealed certain preferences for individual nucleotides based on their relative occurrence at each residue and allowed to derive a consensus sequence, it was experimentally proven that a stretch of six matching nucleotides suffices to efficiently direct the U1 snRNP

to a 5'ss (162, 253). A comparative analysis of 46,308 annotated human 5'ss sequences later on showed that 5'ss sequence logos can be distinguished based on whether continuous basepairing predominantly involves exonic, centred or intronic nucleotides of the 5'ss (202).

Resolution of the U1 snRNP structure uncovered the relative arrangement of the U1 snRNA and the ten associated proteins (421). Herein, U1 snRNA forms four stem loops (SLI-IV) and serves as a platform for the binding of the seven Sm proteins and the three U1-specific proteins U1-70K, U1-A and U1-C. The seven Sm proteins assemble around the Sm site nucleotides (AAUUUGUGG) between stem loops III and IV to form a heptameric ring (251, 577). U1-70K and U1-A bind to stem loops I (SLI) and II (SLII) of the U1 snRNA, respectively, which is mediated by their RNA recognition motifs (RRMs). Within the U1 snRNP particle U1-70K wraps its N-terminal peptide underneath the Sm protein complex to finally form an interfacial surface with the Sm protein D3 that is the basis for the incorporation of U1-C (383, 421). The resolution map revealed that U1-C localizes at a position in close proximity to the 5'-end of the U1 snRNA (421, 519). Furthermore it showed that U1-C interacts with the minor groove formed by the RNA duplex between U1 and the 5'ss. It has been suggested that U1-C stabilizes the U1/5'ss interactions and thereby, facilitate E complex formation (88, 208, 584). In consistence with this notion, loss of Prp28, an ATPase, which unwinds the RNA duplex to release the 5'ss for pairing with the U6 snRNA during spliceosome assembly, is compensated by mutations in U1-C detaching its interaction with the 5'ss/U1 RNA duplex (88). However, the long-standing model that U1-C is essential for spliceosome assembly derived from yeast and *in vitro* experiments (538) is difficult to reconcile with more recent genomewide studies that could not find evidence for a global splicing defect in mutant zebrafish (441). Nevertheless, the data suggested a striking influence on a significant number of alternatively spliced target genes.

Since human 5'ss are highly degenerated, their recognition exclusively by RNA-RNA interactions with the U1 snRNA cannot be the only determinant guiding the U1 snRNP to the 5'ss. SREs in the immediate neighbourhood were found to play a pivotal role for the activation of an adjacent 5'ss. SREs can be placed within the upstream exon or a downstream intron of the 5'ss to enhance U1 snRNP recruitment and are therefore classically termed exonic splicing enhancer (ESE) or intronic splicing enhancer (ISE).

A**B****Fig.I-6: 5'ss recognition**

(A) Spliceosome assembly starts with recognition of the 11nt long 5'ss by the U1 snRNP. The 165nt long U1 snRNA folds into a cloverleaf-like structure and is complexed with ten proteins. Seven Sm proteins (E, F, G, B, D1, D2 and D3) assemble around the Sm binding site between stem loops III and IV, thereby forming a heptameric ring. The U1 snRNP-specific proteins U1-70K and U1-A bind to stem loop I and II, respectively, whereas U1-C interacts with U1-70K for incorporation.

(B) Schematic diagram of the protein domain organization of the U1 snRNP-specific proteins U1-70K, U1-C and U1-A. Domain positions were obtained from the Swiss-Prot database (www.uniprot.org). RRM: RNA recognition motif; RS: arginine-serine rich; R: arginine-rich; Zn: zinc-finger; P: proline-rich]

The well-described pathways by which enhancer bound splicing proteins promote U1 snRNP binding to the 5'ss include interactions of SR proteins or the T cell-restricted cellular antigen-1 (TIA-1) with the U1-specific proteins U1-70K (91, 276, 589) and U1-C (157), respectively. TIA-1 was shown to assist U1 snRNP recruitment through specific binding to uridine (U)-rich elements immediately downstream of 5'ss (123, 157). U-rich elements were found to be particularly overrepresented within genes that

changed their alternative splicing pattern in U1-C deficient zebrafish (441), supporting a crucial role for U1-C in the regulation of alternative splicing. Summarized, recognition of a 5'ss by the U1 snRNP cannot simply be predicted by calculating the basepairing potential of a target 5'ss, since there is a strong dependence on positive regulatory elements in their proximity. Weak splice site can nevertheless be activated through the assistance of enhancer elements. However, 5'ss recognition in terms of U1 snRNP binding to the pre-mRNA does not necessarily imply subsequent usage or successful commitment to the splicing pathway. There are few examples of “dead-end” U1 snRNPs deprived of their capability to promote progression into a functional spliceosome (375, 478, 548). Moreover, the population of U1 snRNPs seems to be heterogeneous, introducing potential regulatory steps subsequent to initial 5'ss recognition. For instance, the U1 snRNP can incorporate two different isoforms of U1-70K of which isoform 2 is lacking one phosphorylation site (Ser 226), which enhances the interaction with U1-C but decreases that with Sm-B/B', respectively (209). It was proposed that the phosphorylation state of U1-70K might alter the network of interactions between U1-70K, Sm-B/B' and U1-C and thereby, influence the association with other splicing factors, importantly that with the U2 snRNP during splice site pairing (132, 209). Furthermore, *Drosophila melanogaster* was shown to express two distinct U1 snRNP particles of which one lacked the U1-A homolog SNF and an associated kinase activity rendering U1-70K less phosphorylated (290). Although both were functional in the context of an *in vitro* splicing reconstitution assay, distinct functions in alternative splicing cannot be excluded. Another study identified several U1-like snRNAs defined by sequence variations within the free 5'-end. Sequence alignments with the canonical U1 snRNA suggested a property to assemble functional U1 snRNPs (289). Whether these U1-like snRNAs have specialized functions in alternative splicing needs to be clarified. However, together these studies indicate that cells may express a repertoire of functionally distinct U1 snRNPs.

I.3.3.2 3'ss recognition

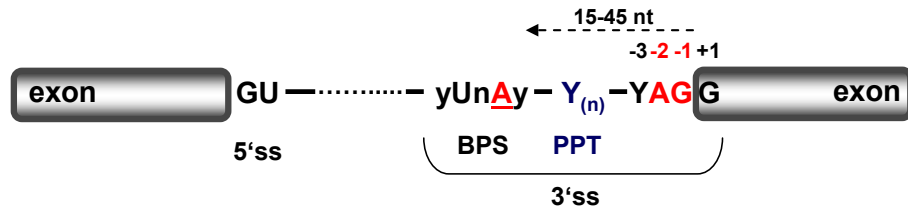
A set of three *cis*-regulatory sequences builds up the 3'ss and thereby defines the downstream exon/intron boundary of canonical introns. The branch point sequence (BPS) is sitting most upstream, followed by a polypyrimidine tract (PPT) of variable length and finally the invariant AG-dinucleotide, which delimits the exon/intron junction (Fig.I-7A). The efficiency by which the 3'ss is recognized through components of the splicing machinery is supposed to critically depend on the functional strength of the PPT and the complementarity between the BPS and the U2 snRNA (106, 254). Furthermore, the nucleotides -3 and +1 flanking the AG-dinucleotide at the downstream exon/intron junction have been implicated to affect 3'ss recognition under certain conditions. A pyrimidine at position -3 (Y-3) is preferred by the splicing machinery to specify the 3'ss and was shown to improve the AG-dinucleotide recognition for the second catalytic step of the splicing reaction (103, 323, 503). Additionally, a guanosine (G) at the first position of the exon is favoured, finally defining the consensus sequence YAG/G. However, mutations from G to C in position -1 revealed that the requirement for G+1 depends on the length and the pyrimidine content of the upstream located PPT, suggesting that the position -1 is only crucial for a specific subset of introns (165, 201). The human branch point consensus is an extremely variable sequence (yUnAy) with the branch point adenosine (BP A) being the fourth nucleotide (169). Most branch point sequences are situated in a distance of 15-45 nt upstream relative to the 3'-end of the intron. However, BPS were also found deeper or multiple times within the intron (59, 107, 178, 195, 389, 486). The introns with more than one BPS undergo unusual splicing events such as the recursive mechanism by which a large intron in *Drosophila* is consecutively removed in a piece-by-piece fashion (62, 486). While yeast introns have a nearly invariant "UACUAAC" sequence, the human BPS is very short and highly degenerated, which complicates to identify candidate BPS based on sequence alone (169, 277, 409). This indicates that recognition of the BPS is also dictated by other features of the 3'ss and that the BPS cannot be considered separately.

Within the PPT uridines are preferred over cytidines and it was demonstrated that a PPT composed of 11 continuous uridines is highly efficient in splicing (106, 376). For the respective introns with a functionally strong PPT, the early AG-dinucleotide recognition is dispensable. On the other hand, short or degenerated PPTs require the

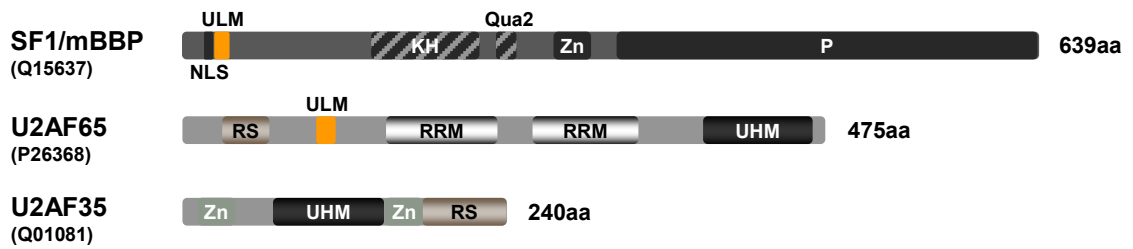
simultaneous recognition of the AG-dinucleotide and are more vulnerable to mutations at positions Y-3 or G+1 (431). Based on these observations introns have been classified in AG-dependent and AG-independent (192, 193, 590). However, human introns display a large diversity of PPT sequences, which differ in nucleotide composition and length, revisiting the difficulty to predict 3'ss recognition by solely regarding one element without the others.

Combinatorial control of the 3'ss recognition is predominantly provided by synergistic binding of general splicing factors to the three elements. Weakening or improving the recruitment of one of these splicing factors can determine the need for the other (107, 169). Splicing factor 1/mammalian branch point binding protein (SF1/mBBP) binds to the BPS (30, 314), while the U2 snRNP auxiliary factor (U2AF) 65-kDa and 35-kDa subunits recognizes the PPT and the AG-dinucleotide, respectively (361, 448, 590, 609, 610, 627) (Fig.I-7B-C). Strong PPTs were shown to functionally compensate for the absence of an AG-dinucleotide. Conclusively, recognition of the AG-dinucleotide by U2AF35 plays a crucial role for some but not all introns during early spliceosome assembly (59, 192). However, AG-dependent introns with short, divergent PPTs can only be spliced efficiently when both U2AF subunits are present within the reaction (192, 401, 590, 627, 629). Although a subset of introns does not require the presence of an AG-dinucleotide, absence of U2AF35 impairs viability of several model organisms such as the fruit fly *Drosophila melanogaster* (255, 447), the fission yeast *Schizosaccharomyces pombe* (575, 576, 580), the nematode *Caenorhabditis elegans* (627) or the zebrafish *Danio rerio* (177). In the budding yeast *Saccharomyces cerevisiae* there has been no orthologue of U2AF35 described so far. Within vertebrate cells, alternative splicing generates two distinct U2AF35 isoforms, U2AF35a and U2AF35b, which differ in only seven amino acids from one another (402). The minor isoform U2AF35b shows 9-18 fold lower expression levels in various human tissues, but could functionally replace U2AF35a in biochemical complementation assays. Depletion of each U2AF35 isoform in HeLa cells decreases cells viability by changes in the alternative splicing of cell cycle associated genes (403). However, the overall effects on the cells are relatively moderate compared with those following loss of other constitutive splicing factors, supporting surrogate mechanisms by which the requirement for U2AF35 can be bypassed or at least mitigated. Furthermore, alterations in the relative ratio between U2AF35a and U2AF35b may affect the specificity for 3'ss selection.

A



B



C

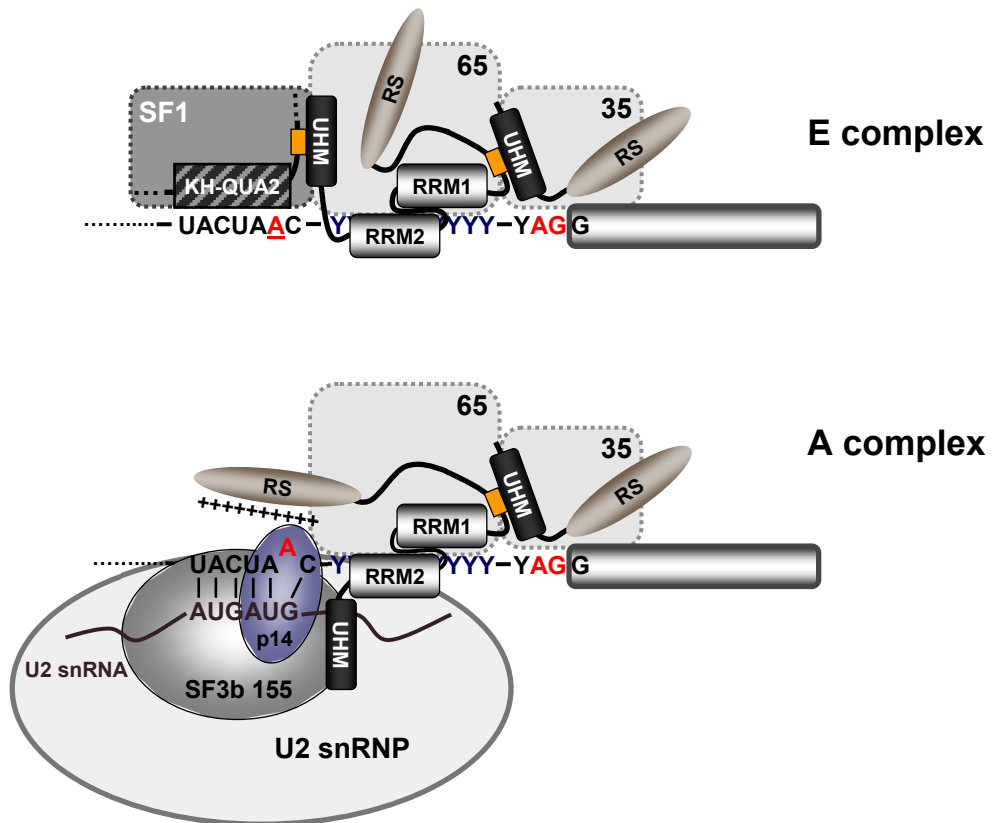


Fig.I-7: 3'ss recognition [continued on next page]

(A) The tripartite metazoan 3'ss contains the AG-dinucleotide (red), the PPT (blue) and the BPS. The conserved branch point adenosine is indicated (red).

U2AF65 binding to the PPT is mediated by two RNA recognition motif (RRM) domains (RRM1-2) (231, 495, 542, 610). A third RRM-like domain located in the C-terminal part and called U2AF homology motif (UHM) functions as binding domain for the interaction with the U2AF ligand motif (ULM) of mBBP/SF1 and other splicing factors (184, 221, 265, 428, 473).

U2AF65 contains a ULM domain of its own, which is essential for heterodimerisation with U2AF35 (446). An intrinsically unstructured arginine-serine (RS) rich domain resides in the N-terminal part of U2AF65 and contacts the BP when U2AF65 is bound to the PPT. The RS domain promotes stable RNA duplex formation between the U2 snRNA and the BPS by neutralizing the negatively charged phosphate backbone of the approaching U2 snRNA (481, 554). U2AF65 can adopt two alternative conformations (326). Conformational transition from closed to open state releases domain surfaces of RRM1, which are otherwise occluded by intramolecular interactions and promote stable association with the PPT. The open conformation may also switch on interactions between U2AF65 and other splicing components, which parallel functional spliceosome assembly (326). A model was proposed in which the U2AF65 shift towards the open state is triggered by the quality of the PPT, preventing promiscuous spliceosome assembly at short and unspecific pyrimidine sequences randomly found within the pre-mRNA.

Fig.I-7: continued

(B) Both subunits of U2AF and SF1/mBBP cooperatively bind to the 3'ss (E complex, E). Herein, UHM:ULM interactions between SF1/mBBP and U2AF65 or U2AF65 and U2AF35, respectively, stabilize binding of the protein complex to the 3'ss. Upon transition into A complex SF1/mBBP is displaced from the BPS by the U2 snRNP. The U2-snRNP specific protein p14 directly contacts the branch point sequence. Basepairing interactions between the U2 snRNA and the BPS are stabilised by the arginine-serine (RS) domain of U2AF65. Moreover, the UHM of U2AF65 is released for interactions with SF3b155. [UHM: U2AF homology motif; ULM:U2AF ligand motif; U2: U2 snRNP] [adapted to (567)]. **(C)** Schematic diagram of the protein domain organization of SF1/mBBP and both subunits of U2AF. Domain positions were obtained from the Swiss-Prot database (www.uniprot.org). [BPS: branch point sequence; PPT: polypyrimidine tract; A: branch point adenosine; Y: pyrimidine; UHM: U2AF homology motif; ULM: U2AF ligand motif; KH: K homology; QUA2: Quaking homology 2; RRM: RNA recognition motif; P: proline-rich; Zn: zinc finger]

Several studies provide a body of evidence that there may exist alternative pathways of 3'ss recognition to overcome the requirement for U2AF65. For example it was shown that the structurally related (U) binding splicing factor 60kDa (PUF60) and the SR protein SRSF2 can functionally replace U2AF65 *in vitro* (203, 327).

SF1/mBBP was found to be essential for mammalian cell viability (537) and deficiency has been linked to increased cancer susceptibility and other diseases (78, 490, 491). SF1/mBBP recognizes the 3'ss along with U2AF65 in the early stages of spliceosome assembly (29, 30). A central K homology (KH) motif and a Quaking homology 2 (QUA2) region guide SF1/mBBP to the BPS (314, 428). Furthermore, interactions between the UHM domain of U2AF65 and the ULM of mBBP/SF1 facilitate cooperative binding to the 3'ss (240, 495). Specific serines located within the region spacing the KH-QUA2 region and the ULM were shown to be highly conserved from yeast to humans and undergo extensive phosphorylation to either inhibit (Ser20) or stimulate (Ser80/Ser82) the interaction between SF1/mBBP and U2AF65, respectively (336, 573). Despite synergistic recognition of the 3'ss by SF1/mBBP and U2AF65, SF1/mBBP can also bind separately to the intron (264). However, the onset of spliceosome assembly requires the coincidental association of U2AF65 with the PPT. As already described for U2AF65 and U2AF35, SF1/mBBP is essential for RNA splicing of many but not all genes indicating compensatory pathways of spliceosome assembly by which absence of SF1/mBBP can be mitigated (409, 449, 537).

Chasing into pre-spliceosomal A complex correlates with stable U2 snRNP binding and displacement of SF1/mBBP from the BPS (384, 408, 567, 588, 625) (Fig.I-7C). Conformational rearrangements mediated by the ATPases UAP56 and Prp5 precede basepairing of the U2 snRNA with the BPS (154, 214, 414, 415, 596, 612). Furthermore the two U2 snRNP associated protein subcomplexes SF3a and SF3b and the RS domain of U2AF65 promote stable U2 snRNP binding (183, 184, 481, 554, 567).

I.3.4 Alternative splicing and *cis*-regulatory elements

Alternative splicing allows the regulated activation of splice sites and thereby the flexible production of multiple mRNA and protein isoforms from one primary transcript. The majority of human genes is estimated to undergo alternative splicing (406). This effectively expands the proteomic output on basis of a comparatively finite number of protein coding genes (339, 388). Protein isoforms expressed from one gene are often defined by opposing physiological functions so that even small changes in the relative expression ratio can have a huge outcome for the cell (347, 470). Furthermore, many human diseases have been linked to erroneous pre-mRNA splicing (64, 75, 134, 398, 570) so that much interest has been generated to understand the mechanisms by which alternative splicing is regulated. Modes of alternative splicing range from entire exclusion of an exon, alternative 5' or 3'ss selection to the retention of an intron (Fig.I-8).

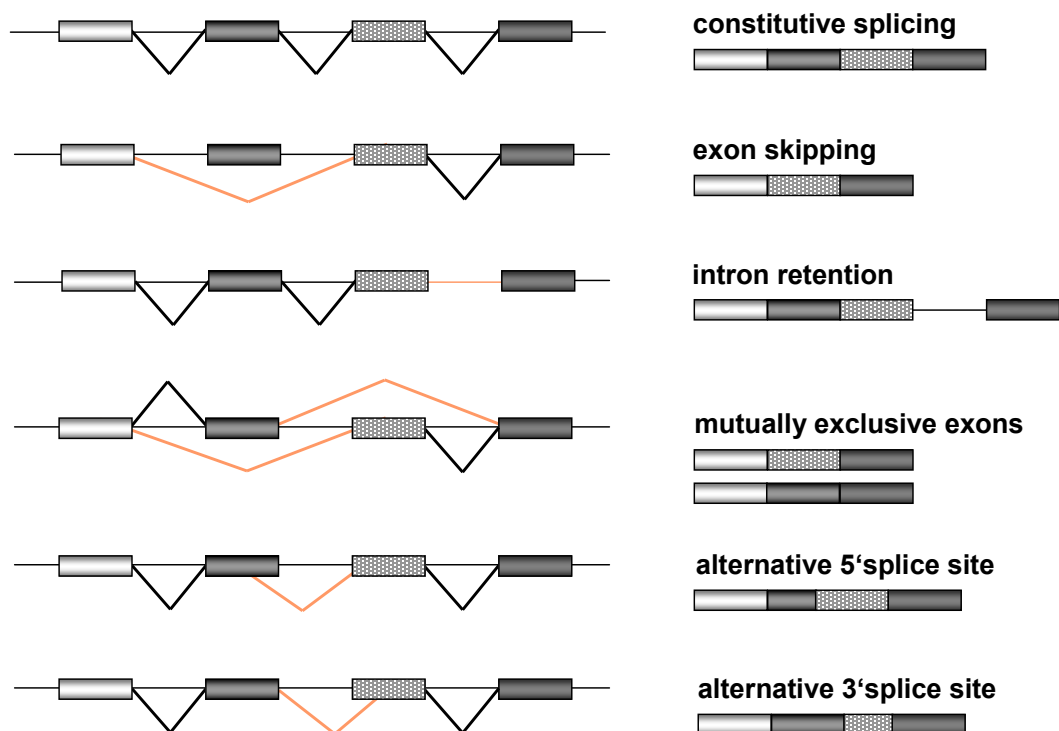


Fig.I-8: Modes of Alternative Splicing

Schematic diagram of the different modes of alternative splicing. These range from inclusion (constitutive splicing) or exclusion of internal exons (exon skipping) to the retention of an intron (intron retention), mutually exclusive recognition of middle exons (mutually exclusive exons) and use of alternative 5'ss (Alternative 5'splice site) and 3'ss (alternative 3'splice site).

Therefore, alternative splicing confers an immanent plasticity to each pre-mRNA, allowing its processing into a great diversity of different transcript isoforms. The splicing machinery attains its instructions by what is called the “splicing code” in which the information where to “apply the scissor” is hidden in the coordinated outcome of the splice site itself, neighbouring sequences and RNA secondary structures (21, 484, 574). Human splice sites often loosely match the consensus sequence, but are nonetheless faithfully recognized by the spliceosome. By contrast, cryptic splice sites are frequently found to be even stronger than the true ones, but are nonetheless not used by the splicing machinery. This raises the question of how the spliceosome handles the arduous challenge to discriminate between false and true splice sites, when their sequence alone is clearly not sufficient (86). The gap between splice site sequence and its recognition is primarily bridged by auxiliary elements located within the exons or flanking introns that specify bona-fide splice sites. These splicing regulatory elements (SREs) are bound by nuclear RNA binding proteins and can enhance or repress recognition of a proximal splice site (310, 347, 561, 574). Often multiple SREs are interconnected to combinatorially control splice site use immediately converting intracellular concentrations of their protein binding partners into appropriate expression levels of the different mRNA isoforms (37, 196, 504). Transcriptome-wide analyses could even reveal coordinated alternative splicing programs in response to certain tissue-specific splicing factors such as Nova or Fox1/2, which specifically shape cellular functions (550, 603, 614).

SREs can be classified on basis of their relative location and activity into intronic/exonic splicing enhancers (ESE/ISE) or exonic/intronic splicing silencers (ESS/ISS) (Fig.I-9). SREs were shown not to be exclusively involved in alternative splicing events but also to play a critical role for constitutively used splice sites (461, 617). The importance of SREs is underlined by a continuously growing body of examples that could relate their disruption by mutations to faulty pre-mRNA splicing and disease (75, 404, 506). However, SREs are defined by their short (5-10nt) and relatively degenerated consensus sequences that are specifically recognized by nuclear splicing factors. The mechanisms by which these splicing factors act throughout splicing activation and repression are in detail described in chapters I.3.5.1.1 and I.3.5.2.1.

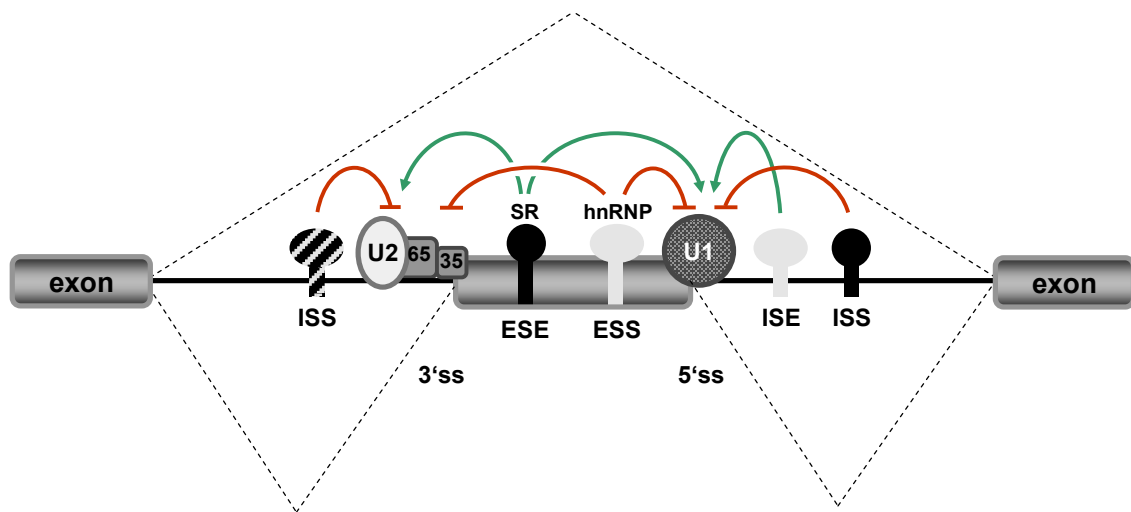


Fig.I-9: Alternative splicing is controlled by splicing regulatory elements (SREs) in proximal locations relative to the regulated exon.

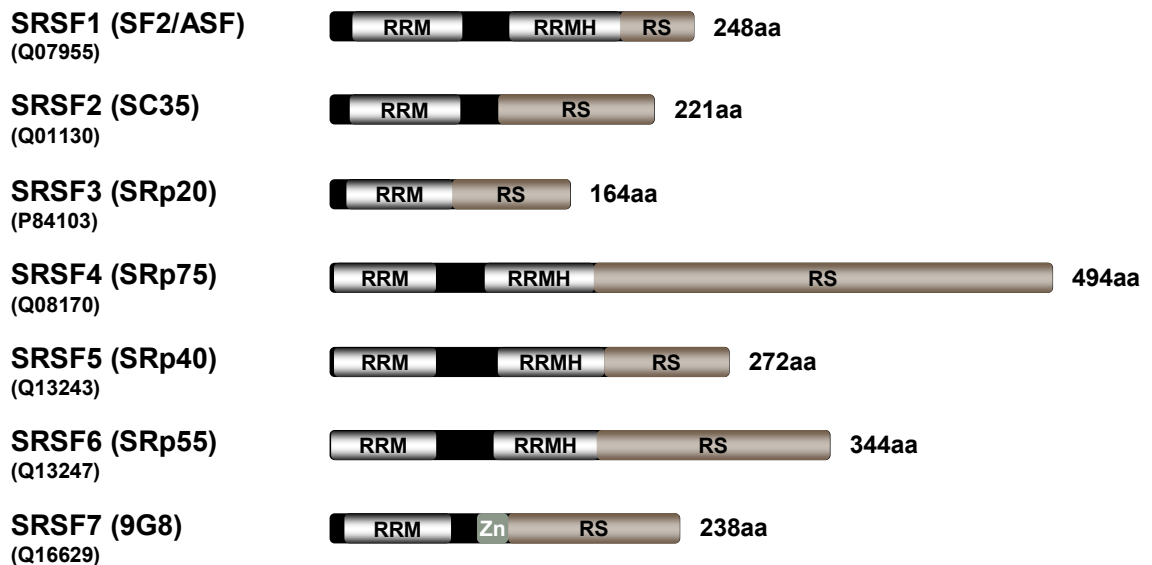
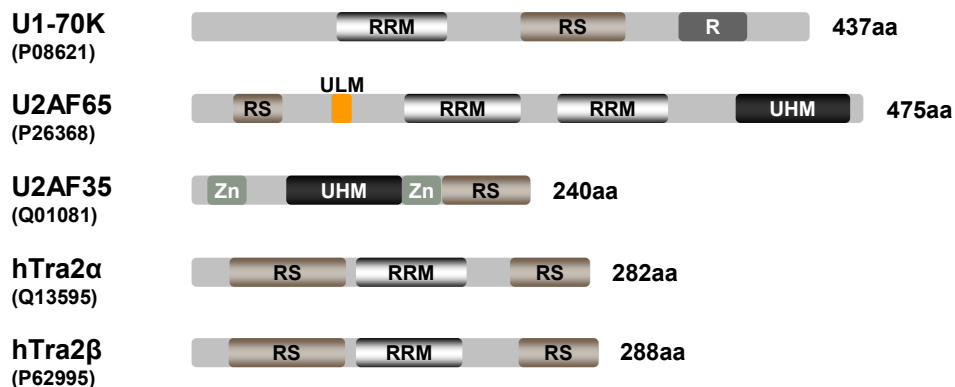
SR (black) and hnRNP proteins (grey) bind to splicing regulatory elements (SREs) embedded within the exon or intron and can either act positively (green) or negatively (red) on the inclusion of a regulated exon (indicated by dashed lines). Based on their splicing activity and relative location, SREs are classified into exonic or intronic splicing enhancer (ESE, ISE) or silencer (ESS, ISS). [U2: U2 snRNP; 65: U2AF65; 35: U2AF35; U1: U1 snRNP] [modified to (574)].

I.3.5 Splicing regulatory proteins

I.3.5.1 SR proteins

The SR proteins represent a family of multifunctional RNA binding proteins, which has been best described for their roles in the regulation of constitutive and alternative pre-mRNA splicing. In addition, they act at later steps of RNA metabolism, ranging from mRNA export and nonsense mediated decay (NMD) to translation [reviewed in (185, 317, 485)]. The first SR proteins were discovered in the late 80's in *Drosophila melanogaster* (8, 44, 95) and the early 90's in human cells (164, 171, 283), respectively. They were characterized by a common protein domain particularly enriched in arginine-serine (RS) dipeptide repeats and therefore termed RS domain. Ultimately, SR proteins were classified as such when they fulfilled the following requirements: i) reactivity with the monoclonal antibody mAB104 that recognizes

phosphorylated RS domains (444), ii) the property to reconstitute pre-mRNA splicing in cytoplasmic splicing deficient HeLa S100 extracts (607) and iii) their phylogentic conservation across vertebrates and invertebrates. Based on these criteria, twelve SR proteins have been described – SF2/ASF, SC35, SRp20, SRp75, SRp40, SRp55, 9G8, SRp46, SRp30c, SRp38, SRp54 and SRp35 (Fig.I-10A). All SR proteins share a modular structure with separated functional domains classically either mediating recruitment to the pre-mRNA or serving as effector domain for interactions with other splicing associated proteins (185, 276, 589). Here, at least one copy of an RNA recognition motif (RRM) domain within the N-terminal part of SR proteins dictates sequence-specific binding to the pre-mRNA, while an RS domain of variable length within the C terminus contacts general splicing components throughout splice site recognition. Beside their function to interact with other splicing-related proteins, the RS domain serves as a nuclear localization signal (NLS) and targets the SR proteins for nuclear import through an interaction with the SR protein specific nuclear import receptor transportin-SR (66, 260, 291). A genom-wide survey in metazoans could identify other RS domain-containing proteins, which lacked the classical RRM domain, failed to react with the mAB104 antibody or could not complement S100 extracts and were therefore collectively referred as SR-like or SR-related proteins (48). However, despite the lack of an RRM domain some of these SR-like proteins can bind to the RNA via functionally equivalent domains such as the PWI motif of the splicing factor SRm160 (42, 532). The SR-like proteins include both subunits of the U2AF heterodimer, the U1 snRNP specific protein U1-70K and the splicing coactivator SRm160/300 (43) (Fig.I-10B). Because of the expanding number of proteins with structural and functional overlap to SR proteins and the heterogeneous naming among the individual SR protein family members recently a novel unified nomenclature has been introduced (340). SR proteins are now defined by one or two N-terminal RRM domains and a C-terminal RS domain of at least 50 amino acids that is predominately composed of consecutive RS dipeptides. The 12 human SR proteins, which fulfilled the criteria, were renamed to SRSF (SR splicing factor) 1-12 to clarify that they belong together. SR proteins are expressed in all metazoan species and some lower eukaryotes with exception of the budding yeast *Saccharomyces cerevisiae* (188, 322, 607). However, *Saccharomyces cerevisiae* contains three SR-like proteins of which at least Npl3 was found to promote pre-mRNA splicing (285).

A**SR proteins****B****SR-like proteins****Fig.I-10: SR and SR-related proteins**

(A) SR proteins exhibit a modular organization into separated domains, including one or two N-terminal RNA recognition motifs (RRMs) and a C-terminal arginine-serine (RS) rich domain of variable size (here representatively shown for SRSF1-7). The RRM domain essentially permits binding to the pre-mRNA, whereas the RS domain is supposed to play an important role for the recruitment of spliceosomal components to adjacent splice sites. Domain positions were obtained from the Swiss-Prot database (www.uniprot.org).

(B) Several RS domain-containing proteins have been identified, which diverge from the canonical structural organization of SR proteins and were therefore referred to as SR-related or SR-like proteins. [RRM: RNA recognition motif; RRMH: RRM homology motif; RS: arginine-serine rich; ULM: UHM ligand motif; UHM: U2AF homology motif; Zn: zinc finger; R: arginine-rich]

The RRM domain is conferring the sequence-specific association of SR proteins with the pre-mRNA. Some of the SR proteins contain a second RRM, which deviates from the canonical motif and is therefore delineated as RRM homolog (RRMH). SR proteins specifically bind to a plethora of short, degenerated RNA sequences. Using “SELEX” (selected evolution of ligands through exponential enrichment), which is based on the selection of high affinity target sites from randomized pools of RNA sequences (549), motifs for SRSF1, SRSF2, SRSF3 and SRSF7 could be isolated (79, 534, 535). The base composition of the identified sequences revealed a high purine content as it was observed for exonic splicing enhancers (ESEs) before. A refined concept positively selecting functional RNA sequences that are not only bound by SR proteins, but also enhance splicing, led to the identification of binding motifs for the SR proteins SRSF1, SRSF2, SRSF5 and SRSF6 (108, 312, 313). Data were integrated into a web-based prediction tool (ESEfinder) in which candidate sequences can be scanned for binding sites of the four SR proteins (76). More recently combined cross-linking immunoprecipitation and high throughput sequencing (CLIP-seq) profiled natural target sites for SRSF1 on a genome-wide scale (457). 83% of all identified SRSF1 RNA targets were purin-rich sequences embedded within annotated exons of protein-coding genes. The structural concepts how the SR proteins bind these sequences are yet only vaguely understood. To date a model of a complete SR protein is not available due to the poor solubility of the RS domain. However, at least the RRM domains of individual SR proteins could be resolved by nuclear magnetic resonance (NMR) spectroscopy. The solution structure of the SRSF3 RRM in complex with a short “CAUC” RNA oligo elucidated a binding mechanism in which all four nucleotides form rather unspecific molecular interactions with the RRM (with the exception of the 5'-C) (199). This provides an explanation for the relatively loose RNA sequence specificity of SRSF3 and how a large spectrum of physiological target sites can be recognized by SR proteins (79, 462). In a more recent study, *in vitro* binding assays using SRSF1 supported a semi-conservative RNA binding mode for SR proteins. SRSF1 prefers binding to decameric, purin-rich sequences. However a stringent location of the purine residues is not required, while uracils inside or outside of a tested ESE sequence are not tolerated and significantly alleviate the binding efficiency of SRSF1 to the pre-mRNA (90). Thereby, a glycine-rich linker separating both RRMs coordinates the flexibility in substrate-binding of SRSF1.

SR proteins accumulate in subnuclear compartments termed speckles, which are distributed throughout the nucleus (293). There are two distinct types of speckles known: interchromatin granule cluster (IGCs) and perichromatin fibrils. IGCs represent the cells “storehouses” for splicing-related factors, where SR proteins wait on demand until required for splicing. Perichromatin fibrils are the sites of cotranscriptional splicing to which SR proteins can vastly be mobilized from the IGCs (163, 341, 511, 512). The RS domains target the SR proteins to the speckles and undergo extensive phosphorylation for the recruitment to the perichromatin fibrils (66, 207, 243, 371, 372). Furthermore, recent studies indicated that the release of SR proteins from the IGCs might be regulated by highly conserved long non-coding RNAs (ncRNAs), trapping SR proteins to reduce their movement to the active sites of splicing (9, 31, 547).

Initially, SR proteins were supposed to be functionally redundant because of their common capability to rescue constitutive *in vitro* splicing in splicing-deficient HeLa S100 extracts (607). However, unique functions were indicated by their sequence-specific RNA binding and their distinct outcomes for alternative splicing (471, 572, 608). The non-redundant functions of SR proteins were afterwards pronounced by severe developmental errors following depletion of specific SR proteins in various model systems such as *Drosophila melanogaster* (151, 216, 411, 429, 432), *Caenorhabditis elegans* (263, 318, 319) and mouse (125, 247, 374, 595). Misregulation of splicing is a common source of human diseases and was shown to be caused by mutations within exonic splicing enhancer (ESE) bound by SR proteins (75, 555). Furthermore, aberrant expression of SR proteins has been related to the development of cancerous tissues (152, 242, 257, 520), reiterating the importance of SR proteins for control of alternative splicing and other aspects of RNA metabolism (see chapter I.3.5.1.1 and I.3.5.1.2).

I.3.5.1.1 SR protein functions in alternative and constitutive splicing

Human splice site sequences are – as a result of their poor match to the consensus – often not sufficient to recruit the spliceosome by themselves. This is the reason why their immediate sequence environment in many cases contains additional elements, which can support the assembly of spliceosomal components. These auxiliary elements are known as ESEs and ISEs (exonic and intronic splicing enhancers, respectively). SR proteins are well-described to bind to exonic splicing enhancer (ESE) sequences to promote recognition of nearby splice sites and early (E) complex formation (37, 41, 185) (Fig.I-11A). The underlying mechanisms by which the splicing activation is achieved, can range from RS domain-dependent recruitment of spliceosomal components such as U1-70K or U2AF65 (164, 276, 589) to the displacement of silencer proteins from overlapping exonic splicing silencers (ESS) (624). SR proteins binding to exonic splicing enhancers (ESEs) can be described for both, alternatively and constitutively spliced exons (148, 461). The probably best known mode of activation is their targeting of U1-70K and U2AF65 via RS-RS protein domain mediated interactions, which directs spliceosome assembly at the 5'ss and the 3'ss. Activation of 3'ss positively correlates with the number of SR proteins bound to the exon and the extent of arginine-serine (RS) dipeptides within their RS domain (186). Furthermore, the relative distance of the ESE complex to the 3'ss determines the magnitude of splicing activation. However, more recent studies suggest a reassessment of the classical view that the RS domain enhances splicing exclusively through protein interactions with general splicing components, while RRM domain of SR proteins solely mediates their RNA binding (91, 481, 482). In this regard, the positively charged RS domains of SR proteins were shown to directly contact the pre-mRNA to neutralize repulsive powers between the negative phosphate backbones of the splice sites and the approaching U snRNAs (481, 482). Furthermore, mutational analysis indicated that rather the RRM domains of SRSF1 interact with U1-70K for recruitment of the U1 snRNP to the 5'ss, releasing the RS domain for other activities such as shielding of the negative phosphate backbone of the 5'ss (91). SR proteins can also enhance splicing regardless of their binding to ESE sequences (212). These ESE-independent activities include the enhancement of U4/U6•U5 tri-snRNP incorporation into the spliceosome and the stimulation of inter-splice site interactions across the intron (332, 440, 589) (Fig.I-11B). SR proteins are well characterized

antagonists of hnRNP proteins, classically defined as global repressors of splicing (68, 143, 144, 352, 353, 600) and it was proposed that their relative expression levels varying naturally over a wide range of tissues may contribute to portray cell specific alternative splicing patterns (198, 252). Besides their association with ESE sequences, SR proteins bind to intronic sequences to regulate alternative splicing, although there are by far fewer examples reported. However, a comprehensive study identified 17,365 RNA target sites for SRSF1 within protein-coding genes in the genome of cultured HEK 293T cells (457). Although most of them were mapped to exons, at least 2,911 were found to be embedded within flanking introns (17%). The importance of SR protein binding sites within introns remains largely unexplored. However, the few cases in which intronic positions of SR proteins were analyzed, these were associated with repression of splicing (58, 120, 213, 229, 256, 426, 483). At present, there is little understanding of the mechanisms that confers the repressive phenotype to intronic bound SR proteins. Several modes of repression were discussed, including steric hindrance (256) (Fig. I-11C) or redirection of spliceosomal components to intronic “decoy” sequences that lack functional splice sites (58). Recently, it was shown that SR-like protein Tra2 solely activate splicing when positioned to the exon, but that its tethering to an intronic location interferes with proper intron excision (483). The same study revealed that while splicing activation was mediated by the proteins’ RS domain, repression required its RRM domain. However, other SR proteins failed to inhibit splicing, when bound to the intron (e.g. SRSF1) so that it is still unclear whether repression through intronic positions can be considered as a unified feature of SR proteins (483).

Reversible phosphorylation of the RS domain of SR proteins is a key determinant of its function during pre-mRNA splicing. The phosphorylation state of SR proteins dynamically changes during the course of a splicing reaction, but above that also controls subcellular localisation and SR protein activities in other post-splicing events [(424, 533, 591, 592); reviewed in (510)]. For spliceosome formation the RS domain undergoes extensive phosphorylation on its serine residues (444), which was shown to enhance the interaction of SR proteins with other RS domain-containing splicing factors such as U1-70K (70, 276, 592). Cho *et al.* reported that hyperphosphorylation of the SRSF1 RS domain dissolves its intramolecular interaction with the RRM domains, resulting in an open conformation required for recruitment of the U1 snRNP (91).

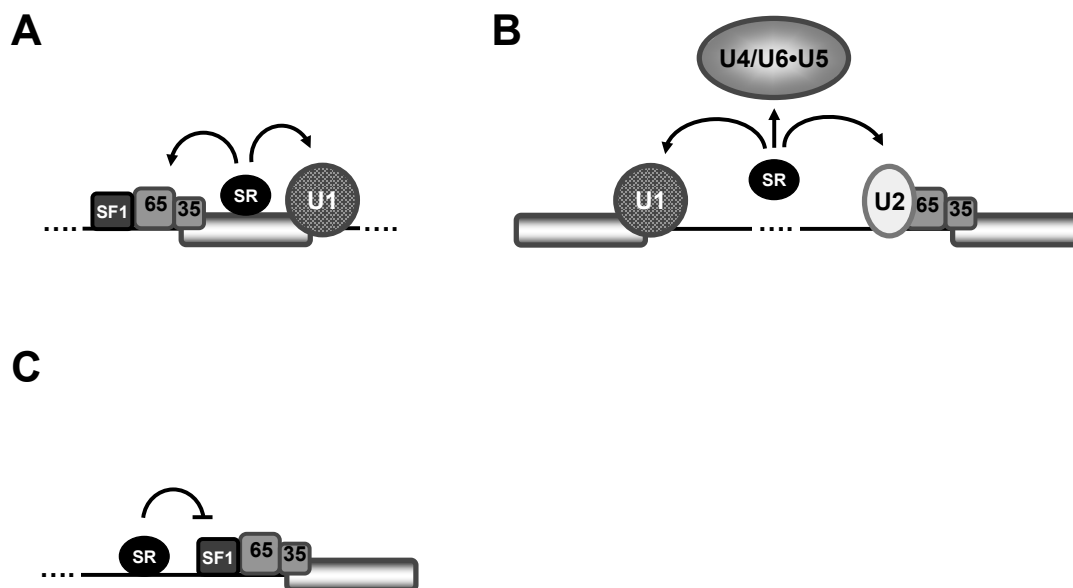


Fig.I-11: Modes of splicing by SR proteins

SR proteins bound to an exonic splicing enhancer (ESE) promote splicing through stabilization of spliceosomal components at the flanking splice sites (164, 276, 589) **(A)**. Apart from that, SR proteins were described to stimulate intron-definition in an ESE-independent manner (212) **(B)**. However, SR proteins were also found to inhibit splicing at the flanking splice sites when bound to an intron (256) **(C)**. [U1: U1 snRNP; 65: U2AF65; 35: U2AF35; U2: U2 snRNP]

However, several SR protein kinases phosphorylate RS domains, including SR protein kinase 1 (SRPK1), Clk/Sty and DNA topoisomerase I (104, 191, 396, 443, 509, 571). SR protein phosphorylation was also shown to be regulated via signal cascades, which link the splicing machinery to extrinsic stimuli (324, 488). The responsiveness of SR proteins to cells environment was supported by the findings that the Akt kinase is activated by ligands such as insulin and subsequently phosphorylates SR protein to modify alternative splicing patterns (39, 410). However, while phosphorylated SR proteins drive progression of the spliceosome into A and B complexes (71, 363, 440), transester reactions and export of the spliced mRNA require their dephosphorylation (71, 226, 591), which is mediated by the phosphatases 1 and 2A (325, 362, 363). Accordingly, dynamics in the phosphorylation state of the RS domains are crucial for SR protein functions in pre-mRNA splicing.

I.3.5.1.2 SR protein functions in other aspects of RNA metabolism

The last years have added several functions to the list of known functions in constitutive and alternative splicing, which have been described for SR proteins and emphasized their importance as general regulators of gene expression. Their extra functions ahead of splicing include mRNA nuclear export, NMD (nonsense-mediated decay) and translation [reviewed in (225, 317, 485)]. Most of the SR proteins with exception of SRSF2 can constantly shuttle between the nuclear and cytoplasmic compartment (67, 458). This shuttle activity was related to their activity to facilitate the transport of cellular RNAs into the cytoplasm by interaction with the NFX1/TAP export pathway (205, 223, 224). The affinity for NFX1/TAP is increased when the RS domain is hypophosphorylated (226). When SR proteins enter the cytoplasm they can again be phosphorylated by SR protein kinases on their RS domain and return to the nucleus to perform a novel round of splicing (126). SR proteins can affect translation (25, 456, 529, 531) either indirectly through a change in the alternative splicing pattern of factors involved in translation (257) or directly by suppression of 4E-BP, a negative regulator of CAP-dependent translation (367). SR proteins also enhance nonsense-mediated decay (NMD), a cellular surveillance pathway, which identifies erroneous spliced mRNAs with premature termination codons (PTCs) and targets them for degradation (459, 618). SR proteins use NMD to negatively autoregulate their own gene expression (296, 386). Therefore, they are supposed to trigger the inclusion of a PTC-containing exon (or intron), which then marks the mRNA for destruction by NMD.

I.3.5.2 hnRNP proteins

Heterogeneous nuclear RNAs (hnRNA) collectively describe all mRNAs in the nucleus regardless of their processing state. These hnRNAs are packaged by various proteins leading to the formation of heterogeneous nuclear ribonucleoprotein (hnRNP) particles [reviewed in (137)]. The term “hnRNP proteins” is therefore historically rooted due to their identification within high molecular mass hnRNP particles in the late 70’s (34, 138, 557). Finally, immunopurification experiments could differentiate a list of 20 proteins, termed hnRNP A-U (417). Reminiscent to SR

proteins, the hnRNP family constitutes a collection of RNA binding proteins acting as molecular adapters to various cellular pathways, most prominently pre-mRNA splicing, but also with functions during other steps of RNA metabolism [reviewed in (197, 284, 344)].

HnRNP proteins display a modular protein organisation with separated domains for RNA binding and protein-protein interactions interrupted by variable linker regions of distinct size (Fig.I-12). However, despite some commonalities, hnRNP proteins are highly divergent regarding their composition and functional properties. One widespread structural feature amongst hnRNP proteins is the RRM (RNA recognition motif) protein domain (217) with a typical “ $\beta\alpha\beta\beta\alpha\beta$ ” structure and two RNP consensus sequences (RNP-1 and RNP-2) (36, 320). The RRM folds into a four stranded antiparallel β -sheet backed by the two α -helices. The conserved RNP1 and RNP2 motifs are located in the central strands of the β -sheet and contain aromatic side chains, which are exposed from the surface of the β -sheet and stack with RNA nucleotides by hydrophobic interactions (343). These interactions just involve two residues on the RNA and confer the non-specific binding capacity of prototypical RRMs to short RNA sequences during mRNP biogenesis. Sequence specific binding associated with other functions in RNA metabolism can be achieved through extended interactions between the RNA and regions outside of the two conserved RNPs (343). It has been proposed that the packaging of the pre-mRNA by unspecifically bound hnRNP proteins protects it against degradation, while the sequence-specific RNA binding correlates with more specialized functions e.g. in alternative splicing. However, not all hnRNP proteins use a canonical RRM domain to direct association with the pre-mRNA. In other hnRNP proteins functionally equivalent RNA binding domains (129, 130, 498) [quasi (q)RRM; hnRNP K – homology (KH)] have been identified, satisfying the broad spectrum of distinct RNA binding specificities resolved by SELEX and CLIP [reviewed in (139, 180, 197)]. For example, hnRNP F/H proteins contain three qRRM domains that are characterized by poor conservation of the RNP consensus motifs and – in some cases – an extra β -hairpin formed between $\alpha 2$ and $\beta 4$ (74, 75). NMR structure revealed that the aromatic RNP residues normally involved in non-specific RNA binding are occluded within the qRRM by intramolecular interactions (74, 75), potentially accounting for the strict sequence-specific binding of hnRNP F/H proteins to guanosine (G) runs (130, 463).

hnRNP proteins

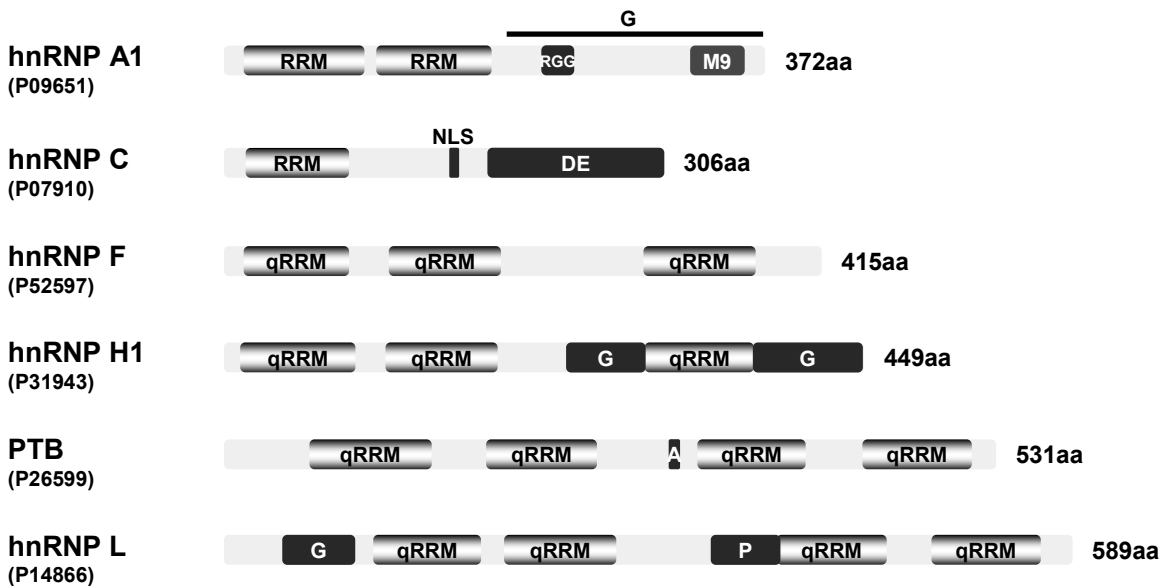


Fig.I-12: hnRNP proteins

The hnRNP protein family (here representatively shown) encompasses a collection of heterogeneous RNA binding proteins. Despite a commonality to at least contain one RNA binding domain (RBD), they differ in their protein domain structure and functional properties. Domain positions correspond to the respective entries in the Swiss-Prot database (www.uniprot.org) [G: glycine-rich domain; RGG: arginine-glycin-rich; RRM: RNA recognition motif; qRRM: quasi RRM; NLS: nuclear localization signal; DE: aspartate-, glutamate-rich; A: alanine-rich; P: proline-rich].

However, in addition to at least one RNA binding motif (RRM, qRRM or KH-type), hnRNP proteins often contain auxiliary domains enriched in glycines, prolines or acidic amino acids (G-rich; P-rich; DE-rich), which act in protein-protein interactions or play a role in subcellular localisation. Almost all hnRNP genes are alternatively spliced and the expressed hnRNP proteins undergo extensive posttranslational modifications [reviewed in (139, 180, 197)], including phosphorylation (368, 556, 594), sumoylation (303, 560), and arginine methylation (89, 210, 387, 407). This significantly expands the repertoire of alternative forms of hnRNPs with functional consequences for their activities, also in splicing (360, 556).

Despite their presence in the cytoplasm, hnRNP proteins are in generally localized diffusely throughout the nucleoplasm without accumulating in discrete areas (417) as it is described for SR proteins (293). However, there are some exceptions such as hnRNP G whose distribution partially overlap with nuclear speckles (18, 174, 419,

562). The fact that most of the hnRNPs preferentially localize to the nucleus has been related to their functions in mRNA export (19, 474). Likewise SR proteins, hnRNP proteins can continuously shuttle between the nucleus and the cytoplasm, which is supposed to guide the associated mRNAs out of the nucleus (67, 418). Shuttling hnRNP proteins often cooperatively assemble to homomeric or heteromeric complexes for packaging of their mRNA substrates (266) and different subsets of hnRNPs are specifically present on each mRNA during trafficking and mRNP biogenesis (369, 563). Many of the hnRNP proteins remain associated with the RNA during transport to the ribosomes, indicating a linkage to cytoplasmic pathways such as translation (87, 137). The hnRNP complexes are highly dynamic and subjected to a continual coming and going of hnRNP and other proteins, which specifies the outcome for the bound RNA and was termed “mRNA code” (500).

HnRNP family members have many features in common with other RNA binding factors, including SR proteins. This starts structurally with a modular protein organization into an RNA binding domain and an auxiliary domain used as a platform for protein-protein interactions. Moreover, hnRNP proteins functionally overlap with other RNA binding proteins in terms of their activities in splicing and post-splicing events (455, 500) (see chapter I.3.5.2.1). The high divergence amongst hnRNP proteins and the looseness of delineation from other RNA binding factors calls for a reassessment of their classification and a better resolution of the existing nomenclature (137, 138, 197, 344).

I.3.5.2.1 hnRNP protein functions in alternative splicing

The majority of hnRNP proteins participates in the regulation of alternative splicing. Early studies suggested an exclusive role as global repressors of splicing through binding to exonic and intronic silencer elements within numerous target genes. However, this is difficult to reconcile with the increasing number of cases in which hnRNP proteins, especially of the hnRNP F/H group, enhance inclusion of alternative exons. Several modes of repression by hnRNP proteins have been discussed so far, while there is only little understanding of how they can promote splicing.

HnRNP A/B proteins are perhaps best known as negative regulators of splicing, reflected by a plethora of exonic (122, 131, 176, 258, 353, 354, 438) and intronic

silencer elements (80, 259, 540), which are bound by members of this subgroup [reviewed in (344)]. The most simple scenario of repression envisions a sterical hindrance that impedes binding of SR proteins or general splicing components to exonic splicing enhancer or splice sites, respectively, through association of hnRNP A/B proteins with overlapping sites (128, 194). However, the observations that hnRNP proteins are capable to self-interact and in many cases silencer binding sites are distantly placed from splice sites or positive regulatory elements, rationalized another mechanism; this is that hnRNP A/B proteins are initially recruited to an high affinity binding site, which enhances the cooperative binding of additional hnRNP proteins to remote lower affinity binding sites (80, 397, 624). Spreading of hnRNP A/B proteins along the pre-mRNA establishes a “zone of silencing” that either displaces SR proteins from exonic splicing enhancer or sterically blocks the access of spliceosomal components to the splice sites (Fig.I-13A). This model was reinforced by recent findings, showing that hnRNP A/B proteins preferentially propagate in a 3'-to-5' direction along HIV-1 *tat* exon 3 to remove the SR protein SRSF2 from an exonic splicing enhancer and inhibit recognition of the upstream 3'ss (397). This process was so efficient that it even overwhelmed small RNA secondary structures placed into the exon to arrest nucleation. The property of hnRNP A/B proteins to self-interact has also been implicated in another mode of repression in which hnRNP A/B proteins bind to sites flanking an alternative exon to loop out the internal splice sites (38, 81, 381) (Fig.I-13B). However, “looping out” can also enhance splice site pairing, when hnRNP A/B proteins bind to the terminal ends of a large intron and fold out the internal RNA sequences to shorten the distance between the external splice sites (345) (Fig.I-13C). That “looping-out” might be a common strategy to facilitate splice site pairing was supported by computational analyses of overall 156,525 human introns, which revealed that hnRNP A1 binding sites at the intron ends are significantly overrepresented in large introns (>330nt) where the splice sites are far apart from each other (345).

A more general role of self-interaction among hnRNP proteins for splicing has been suggested for hnRNP C, which likewise hnRNP A/B proteins bind to pre-mRNA in a non-sequence specific manner (220, 364, 530, 581). Individual-nucleotide resolution UV-crosslinking and immunoprecipitation (iCLIP) has been applied to globally map hnRNP C RNA targets and revealed a loose preference for uridine-rich sequences with repetitive binding of hnRNP C in defined distances (165-300nt) within the

genome (280). While RNA targets were identified within both exons and introns, splice sites were specifically left out. A model was derived in which hnRNP tetramers occupy two binding sites and the intervening RNA sequence is compacted by unspecific protein-RNA interactions, which might bury alternative exons inside of the internal sequence to inhibit splicing or promote pairing of splice sites outside of the internal sequence by bringing them closer together (280).

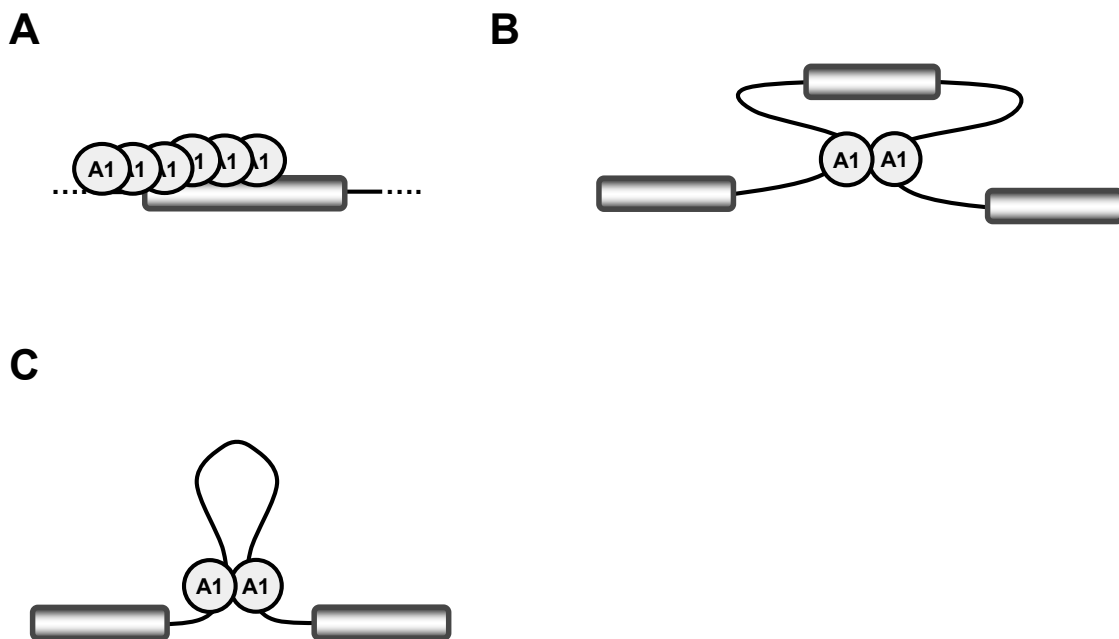


Fig.I-13: Modes of splicing by hnRNP proteins

Exonic hnRNP proteins multimerize across the exon and thereby occlude an adjacent splice site [“zone of silencing”, **(A)**] (80, 397, 624). In addition, hnRNP proteins bind to the flanking introns of an alternatively spliced exon **(B)**. Self-interactions looping out the internal splice sites then lead to splicing repression by interference with proper intron definition (38, 81, 381). However, hnRNP proteins can also bind to the terminal ends of an intron, folding out the internal sequences **(C)**. The flanking exons are brought into proximity, alleviating end-to-end bridging (345).

Members of the hnRNP F/H subgroup specifically recognize poly (G) runs (74, 129, 130, 463) to repress (57, 146, 155, 238, 436) or enhance splicing of alternatively spliced genes (94, 170, 196, 204, 370). Therefore, hnRNP F/H proteins appear to use common mechanisms as they have been suggested for hnRNP A/B proteins. Repression and activation by hnRNP A/B proteins were described to essentially rely on their property to self-interact, which requires an extensive glycine-rich domain (GRD) in the C-terminus (122, 354). This domain can also be found within hnRNP

F/H proteins. In line with this observation, GRDs originating from both hnRNP A1 and hnRNP H were documented to be functionally interchangeable for “looping out” of internal RNA sequences throughout splicing regulation (153). However, in contrast to hnRNP A/B proteins, hnRNP F/H proteins have also been widely described to act as enhancers of 5'ss through binding to G runs located in the downstream intron (94, 170, 196, 204, 370). Herein, it was shown that G runs are particularly enriched within the downstream regions of alternatively spliced exons (196, 331) and supposed to safeguard activation of 5'ss with weaker complementarities to the U1 snRNA (593). However, the molecular mechanism behind the enhancement of upstream 5'ss is still poorly understood.

HnRNP I is better known as polypyrimidine tract binding protein (PTB) because it was shown to negatively act on recognition of 3'ss via binding to upstream polypyrimidine tracts (PPTs) (412, 413, 501). SELEX analysis revealed that PTB preferentially binds to “UCUU” motifs (412, 501), however, related sequences such as “CUCUCU” have also been demonstrated to function in alternative splicing (83, 84). PTB contains four qRRM domains, however, the relative contributions of each qRRM to recognition of the pre-mRNA are still not fully understood and the data are conflicting (105, 175, 356, 392, 393, 498, 566). PTB is widely expressed in human tissues and was demonstrated to control a plethora of alternatively spliced exons (15, 513, 514). In addition to the modes of repression, including “sterical hindrance” (179, 189, 232, 309, 348, 377, 389) and “looping out” (7, 292, 514, 598) that have also been proposed for other hnRNP proteins, PTB was shown to specifically interfere with splicing at later steps of spliceosome assembly (477, 478). Herein, it was demonstrated that PTB stalls progression into a functional spliceosome at a step later than splice site recognition by interference with transition of exon- to intron-definition complexes, leading to the accumulation of “dead-end” pre-spliceosomal A-like (A') complexes (477). A more recent study revealed that PTB bound immediately downstream of the c-src N1 exon directly contacts the stem loop IV (SLIV) of the U1 snRNA, which renders the U1 snRNP inactive for further spliceosome assembly, potentially by favouring an inappropriate conformation (478). However, despite its thoroughly investigated role as negative regulator of alternative splicing, a comprehensive study identified a substantial number of alternative exons, which are positively regulated by PTB through binding to the downstream intron of regulated 5'ss (316). The commonality between hnRNP F/H and PTB to promote 5'ss use from

downstream intronic sequences inflicts the notion of a general property among hnRNP proteins to activate splicing in a position-dependent manner by a yet unidentified mechanism.

I.3.5.3 hnRNP-like splicing factors

Besides the two protein families of SR proteins and hnRNP proteins, there are further RNA binding proteins described to function in regulation of alternative splicing such as TIA-1, Fox1/2 or Nova (Fig.I-14). These proteins often show a tissue-specific or developmentally controlled expression to reprogram cellular splicing patterns and thereby dictate cell's differentiation and fate. Consistent with studies showing position-dependent splicing phenotypes for individual splicing factors such as PTB (316, 598) or Tra2 (483), certain hnRNP-like proteins were also shown to exhibit dual effects on splicing regarding the location of their binding sites relative to an alternatively spliced exon (550, 603, 614).

The T cell-restricted cellular antigen-1 (TIA-1) has been shown to specifically bind to uridine-rich sequences downstream of 5'ss, which enhances U1 snRNP recruitment (121, 123, 156, 232, 297, 357, 492, 628). TIA-1 contains three RNA recognition motifs (RRM) and an auxiliary glutamine (Q)-rich domain, which is located within the C-terminus (24, 157, 543). A second gene encodes the nearly identical TIA-1 related (TIAR) protein (123, 261, 297, 622) and there are at least two alternatively spliced isoforms described for both TIA-1 and TIAR (24, 233, 262). The expression of TIA-1 and TIAR isoforms is restricted to certain human tissues and cell lines, suggesting specialized functions in splicing regulation or other aspects of mRNA metabolism (233). Despite coprecipitation of the most C-terminal RRM domain (RRM3) of TIA-1 with cellular RNAs, a specific interaction with U-rich sequences could not be confirmed *in vitro* (123). Instead, the central RRM domain (RRM2) exclusively confers specific binding to U-rich sequences (123, 157, 287). TIA-1 was described to directly interact with the U1 snRNP specific protein C (U1-C) to facilitate U1 snRNP recruitment to upstream 5'ss (157). These interactions between TIA-1 and U1-C are relatively weak and a cooperative binding mutually stabilizes binding of TIA-1 and U1 snRNP to the pre-mRNA. Furthermore, proper interaction between both proteins requires a cystidine/histidine zinc finger like motif within U1-C (157), which is also

needed to incorporate U1-C into the U1 snRNP particle (382). Mutations disrupting the structural integrity of the zinc finger-like domain do not abrogate the interaction with TIA-1, suggesting that distinct surfaces mediate incorporation into the U1 snRNP particle and the interaction with TIA-1. However, RRM1 and the Q-rich domain of TIA-1 both contribute to U1-C binding and recruitment of the U1 snRNP to the 5'ss. It is striking to notice that while RRM1 is dispensable for binding of TIA-1 to the pre-mRNA and it fails to interact on its own with U1-C, it can significantly enhance recruitment of the U1 snRNP even in absence of the Q-rich domain (157). Whether the RRM1 of TIA-1 directly contacts the U1 snRNA as it was documented for the RRM domains 1 and 2 of PTB has to be clarified (478).

Fox1/2 proteins constitute a family of evolutionary conserved regulators of tissue-specific alternative splicing in metazoans [reviewed in (288)]. Mammals contain three family members, Fox-1, Fox-2 and Fox-3, however, multiple promoters and alternative splicing significantly expand the number of alternative forms (380, 553, 599). Metazoan Fox-1 expression was exclusively detected in brain, skeletal muscle and heart (244, 489, 553), while Fox-2 mRNA was found to be highly expressed in whole embryos and in adult brain, heart and ovary (553). Fox proteins specifically bind to (U)GCAUG sequences within regulated exons or their flanking introns and can either promote (227, 244, 307, 373, 422, 539, 599, 603, 614) or repress splicing (20, 117, 166, 244, 539, 603, 614, 621) for cell-type dependent regulation of alternatively spliced genes. Fox proteins contain a highly conserved central RNA recognition motif (RRM), which mediates specific binding to the pre-mRNA (16, 117, 380). The N-terminal and C-terminal regions separated by the RRM are less conserved, but a common feature among the Fox1/2 family members. The C-terminal part is also crucial for the splicing activity of Fox1/2 protein, while the N-terminal sequences are dispensable (20, 166, 526). Binding of Fox1/2 proteins to the (U)GCAUG sequence is exceptionally specific and less flexible than seen for other splicing factors, which was revealed by SELEX analysis (422). NMR studies resolved that the RRM adopts a classical “ $\beta\alpha\beta\beta\alpha\beta$ ”-fold with the two α -helices packed against the four stranded β -sheet (16). The 3'-half of the complexed RNA ligand is recognized by the canonical interface of the β -sheet, while the first four nucleotides “UGCA” are contacted by two loops, which wraps the U, G and C around a crucial phenylalanine and includes mismatch basepairing of G+2 and A+4 (16). This unusual binding mode confers high RNA binding affinity and specificity to Fox1/2 proteins.

hnRNP-like proteins

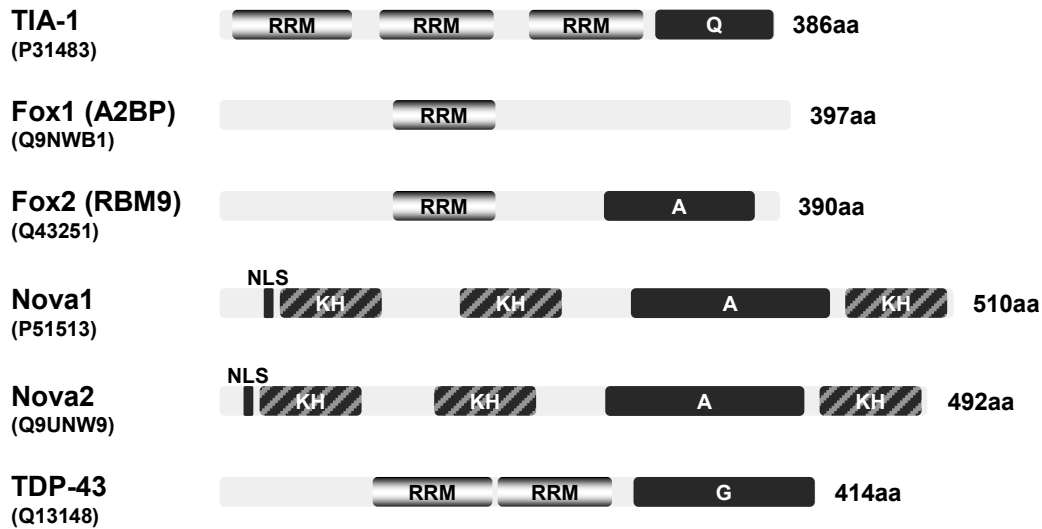


Fig.I-14: hnRNP-like proteins

HnRNP-like proteins show a modular domain organization, which resembles that of hnRNP proteins. Likewise hnRNP proteins, they contain at least one RNA binding domain and auxiliary domains rich in alanine, glutamine or glycine residues. Domain positions correspond to the respective entries in the Swiss-Prot database (www.uniprot.org) [KH: H homology motif; RRM: RNA recognition motif; Q: glutamine-rich; A: alanine-rich; G: glycine-rich]

Fox1/2 proteins can positively and negatively act on alternative splicing, which seems to be predominantly regulated by their position relative to the splice sites. While binding of Fox proteins to upstream intronic regions or within an alternatively spliced exon was shown to silence splicing (20, 117, 166, 244, 539, 603, 614, 621), their recruitment to sequences immediately downstream of a 5'ss was shown to enhance it (227, 244, 307, 373, 422, 539, 599, 603, 614). In consistence with the role of Fox proteins as enhancers of upstream 5'ss, (U)GCAUG motifs were found to be significantly overrepresented within proximal downstream intronic regions of alternatively spliced exons (77, 524, 603, 604, 614). However, while it was suggested that Fox proteins inhibit splicing by occlusion of exonic splicing enhancer or splice sites (349, 621), the mechanism beyond activation of upstream 5'ss is still unclear. Fox proteins were shown to interact with U1-C in two-hybrid assays and the interaction of both proteins could be a plausible mechanism by which U1 snRNP recruitment to the 5'ss is enhanced (394).

Nova was shown to be a key regulator within alternative splicing networks of the central nervous system. Nova modulates alternative splicing by binding to cluster of “YCA Y” motifs (where Y stands for a pyrimidine) (52, 241, 302). There are two Nova genes (Nova 1 and Nova 2) expressed in mammals, exclusively within the central nervous system (CNS) (53, 54, 601), which modulate approximately 700 alternative splicing events (613). The targeted exons contain “YCA Y”-cluster either within the regulated exon or the flanking introns in proximity to the splice sites. Reminiscent to the position-dependent splicing activities of Fox1/2 proteins, Nova can do both, activate or repress splicing of a particular exon. Comprehensive studies combining bioinformatics, biochemistry and genetics identified functional Nova RNA targets in a genome-wide scale and revealed that the asymmetric actions of Nova correlate with the location of its binding sites within the target genes (305, 550). However, the underlying principles of Nova mediated splicing enhancement and repression are still unclear, although mechanisms such as a “looping out” as described for hnRNP proteins have been suggested.

I.4 Alternative splicing by HIV

A single primary RNA is transcribed from the 5'-LTR promotor of the integrated provirus. Since scanning ribosomes most often start translation at the CAP-proximal AUG, the position of the Gag and Gag/Pol open reading frames (ORFs) closest to the 5'-end of the unspliced mRNA privilege them to be the first and sole one recognized. However, expression of the entire proteome and proper replication also require the translation of the seven further downstream located viral genes, including *tat* and *rev*. Alternative splicing sidesteps the hurdle of inhibitory upstream AUGs by removing them and juxtaposing the downstream reading frames to the 5'-end of the mRNA. Which particular HIV-1 protein is encoded by a spliced mRNA, is determined by the ORF immediately downstream of the used 3'ss. In addition, alternative splicing of the HIV-1 pre-mRNA permits regulation of the extent of splicing at the individual viral 3'ss and therefore allows an adjustment of the relative abundance of each of the viral mRNA species.

The HIV-1 pre-mRNA undergoes a complex series of alternative splicing events, which generates more than 40 different mRNAs (425). The functional strength of the

viral 3'ss is determined by their sequence composition plus the activity of splicing regulatory elements (SREs) in their vicinity [reviewed in (521, 522)]. These SREs act together in an intricate and dynamic splicing network that controls appropriate and timely ordered splice site activation and thereby essentially portrays the unique expression profile of each of the different viral proteins. Therefore some of the HIV-1 mRNAs occur in a relatively high abundance such as *env/vpu*- and *nef*-mRNAs whereas others are expressed at lower amounts such as *vpr*, *vif* and *tat*-mRNAs (425), reflecting the relative splicing efficiencies at each of the viral 3'ss.

Alternative splicing of the HIV-1 pre-mRNA is rather inefficient to allow the accumulation of unspliced *gag* and *gag/pol* mRNAs, which are used as a template for the expression of viral structural proteins or encapsidated as genomic RNA into new virions. A balanced ratio between unspliced and spliced viral mRNA levels is essential for proper HIV-1 replication as was shown by particular mutants with an excessive splicing phenotype, exhibiting a severe defect in virus particle production [e.g. (329), reviewed in (521)].

Intron-containing mRNAs are normally retained within the nucleus and marked for rapid degradation to prevent the expression of aberrant or poisonous proteins within the cell. However, the viral regulatory protein Rev mediates the escape of intron-containing 4kb and unspliced 9kb mRNAs from cellular RNA surveillance by coupling them to the Crm1 receptor-nuclear export pathway [reviewed in (47, 109, 200, 420)]. Therefore, Rev binds to a highly structured region within the *env*-intron, termed Rev-responsive element (RRE). In the early phase during infection all introns are removed from the primary transcript, which gives rise to the 1.8kb size mRNA class encoding the viral regulatory proteins Tat, Rev and Nef (267, 274, 378, 467). Rev then enters the nucleus and facilitates the transport of intron-containing viral mRNAs into the cytoplasm (267, 365, 378, 469). Thus, Rev is crucial for a temporal shift in HIV-1 gene expression by skirting the retention of intron-containing viral mRNAs within the nucleus and directing their export into the cytoplasm.

I.4.1 *Cis*-regulatory elements control HIV-1 pre-mRNA splicing

Alternative splicing of the HIV-1 pre-mRNA is under combinatorial control of several splicing regulatory elements (SREs) [reviewed in (521, 522)]. The SREs serve as platforms for cellular splicing factors and are placed adjacent to the viral splice sites they control. Along the HIV-1 pre-mRNA four major 5'ss and eight 3'ss bordered by the canonical GU and AG, respectively, can be found (425) (Fig.I-15A). Most of the splice sites are conserved across all HIV-1 strains with a few exceptions, e.g. the 3'ss cluster A4cab (522). Several studies were carried out to evaluate the intrinsic strength of the viral 5'ss and 3'ss. While the 5'ss D1 and D4 are relatively strong as predicted by their matches to the metazoan consensus motif, 5'ss D2 and D3 are rather weak consistent with a lower complementarity to the U1 snRNA (161, 391). The intrinsic strength of each of the viral 3'ss has been determined using an HIV-1 based one-intron splicing reporter (254). Herein, it was demonstrated that while 3'ss A2 and A3 splice at least at an efficiency of 40% compared to an optimal control 3'ss, 3'ss A1, A4c, A4a, A4b, A5 and A7 are all of weak intrinsic strength. This was attributed to the presence of non-consensus branch points (BPs) with nucleotides other than the preferred adenosine (118, 141) and functionally weak polypyrimidine tracts (PPTs) interrupted by purine residues. In agreement with this notion, mutations optimizing the PPT and the BPS towards their respective consensus sequences largely increase the efficiency of splicing at 3'ss A3, A5 and A7 (239, 254, 493, 517). However, addition of the downstream exonic sequences significantly change the efficiency of splicing at the different viral 3'ss, underlining the importance of splicing elements as integral part of the functional splice site strength. Use of 3'ss A1, A4cab, A5 and A7 is considerably increased in presence of their respective downstream exonic sequences, whereas the splicing efficiency at 3'ss A2 and A3 is decreased (254). Viral mRNAs, which encode for Vif, Vpr and Tat, are expressed at relatively low levels in infected cells, indicating that 3'ss A1, A2 and A3 are rarely spliced. By contrast, viral *rev*-, *nef*-, *env/vpu*-mRNAs are expressed at higher levels, suggesting that splicing at the 3'ss A4cab and A5 occurs more efficiently (425). In addition, approximately one half of all spliced viral mRNAs remove the downstream *env*-intron, indicating that 3'ss A7 is also used with high efficiency.

Splicing at each of the viral splice sites is tightly regulated by neighbouring splicing elements [reviewed in (521, 522)], including exonic and intronic splicing silencers

(ESSs/ISSs) and exonic splicing enhancers (ESEs) (Fig.I-15B). The exonic and intronic splicing silencer elements are bound by hnRNP family proteins, while SR proteins were identified to activate splice site use through binding to ESE sequences. The viral 3'ss A2, A3 and A7 all contain ESS elements within their downstream exonic sequences, which were found to negatively act on splice site activation (4, 6, 35, 73, 238, 329, 493, 494, 516, 624). Almost all of the as yet documented ESS (ESSV, ESS2, ESS3) elements contain “(Py)UAG” motifs, which match the typical core binding site for hnRNP A/B proteins (35, 73, 624).

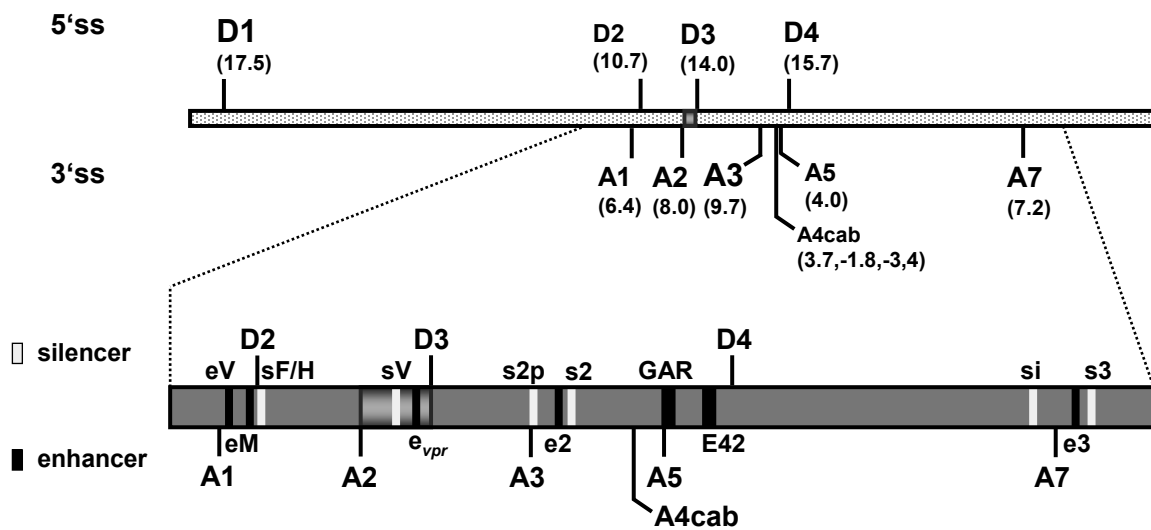


Fig.I-15: Alternative splicing of the HIV-1 pre-mRNA is controlled by a network of splice sites and splicing regulatory elements.

(A) The intrinsic strength of the 5'ss (D1 to D4) and 3'ss (A1-A7) distributed along the HIV-1 pre-mRNA. The size of the splice sites correspond to their predicted intrinsic strength

[5'splice site (5'ss): HBond Score, www.uni-duesseldorf.de/rna; 3'splice site (3'ss): MaxEntScore, http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html].

The location of exon 3 is indicated by a dark grey box. The nomenclature of the viral splice sites refers to (111).

(B) Positions of the splicing regulatory elements within the HIV-1 pre-mRNA. The 3'-part of the pre-mRNA containing known enhancer (black) and silencer (grey) sequences is enlarged. Exon 3 is flanked by 3'ss A2 and 5'ss D3. [eV: ESEvif (146); eM: ESEM1/M2 (254); sF/H: guanosine (G) rich silencer (146); sV: ESSV (35, 131, 329); e_{vpr}: ESE_{vpr} (this thesis); s2p: ESS2p (238); e2: ESE2 (194, 606); s2: ESS2 (4, 73, 493); GAR: guanosine-adenosine rich (GAR) ESE (11, 72, 253); E42: E42 fragment (11); si: ISS (540); e3: ESE3 (516); s3: ESS3 (6, 494, 516) [adapted to (11, 522)].

However, hnRNP F/H proteins have also been implicated to play a role in silencing of viral splice sites. An “UGGGU” sequence immediately downstream of 3’ss A3 was shown to be bound by hnRNP H for inhibition of the *tat*-mRNA specific 3’ss A3 (238). Beside their negative action through exonic silencer elements, hnRNP A/B and hnRNPF/H proteins were also described to bind to intronic sites, thereby repressing the use of an adjacent viral splice site (146, 540). An hnRNP A/B-dependent ISS element was identified overlapping with one of the BPS sequences used for splicing at 3’ss A7 (540). Disruption of the ISS by mutagenesis increases the splicing efficiency at 3’ss A7. In addition, a “GGGG” motif downstream of 5’ss D2 – supposed to be bound by hnRNPF/H proteins – was characterized to interfere with HIV-1 exon 2 inclusion (146). The mechanisms by which silencing is imposed on the viral 3’ss are not fully understood, although there is a general consensus that hnRNP A/B proteins prevent the access of splicing components to their target sites (5, 131, 494). Thus, it was suggested that repression of splicing by ESS sequences depends on the capability of hnRNP A/B proteins to multimerize via their glycine-rich domain (GRD) along the exon to thereby sterically hinder recognition of the splice site by constitutive splicing factors or of ESE elements by SR proteins, respectively (also see Fig. I-13A, “zone of silencing”). Further evidence for this mode of silencing was provided by the finding that initial high affinity binding of hnRNP A/B proteins to the ESS3 downstream of 3’ss A7 induces a cooperative 3’-to-5’ spreading along the pre-mRNA, which cannot just only unwind local RNA structures, but also displaces ESE-bound SRSF2 (116, 342, 397). Whether hnRNP A/B proteins also propagate from the ESSV and the ESS2 towards 3’ss A2 and A3, respectively, to block binding of general splicing factors or SR proteins to either the splice site or ESE sequences, remains to be answered. However, the ESSV was shown to specifically inhibit 3’ss A2, whereas U1 snRNP binding to the downstream 5’ss D3 is not impaired (131), which agrees with the observation that nucleation of hnRNP A/B proteins preferentially occurs in a 3’-to-5’ direction along the pre-mRNA (397). The negative effects of the silencers on viral splice site use are counteracted by several positive regulatory elements that increase the efficiency by which splice sites are engaged in the splicing reaction.

Exonic splicing enhancer (ESE) elements are recognized by different members of the SR protein family and could be identified to positively regulate viral splice sites in their vicinity (11, 72, 146, 254, 606). Different modes of action have been discussed

underlying SR proteins mediated activation of viral splice sites. Some studies reported that SR protein sequence motifs downstream of the 3'ss A3 and A7 overlap or are at least close to ESS elements, which may interfere with an accessibility for hnRNP A/B proteins (342, 606). Thus, the relative nuclear concentrations of SR proteins and hnRNP proteins could be the adjusting screw, which determines whether the positive or negative splicing element wins the binding competition and can emanate its activity. Another mechanism that was proposed to underlie activation of viral 3'ss is the SR protein-enhanced recruitment of the U1 snRNP to the downstream 5'ss. In consistence with the model of exon definition, downstream binding of the U1 snRNP to the 5'ss positively feed back on the assembly of general splicing factors at the upstream 3'ss and establishes an interaction network for communication between both splice sites (11, 215, 337). This was exemplified by results obtained in our working group, which revealed that a guanosine-adenosine rich (GAR) enhancer within HIV-1 exon 5 bound by the SR proteins SRSF1 and SRSF5 supports recruitment of the U1 snRNP to the flanking 5'ss D4, which in turn is necessary for bridging interactions across the exon and splice site pairing (11, 72, 253). In line with this notion, exon 5 recognition in the absence of the GAR element can be partially bypassed by coexpression of a mutated U1 snRNA perfectly matching 5'ss D4 (11). However, the GAR element was also shown to adjust the relative levels of splicing at the adjacent 3'ss A4cab and A5, hence, determining the ratio between *rev*- and *nef*-mRNA expression (11, 72).

Secondary structures have also been implicated in HIV-1 alternative splicing by possibly regulating the accessibility of ESS and ESE elements [reviewed in (453)]. Furthermore, the question of whether viral proteins such as Rev or Tat themselves commandeer the cellular splicing machinery to alter HIV-1 pre-mRNA splicing remains largely unanswered. Rev interaction with the RRE overrides the cellular prohibition on export of unspliced and intron-containing mRNAs into the cytoplasm. However, an additional role in inhibition of HIV-1 splicing has been discussed. For instance, it was suggested that Rev interacts with the splicing inhibitor p32, thereby bringing it into proximity of ESE-bound SR proteins and negating their activity (321, 416, 541, 620). HIV-1 Tat, which is well-documented to enhance transcriptional elongation through binding of the TAR sequence present within the 5'-end of all nascent viral mRNAs, was also attributed to play a role in alternative splicing regulation (32, 33, 234). In a study carried out recently, Tat was found to negatively

autoregulate the expression of its own mRNA by favouring use of the 3'ss A4cab and A5 at the expense of *tat*-mRNA specific 3'ss A3 splicing (234). The seen effects required the presence of an intact GAR enhancer within HIV-1 exon 5 and it was proposed that Tat-mediated recruitment of the p-TEFb complex to the TAR region would in turn direct the splicing coactivators Tat-SF1 and CA150, which stimulate selective binding of SRSF1 to the GAR element (234). However, other studies suggested that Tat would up- or downregulate the phosphorylation of ESE-bound SR proteins throughout infection by a timely-ordered recruitment of the cellular kinase Cdk13 at first followed by p32 to globally enhance the shift from spliced to unspliced mRNAs (32, 33).

I.4.2 Expression of intron-containing *vpr*-mRNAs and HIV-1 exon 3 splicing regulation

Some of the mRNAs from both major classes additionally include either one or both of two non-coding leader exons, defined by concerted splicing of A1 and D2 (exon 2) or A2 and D3 (exon 3), respectively (425). The alternative inclusion or exclusion of exon 2 and 3 results in the production of numerous isoforms within the 1.8kb and 4.0kb mRNA class and is therefore a major contributor to the huge diversity of over 40 different spliced mRNAs arising from HIV-1 pre-mRNA splicing (425). At present, the function of the two non-coding leader exons remains a puzzling issue, despite the fact, that there have been several studies carried out to clarify their importance for HIV-1 mRNAs (283, 334, 462). Because of the fact that neither exon 2 nor exon 3 create a new open reading frame (ORF) through a lack of an AUG, but become part of the 5'-untranslated region (5'-UTR) of distinct viral mRNA species, they were supposed to function in other aspects of mRNA metabolism such as RNA stability or translation. When their influence on gene expression was tested within the context of reporter genes, results were showing either no (467) or a positive effect for exon 2 and a negative effect for exon 3 (286). However, recent data with 5'ss D2 mutants, entirely excluding exon 2 from the viral mRNA species, could not find evidence for a noteworthy change in viral replication (337). Formation of intron-containing *vif* and *vpr*-mRNAs requires splicing at 3'ss A1 and A2, respectively. However, since the reading frames of Vif and Vpr start each within the downstream intron of either exon

2 or exon 3, splicing at 5'ss D2 and D3 prevents production of Vif- and Vpr-encoding mRNAs. Based on the findings, questioning the importance of exon 2 and 3 for mRNA stability, it was therefore speculated, they might be byproducts, resulting from dampening *vif*- and *vpr*-mRNA expression (521).

Vpr-mRNAs are spliced but intron-containing mRNAs expressed at relatively low levels within an infected cell (approximately 2% of the 4kb mRNA species). They are mainly formed by a splicing event pairing the major 5'ss D1 with 3'ss A2 of the non-coding exon 3. *Vpr*-mRNA expression and exon 3 inclusion into the viral mRNA species are negatively regulated by an exonic splicing silencer, termed ESSV, which is placed within the centre of exon 3 (35, 131, 329). The ESSV contains three (Py/A)UAG motifs within a 16nt long core sequence that bind members of the hnRNP A/B protein family to impose silencing on 3'ss A2. Using *in vitro* studies, it could be demonstrated that the ESSV specifically blocks binding of the general splicing component U2AF65 to the polypyrimidine tract (PPT) of 3'ss A2, whereas U1 snRNP recruitment to the downstream 5'ss D3 was unaffected (131). It was suggested that maintained U1 snRNP binding might be necessary for its potential functions aside from splicing, including RNA stabilisation (253) or inhibition of premature polyadenylation (12, 13, 250). Disruption of ESSV by mutagenesis results in largely upregulated levels of exon 3 inclusion and *vpr*-mRNA (35, 329). A similar splicing phenotype can be found when 5'ss D3 is mutated towards the metazoan 5'ss consensus sequence in the context of a *gag/pol* deleted proviral plasmid (35). By contrast, a lowered intrinsic strength of 5'ss D3 reduces splicing at 3'ss A2 and the amounts of *vpr*-mRNA in the context of ESSV-negative provirus (329). From these findings it was inferred that 5'ss D3 enhances *vpr*-mRNA splicing at 3'ss A2 by stabilizing interactions across the exon. This suggestion was further supported by the observation that a mutated U1 snRNA with an optimized complementarity to the 5'ss D3 sequence strongly increases splicing at upstream 3'ss A2 (338). Removal of the ESSV activity from exon 3 is linked to an excessive recognition of the flanking splice sites A2 and D3, leading to high levels of exon 3 inclusion and *vpr*-mRNA (329). The exon 3 oversplicing phenotype goes along with a dramatic perturbation of the balance between spliced and unspliced viral mRNAs and a severe defect in virus particle production (10 to 20 fold reduction), which is probably due to a failure to accumulate sufficient amounts of structural proteins required for capsid assembly. Therefore, the ESSV appears to be of particular importance to restrict HIV-1 exon 3

splicing to levels permitting both (i) accumulation of unspliced mRNA for structural protein expression and (ii) *vpr*-mRNA formation. Consistent with this key regulatory function, ESSV-negative viruses escape from their replication defect by second site mutations upon prolonged culturing, which either inactivate 3'ss A2 or 5'ss D3, respectively, to switch off unbalanced exon 3 splice site recognition (329). Overexpression of the SR protein SRSF1 selectively stimulates splicing at 3'ss A2 and reduces virus particle production (236, 237, 439). However, although splicing in absence of ESSV is above the levels predicted by the individual splice site strengths of 3'ss A2 and 5'ss D3, an SRSF1-dependent ESE within HIV-1 exon 3 has yet not been identified. Recently, the SR proteins SRSF4 and SRSF6 were reported to regulate *vpr*-mRNA production through inhibition of 5'ss D3 (545, 546). However, deletion analysis revealed that this effect is not linked to a specific target site within HIV-1 exon 3 or the downstream intron (545). Instead it was suggested that the accumulation of *vpr*-mRNA seen after overexpression of SRSF6 results from inhibition of GAR ESE-mediated activation of 3'ss A5 (545). In another study it was proposed that binding of the high mobility group A protein 1A (HMGA1a) to a sequence immediately upstream of 5'ss D3 is needed for *vpr*-mRNA expression (548). In support of this model, *in vitro* splicing of *vpr*-mRNAs is increased following addition of recombinant HMGA1a to HeLa nuclear extracts (548). It was stated that HMGA1a traps the U1 snRNP and ESSV-bound hnRNP A1 to simultaneously inhibit splicing at 5'ss D3, while relieving repression from 3'ss A2, respectively. However, whether it indeed plays a role for HIV-1 exon 3 splicing regulation in the context of replicating virus is an issue of upcoming experiments.

I.4.3 Manipulation of the host cell splicing machinery to optimize viral replication

There is only little understanding yet in whether the intracellular concentrations of nuclear splicing proteins are selectively manipulated during infection for regulation of HIV-1 gene expression. However, given the numerous studies, which demonstrate that HIV-1 replication is severely impaired by mutations within splicing regulatory elements (e.g. ESSV) [reviewed in (521, 522)] or changes in the expression levels of splicing factors that bind to these elements (236, 237, 439), it seems likely that the

expression levels of splicing factors represent targets for HIV-1 to optimally organize its pre-mRNA splicing. The splicing factor overexpression and siRNA experiments taken together further support the hypothesis that tight regulation of HIV-1 splicing is required for efficient virus replication and that this regulation can be abrogated by changes in the levels of cellular splicing factors (236, 237, 439). While a two- to threefold upregulation in the amount of SRSF2 can be detected 48h post infection of the T-cell line H9 (333), the expression of SRSF7 is decreased 60h after infection in the T-cell line M4 (450). However, the importance of these changes for the HIV-1 splicing profile was in neither of the two studies answered. In macrophages, HIV-1 infection is characterized by a productive virus particle production in the first weeks that continuously slows down over the course of time. It was found that this decrease in virus particle production is accompanied by a selective reduction in the expression of *tat*-mRNAs as coexpression of plasmid-driven Tat rescues efficient replication (508). Importantly, the decline in Tat is not induced by changes in the mRNA stability, but due to changes in the expression levels of certain splicing factors (135). *Tat*-mRNA species are formed by specific splicing at 3'ss A3. Two exonic splicing silencer (ESS), termed ESS2 (4, 73, 493) and ESS2p (238) and one exonic splicing enhancer (ESE), termed ESE2 (194, 606), which overlaps with the ESS2, are located in the downstream sequence of 3'ss. ESS2 and ESS2p repress splicing at 3'ss A3 through binding of hnRNP A1 and hnRNP H, while ESE2 binds the SR protein SRSF2 to promote *tat*-mRNA splicing. The decrease in Tat during infection is consistent with changing expression profiles of hnRNP A1, hnRNP H and SRSF2. hnRNP A1 and hnRNP H were found to decrease in the initial phase followed by a reinduction of expression during the following weeks. SRSF2 shows an inverted expression profile, being expressed at high levels during the first week of infection, but declining over the subsequent weeks (135). The differences in the levels of spliced mRNA species in HIV-infected macrophages and T-cells suggest the possibility that levels of host cell factors, binding to HIV-1 splicing elements (for example hnRNP A/B, hnRNP H or specific SR proteins) may differ in these two cell types. Modest increases in SR protein levels therefore appear to be sufficient to produce significant changes in HIV-1 alternative splicing in HeLa cells and reinforce the notion that the virus may exploit such changes within infected cells to alter the levels of specific viral mRNAs and to affect virus replication.

I.5 Aim of this work

Splice site activation is controlled by splicing regulatory elements (SREs), which can be commonly found in the immediate vicinity of human and viral splice sites. There is a long-held classification drawing the line between SR proteins as classical enhancers and hnRNP proteins as classical repressors of splicing. However, this rigid definition calls for a reassessment since it does not regard the growing number of examples in which SR proteins and hnRNP proteins show a mirror-imaged splicing phenotype and face off their activities; in other words: SR protein repress and hnRNP proteins activate splicing. Herein, the location relative to the regulated exon appears to be the determinant for the splicing outcome. To gain insights into the position-dependent activation and repression by SR and hnRNP proteins, this work aims to systematically investigate the effects of known splicing factors on 5'ss activation from an upstream exonic or downstream intronic location. In addition, this thesis intends to unravel the mechanistic principles underlying position-dependent splicing activation and repression.

In the second part of this work, the contribution of SREs to HIV-1 pre-mRNA splicing will be investigated. In particular, it will be examined how SREs embedded within the HIV-1 exon 3 act together to regulate expression of intron-containing *vpr*-mRNA. In addition, the importance of SREs for regulated HIV-1 pre-mRNA splicing and virus replication will be evaluated.

II. Materials and Methods

II.1 Materials

II.1.1 Chemicals, culture media and solvents

If not explicitly mentioned otherwise, Chemicals were used from Invitrogen, Merck, Riedel-de-Haen, Roth, Sigma and Serva. Preparation of growth media and solvents is described in the respective experimental protocols or derives from standard laboratory manuals (14, 454).

Cloning

Ampicillin	Roche	Cat.-No. 10835242001
dNTP Mix	Qiagen	Cat.-No. 201901
High Fidelity Polymerase	Roche	Cat.-No. 1173265001
LB-Broth (Lennox)	Roth	Cat.-No. X964.2
LB Agar (Lennox L Agar)	Invitrogen	Cat.-No. 22900-025
Plasmid Midi Kit	Qiagen	Cat.-No. 12145
Plasmid Maxi Kit	Qiagen	Cat.-No. 12163
Pwo DNA Polymerase	Roche	Cat.-No. 1644955
T4-DNA Ligase	NEB	Cat.-No. M0202S

Gel electrophoresis

1 kb DNA Ladder	Invitrogen	Cat.-No. 15615-024
“Kaleidoscope” SDS-PAGE Standard	Biorad	Cat.-No. 161-0375
LE Agarose	Biozym	Cat.-No. 840004
Prestained SDS-PAGE Standard		
Low Range	Biorad	Cat.-No. 161-0305
Rotiphorese Gel 30		
(37.5:1 acrylamide/bisacrylamide)	Roth	Cat.-No. 3029.1
Sieve GP Agarose		
(Genetic pure, low melting)	Biozym	Cat.-No. 850050

Material and Methods

Protein preparation

Complete Protease Inhibitors	Roche	Cat.-No. 1674498
Dithiothreitol (DTT)	Serva	Cat.-No. 20710
Leupeptin	Sigma	Cat.-No. L8884
NP40/Igepal	Sigma	Cat.-No. I3021
Pepstatin-A	Sigma	Cat.-No. P4265
PMSF	Roth	Cat.-No. 6367.3

RNA preparation

DNAse I recombinant, RNAse-free	Roche	Cat.-No. 04716728001
---------------------------------	-------	----------------------

Eukaryotic cell culture

DMEM	Gibco	Cat.-No. 41966
DPBS-CaCl ₂ -MgCl ₂	Gibco	Cat.-No. 14190
FCS	Gibco	Cat.-No. 10290-106
Geneticin G-418 Sulphate	Gibco	Cat.-No. 11811-031
Optimem	Gibco	Cat.-No. 31985-047
PenStrep	Gibco	Cat.-No. 15140
0.05% Trypsin-EDTA	Gibco	Cat.-No. 25300
Trypan Blue Stain 0.4%	Gibco	Cat.-No. 15250-061

Transfection

FuGENE® 6 Transfection Reagent	Roche	Cat.-No. 1810575
TransIT®-LT1	Mirus Bio LLC	Cat.-No. 731-0028
Lipofectamine™ RNAiMAX	Invitrogen	Cat.-No. 13778

Immunoblotting

ECL™ Western Blotting Detection Reagents	Amersham	Cat.-No. RPN2106
Hyperfilm™ ECL	Amersham	Cat.-No. RPN3103K
Nitrocellulose membrane	Protran	Cat.-No. 10401196
Protein-Assay	Biorad	Cat.-No. 500-0006
ReBlot Plus Stripping Solution	Chemicom	Cat.-No. 2509

RT-PCR

Ampli-Taq® DNA Polymerase	Applied Biosys.	Cat.-No. N0808-0166
LightCycler® DNA Master		
SYBR Green-I Kit	Roche	Cat.-No. 12158817001
Primer p(dT) ₁₅ Oligo-(dT)	Roche	Cat.-No. 10814290001
Recombinant RNasin®		
Ribonuclease Inhibitor	Promega	Cat.-No. N2511
SuperScript Reverse Transcriptase	Invitrogen	Cat.-No 18080-085

RNA affinity chromatography

Adipidic acid dihydrazide-Agarose	Sigma	Cat.-No. A0802
Nuclear extracts of 5x10 ⁹ Hela cells	Cilbiotech S.A.	Cat.-No. CC-01-20-25
ProteinLoBind tubes (1.5ml)	Eppendorf	Cat.-No. 0030108.116
Ribomax™ LargeScale RNA		
Production kit	Promega	P1300

Sample preparation of mass spectrometry (MS)

Formic acid (1M in H ₂ O)	Fluka	Cat.-No. 06473
ProteinLoBind tubes (0.5ml)	Eppendorf	Cat.-No. 0030108.094
Trypsin (proteomics grade)	Sigma	Cat.-No. T6567
ZipTipC ₁₈ (standard bed)	Millipore	Cat.-No. ZTC 18S 008

Flow Cytometry

4',6-Diamidino-2-phenylindole-Dihydrochloride (DAPI)	Polysciences	Cat.-No. 09224
--	--------------	----------------

II.1.2 Enzymes

Restriction enzymes were supplied by New England Biolabs (NEB), Fermentas (MBI), Invitrogen, Promega and Roche. Enzymes were used as recommended by the manufacturer with the corresponding buffers provided.

II.1.3 Cells

II.1.3.1 Bacteria Cells

The *Escherichia coli* (E.coli) strain DH5 α F'IQ (Invitrogen) was transformed with recombinant plasmids. For amplification of DNA plasmids subjected to cloning and restriction analyses with methylation-sensitive enzymes, the E.coli strain DM1 (Gibco) was used, which is deficient in functional methylases.

Genotypes of E.coli strains

DH5 α F'IQ: F ϕ lacZ Δ M15 Δ (lacZYA-argF) U169 *recA1 endA1 hsdR17* (*r_k*⁻,*m_k*⁺) *phoA supE44* *λ*⁻ *thi*⁻1 *gyrA96 relA1*/F' *proAB+* *lacIqZΔM15 zzzf::Tn5* [KmR]

DM1: F⁻ *dam*⁻13::Tn9(Cm^r) *dcm mcrB hsdR*⁻M⁺ *gal1 gal2 ara lac thr leu ton*^r *tsx*^r Su^o *λ*⁻

II.1.3.2 Eukaryotic Cells

The human cervix carcinoma cell line HeLa, the derivate cell line HeLa-T4⁺ and the human embryonic kidney (HEK) cell line 293T were used for transient transfection experiments. The cell line HeLa-T4⁺ is characterized by surface expression of the human CD4 receptor, which is stably integrated into the cells genome (328). Cells containing the transgene were selected by addition of geneticin to the culture media.

II.1.4 Oligonucleotides

Lyophilised oligonucleotides were provided by Metabion AG (Martinsried) and resolved in sterile H₂O to a stock concentration of 100μM. Oligonucleotides in solution were stored at -20°C.

II.1.4.1 Cloning Primers

Recognition sites for restriction enzymes are underlined.

II.1.4.1.1 SV-env/eGFP reporter plasmids

- #640:** 5' CAA TAC TAC TTC TTG TGG GTT GG 3'
- #1848:** 5' GCA GGA TCC GGA GGA GGA GGA GGA ATG GCC CGA TCT GTG 3'
- #1849:** 5' GCA CTC GAG TTA TTT GTA CAG CTC GTC CAT 3'
- #3072:** 5' CCC CGA ATT CTT GTG AAT ATC AAG CAG GAC ATA ACG AGC TCG CAG TAA GTA GCT T 3'
- #3073:** 5' CCC CGA ATT CTT GTT AAT ATC AAG CAG GAC ATA ACG AGC TCG CAG TAA GTA GCT T 3'
- #3075:** 5' CCC CGA ATT CTT GTG AAT ATC TAG CAG GAC ATA ACG AGC TCG CAG TAA GTA GCT T 3'
- #3077:** 5' CCC CGA ATT CTT GTT AAT ATC TAG CAG GAC ATA ACG AGC TCG CAG TAA GTA GCT T 3'
- # 3163:** 5' CCC CGA GCT CCT GGT GAG TAC AAT TAT TGA CCG TAC ACC ATC AGG GTA CGC GAA TTA TTG ACC GTA CAC CAT CAG GGT ACG CGA ATT CAT TAA TCT TAA GCT CTC CGA AGA CAG TGG 3'
- #3164:** 5' GCC TAG GAA TTC AGG CAT TAG CGG CGA CTT CGC TTA AGA GCT CCT GGT GAG TAC CTT AAG CTC TCC GAA GAC AGT GG 3'
- #3180:** 5' GGG GCC TAG GAA TTC TCT TTT TAA GTC GTA CCT AAT CTT TTT AAG TCG TAC CTA AGA GCT CCT GGT GAG TAC CTT 3'
- #3181:** 5' GGC TGA GCT CCT GGT GAG TAC TCT TTT TAA GTC GTA CCT AAT CTT TTT AAG TCG TAC CTA ACT TAA GCT CTC CGA AGA CAG TG 3'
- #3182:** 5' GGG GCC TAG GAA TTC AGA CAA CGA TTG ATC GAC TAA GAC AAC GAT TGA TCG ACT AGA GCT CCT GGT GAG TAC CTT 3'
- #3183:** 5' GGC TGA GCT CCT GGT GAG TAC AGA CAA CGA TTG ATC GAC TAA GAC AAC GAT TGA TCG ACT ACT TAA GCT CTC CGA AGA CAG TG 3'

- #3188 :** 5'GGG GCC TAG GAA TTC GCT TTC ATT TTT GTC TTT TTT TTA
AGC TTT CAT TTT TGT CTT TTT TTT AAG AGC TCC TGG TGA GTA
CCT T 3'
- #3189:** 5' GGC TGA GCT CCT GGT GAG TAC GCT TTC ATT TTT GTC TTT
TTT TTA AGC TTT CAT TTT TGT CTT TTT TTT AAC TTA AGC TCT
CCG AAG ACA GTG 3'
- #3192:** 5' GGG GCC TAG GAA TTC GTT CCA GAT AAG TTC CAG CCG TTC
CAG ATA AGT TCC AGC CGA GCT CCT GGT GAG TAC CTT 3'
- #3193:** 5' GGC TGA GCT CCT GGT GAG TAC GTT CCA GAT AAG TTC CAG
CCG TTC CAG ATA AGT TCC AGC CCT TAA GCT CTC CGA AGA
CAG TG 3'
- #3194:** 5' GGG GCC TAG GAA TTC AGG GAA GGG AGA GCT CCT GGT
GAG TAC CTT 3'
- #3195:** 5' GGC TGA GCT CCT GGT GAG TAC AGG GAA GGG ACT TAA
GCT CTC CGA AGA CAG TG 3'
- #3265:** 5' AGG GCC TAG GAG ACA ACG ATT GAT CGA CTA AGA CAA CGA
TTG ATC GAC TAG AAT TCT CTT TTT AAG TCG TAC CTA ATC TTT
TTA AG 3'
- #3266:** 5' ATA GGG CCT AGG TCT TTT TAA GTC GTA CCT AAT CTT TTT
AAG TCG TAC CTA AGA ATT CAG ACA ACG ATT GAT CGA CTA
AGA C 3'
- #3403:** 5' GGG GCC TAG GAA TTC TTA GGG TTA GGG CGA GCT CCT
GGT GAG TAC CTT 3'
- #3404:** 5' GGC TGA GCT CCT GGT GAG TAC TTA GGG TTA GGG CTT AAG
CTC TCC GAA GAC AGT G 3'
- #3405:** 5' GGG GCC TAG GAA TTC TGC ATG TGC ATG GAG CTC CTG
GTG AGT ACC TT 3'
- #3406:** 5' GGC TGA GCT CCT GGT GAG TAC TGC ATG TGC ATG CTT AAG
CTC TCC GAA GAC AGT G 3'
- #3407:** 5' GGG GCC TAG GAA TTC CTC TCT CTC TCT GAG CTC CTG GTG
AGT ACC TT 3'
- #3408:** 5' GGC TGA GCT CCT GGT GAG TAC CTC TCT CTC TCT CTT AAG
CTC TCC GAA GAC AGT G 3'

- #3409:** 5' GGC TGA GCT CCT GGT GAG TAC CCA AAC AAC CAA ACA ACC
AAA CAA CTT AAG CTC TCC GAA GAC AGT G 3'
- #3410:** 5' GGC TGA GCT CCT GGT GAG TAC CCA AAC AAC CAA ACA ACC
AAA CAA CTT AAG CTC TCC GAA GAC AGT G 3'
- #3419:** 5' ACT AGA GCT CCT GGT GAG TAC CTT AAG TCT TTT TAA GTC
GTA CCT AAT CTT TTT AAG TCG TAC CTA ACT CTC CGA AGA
CAG TGG CAA 3'
- #3418:** 5' AGG GCC TAG GCC AAA CAA CCA AAC AAC CAA ACA AGA ATT
CAG ACA ACG ATT GAT CGA CTA AGA C 3'
- #3423:** 5' GGC CTA GGA ATT CAG ACA ACG ATT GAT CGA CTA GAG CTC
CTG GTG AGT ACC TT 3'
- #3429:** 5' GGC CTA GGA TTC TCT TTT TAA GTC GTA CCT AAG AGC TCC
TGG TGA GTA CCT T 3'
- #3428:** 5' CCC CGA GCT CCT GGT GAG TAC TCT TTT TAA GTC GTA CCT
AAC TTA AGC TCT CCG AAG ACA GTG G 3'
- #3540:** 5' GGC TGA GCT CCA CGT GAG TCC TCT TTT TAA GTC GTA CCT
AAT CTT TTT AAG 3'
- #3541:** 5' GGC TGA GCT CCA CGT GAG TCC AGA CAA CGA TTG ATC GAC
TAA GAC 3'
- #3542:** 5' GGC TGA GCT CCC CGT GAG TCC TCT TTT TAA GTC GTA CCT
AAT CTT TTT AAG 3'
- #3543:** 5' GGC TGA GCT CCC CGT GAG TCC AGA CAA CGA TTG ATC
GAC TAA GAC 3'
- #3544:** 5' GGC TGA GCT CCT CGT ACG TCC TCT TTT TAA GTC GTA CCT
AAT CTT TTT AAG 3'
- #3545:** 5' GGC TGA GCT CCT CGT ACG TCC AGA CAA CGA TTG ATC GAC
TAA GAC 3'
- #3591:** 5' GAA CTA GAA TTC CCA AAC AAC CAA ACA ACC AAA CAA CCA
AAC AAC CAA ACA AGA GCT CCT GGT GAG TAC CT 3'
- #3592:** 5' TCT CCG AGC TCC TGG TGA GTA CCC AAA CAA CCA AAC AAC
CAA ACA ACC AAA CAA CCA AAC AAC TTA AGC TCT CCG AAG
ACA GTG G 3'
- #3611:** 5' ATA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG
GGG TAG GAT CCT TTG TAC AGC TCG TCC ATG C 3'

II.1.4.1.2 LTR ex2 ex3 reporter plasmids

- #2346:** 5' GAA GCG CGC ACG GCA AGA GGC GAG 3'
- #2347:** 5' CGC GAA TTC AGG CCT CTC TC 3'
- #2381:** 5' GGG CTC GAG ACT AGT GGCTGACTTCCTGGATG 3'
- #2384:** 5' GGG ACT AGT CAA GAA ATG GAG AAA AAA A 3'
- #2386:** 5' GGG CAT ATG TAT GTT TCA GGG AAA GCT AGG GGA 3'
- #2588:** 5' CTT TAC GAT GCC ATT GGG A 3'
- #2949:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTG AAT ATC AAG CAG GAC ATA AC 3'
- #2950:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTG AAT ATC TAG CAG GAC ATA ACA AGG TAG G 3'
- #2951:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTT AAT ATC AAG CAG GAC ATA ACA AGG 3'
- #2985:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTT AAT ATC TAG CAG GAC ATA ACA AGG TAG G 3'
- #2986:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTA AAT ATC AAG CAG GAC ATA ACA AGG 3'
- #2987:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTG AAT ATC CAG CAG GAC ATA ACA AGG TAG G 3'
- #3015:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTT AAT ATC CAG CAG GAC ATA ACA AGG TAG G 3'
- #3016:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTT AAT ATC CAG CAG GAC ATA ACA AGG TAG G 3'
- #3017:** 5' TTT CAG ACT CTG CTA TAA GAA AGG CCT TAT TCT GAC ACA TAG TTC TCC CTC TGT GTG AAT ATC TAT CAG GAC ATA ACA AGG TAG GAT C 3'

- #3202:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA TAT CAA TCA GGA CAT AAC
AAG GTA GGA T
- #3203:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA TAT CAA CCA GGA CAT AAC
AAG GTA GGA T 3'
- #3204:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TTT GAA TAT CAA GCA GGA CAT AAC
AAG 3'
- #3205:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TCT GAA TAT CAA GCA GGA CAT AAC
AAG 3'
- #3206:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA GAT CAA GCA GGA CAT AAC
AAG GTA G 3'
- #3207:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA AAT CAA GCA GGA CAT AAC
AAG GTA G 3'
- #3208:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA TAT CAA GCA GGA AAT AAC
AAG GTA GGA TCT CTA CAA TAC 3'
- #3209:** 5' TTT TCA GAC TCT GCT ATA AGA AAG GCC TTA TTC TGA CAC
ATA GTT CTC CCT CTG TGT GAA TAT CAA GCA GGA GAT AAC
AAG GTA GGA TCT CTA CAA TAC 3'
- #3297:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA AGC AGG ACA TAA C
3'
- #3298:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG ATC 3'
- #3299:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA AGC AGG ACA TAA
CAA GG 3'

- #3300:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG 3'
- #3301:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG 3'
- #3302:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA AGC AGG ACA TAA
CAA GG 3'
- #3303:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG TGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG 3'
- #3813:** 5' TTT TCA GAA TCT GCT ATA AGA AAT ACC ATA TTA GGA CGT
ATA GTT AGT CCT AGG TGT GAA TAT CAA GCA GGA CAT AAC
AAG GTA GGT AGT CTA CAG TAC TTG GCA CTA GCA G 3'
- #3814:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG TAG TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3815:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA AGC AGG ACA TAA
CAA GGT AGG TAG TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3816:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT AGG TAG TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3817:** 5' TTT TCA GAA TCT GCT ATA AGA AAT ACC ATA TTA GGA CGT
ATA GTT AGT CCT AGG TGT GAA TAT CAA GCA GGA CAT AAC
AAG GTT GGT TCT CTA CAG TAC TTG GCA CTA G 3'
- #3818:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'
- #3819:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA AGC AGG ACA TAA
CAA GGT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'

- #3820:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA GGC AGG ACA TAA
CAA GGT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'
- #3883:** 5' TTT TCA GAA TCT GCT ATA AGA AAT ACC ATA TTA GGA CGT
ATA GTT AGT CCT AGG TGT GAA TAT CAA GCA GGA CAT AAC
CAG CTA AGT ATT CTA CAG TAC TTG GCA CTA GCA G 3'
- #3884:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA GGC AGG ACA TAA
CCA GCT AAG TAT TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3885:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA AGC AGG ACA TAA
CCA GCT AAG TAT TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3886:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA GGC AGG ACA TAA
CCA GCT AAG TAT TCT ACA GTA CTT GGC ACT AGC AG 3'
- #3887:** 5' TTT TCA GAA TCT GCT ATA AGA AAT ACC ATA TTA GGA CGT
ATA GTT AGT CCT AGG TGT GAA TAT CAA GCA GGA CAT AAC
AAG CTT GGT TCT CTA CAG TAC TTG GCA CTA G 3'
- #3888:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT TAG TCC TAG GTG CGA ATA TCA GGC AGG ACA TAA
CAA GCT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'
- #3889:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG TGA ATA TCA AGC AGG ACA TAA
CAA GCT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'
- #3890:** 5' TTT TTC AGA ATC TGC TAT AAG AAA TAC CAT ATT AGG ACG
TAT AGT ATC GCC ACG TTG CGA ATA TCA GGC AGG ACA TAA
CAA GCT TGG TTC TCT ACA GTA CTT GGC ACT AG 3'

II.1.4.1.3 Bacteriophage coat fusion protein expression plasmids

Primers used for cloning of MS2 coat fusions

#1781:	5' AGA <u>CTC GAG</u> TTA AGA GGA CAC 3'
#1838:	5' GGT <u>GGA TCC</u> CGC ACA AGC CAT AGG CGA TC 3'
#1839:	5' AGA <u>CTC GAG</u> TTA ATC TCT GGA ACT CGA CCT GG 3'
#2694:	5' ATC <u>GGA TCC</u> GAG AAA AAG AAA ATG GTA ACT CAG GG 3'
#2695:	5' GAT <u>CTC GAG</u> TTA GTA GGG GGC AAA TCG GC 3'
#2954:	5' GGG GGG <u>GGA TCC</u> TCT AAG TCA GAG TCT CCT AAA GAG CCC 3'
#2955:	5' GGG GGG <u>CTC GAG</u> TTA AAA TCT TCT GCC ACT GCC ATA GCT AC 3'
#2980:	5' CCC CCG <u>GAT CCC</u> TGG GCC CTG AGG GAG GTG A 3'
#2981:	5' CCC CCC CTC GAG TTA GTC ATA GCC ACC CAT GCT GTT CT 3'
#2850:	5' GGA TCT <u>GCT AGC</u> GCT 3'
#2862:	5' GGG <u>CTC GAG</u> TTA GAT GGT GGA CTT GGA GAA GG GT 3'
#2869:	5' TGG <u>TGG ATC CGA</u> GGA CGA GAT GCC C 3'
#2870:	5' TGG <u>TCT CGA GTT</u> ACT GGG TTT CAT ACC CTG C 3'
#2954:	5' GGT <u>GGA TCC</u> GAC GGC ATT GTC CCA GAT A 3'
#2955:	5' CCC <u>CTC GAG</u> TTA GGC CGC GGC CAT GGT 3'
#2956:	5' AAA <u>ACC GCG GCC</u> TTC GGT G 3'
#3732:	5' CTA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCT GTA CGA GAG CGA GAT CTG C 3'
#3733:	5' CTA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCA GAG GAC ACC GCT CCT TCC 3'
#3734:	5' CTA <u>GAC TCG AGT</u> TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCA TCT CTG GAA CTC GAC CTG G 3'
#3735:	5' CTA <u>GAC TCG AGT</u> TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCG TCC ATT CTT TCA GGA CTT GC 3'
#3736:	5' CTA <u>GAC TCG AGT</u> TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCC TGG GTT TCA TAC CCT GCC 3'
#3737:	5' CTA <u>GAC TCG AGT</u> TAG GCG TAG TCG GGC ACG TCG TAG GGG TAA CTA GTA CCA CCG TAG GGG GCA AAT CGG CTG 3'

#3738: 5' CTA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG
GGG TAA CTA GTA CCA CCG TCA TAG CCA CCC ATG CTG T 3'

#3739: 5' CTA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG
GGG TAA CTA GTA CCA CCA AAT CTT CTG CCA CTG CCA TAG 3'

#3740: 5' CTA GAC TCG AGT TAG GCG TAG TCG GGC ACG TCG TAG
GGG TAA CTA GTA CCA CCG ATG GTG GAC TTG GAG AAG G 3'

II.1.4.1.4 U1 snRNA expression plasmids

#3924: 5' GCC CGA AGA TCT CGA TCC TAG CTT GCA GGG GAG ATA CCA
TGA TC 3'

#3925: 5' GCC CGA AGA TCT CGA TCC TAC CTT GCA GGG GAG ATA CCA
TGA TC 3'

#3926: 5' TTT TCA CTC GAG CCT CCA CTG TAG 3'

II.1.4.2 Sequencing Primer

#2588: 5' CTT TAC GAT GCC ATT GGG A 3'

#640: 5' CAA TAC TAC TTC TTG TGG GTT GG 3'

#2850: 5' GGA TCT GCT AGC GCT 3'

#2851: 5' CAC CAC AGA AGT AAG GTT CC 3'

II.1.4.3 RT-PCR Primer

II.1.4.3.1 Primers used for semi-quantitative RT-PCR

#640: 5' CAA TAC TAC TTC TTG TGG GTT GG 3'

#1224: 5' TCT TCC AGC CTC CCA TCA GCG TTT GG 3'

#1225: 5' CAA CAG AAA TCC AAC CTA GAG CTG CT 3'

#1544: 5' CTT GAA AGC GAA AGT AAA GC 3'

#2588: 5' CTT TAC GAT GCC ATT GGG A 3'

#3210: 5' TGA GGA GGC TTT TTT GGA GG 3'
 #3211: 5' TTC ACT AAT CGA ATG GAT CTG TC 3'
 #3323: 5' CTG AGC CTG GGAGCT CTC TGG C 3'
 #3392: 5' CGT CCC AGA TAA GTG CTA AGG 3'
 #3632: 5' TGG ATG CTT CCA GGG CTC 3'
 #3876: 5' ATC ACG TTC TCG ATC AAA AGG C 3'
 #3877: 5' TCG GGA CCG AGA CCT GC 3'

II.1.4.3.2 Primers used for quantitative RT-PCR

#3239: 5' TGA GGA GGC TTT TTT GGA GG 3'
 #3240: 5' TTC ACT AAT CGA ATG GAT CTG TC 3'
 #3242: 5' TCT TCC AGC CTC CCA TCA GCG TTT GG 3'
 #3243: 5' CAA CAG AAA TCC AAC CTA GAG CTG CT 3'
 #3323: 5' CTG AGC CTG GGA GCT CTC TGG C 3'
 #3324: 5' GGG ATC TCT AGT TAC CAG AG 3'
 #3387: 5' TTG CTC AAT GCC ACA GCC AT 3'
 #3388: 5' TTT GAC CAC TTG CCA CCC AT 3'
 #3389: 5' TTC TTC AGA GCA GAC CAG AGC 3'
 #3390: 5' GCT GCC AAA GAG TGA TCT GA 3'
 #3391: 5' TCT ATC AAA GCA ACC CAC CTC 3'
 #3392: 5' CGT CCC AGA TAA GTG CTA AGG 3'
 #3395: 5' GGC GAC TGG GAC AGC A 3'
 #3396: 5' CCT GTC TAC TTG CCA CAC 3'
 #3397: 5' CGG CGA CTG AAT CTG CTA T 3'
 #3398: 5' CCT AAC ACT AGG CAA AGG TG 3'
 #3629: 5' GGC GGC GAC TGG AAG AAG C 3'
 #3631: 5' CGG CGA CTG AAT TGG GTG T 3'
 #3632: 5' TGG ATG CTT CCA GGG CTC 3'
 #3633: 5' CGA CAC CCA ATT CTT GTT ATG TC 3'
 #3636: 5' CCG CTT CTT CCT TGT TAT GTC 3'
 #3637: 5' GAG AAG CTT GAT GAG TCT GAC 3'
 #3832: 5' GAG CTC CTG ACC CAC 3'
 #3876: 5' ATC ACG TTC TCG ATC AAA AGG C 3'

#3877: 5' TCG GGA CCG AGA CCT GC 3'

II.1.4.4 siRNA

siRNA (GCAAAGCCAAGAUGAGCCUdTdT) directed against TDP-43 was kindly provided by Dr. Emanuele Buratti.

II.1.4.5 Primers for *in vitro* transcription and RNA affinity chromatography

Natural binding sites derived from SV-env/eGFP

T7 promotor sequence is underlined and MS2 RNA hairpin sequence is italicized.

env (fwd): 5' TAA TAC GAC TCA CTA TAG G AC ATG AGG ATC ACC CAT GTA
AGT TCC TAT AGG CTT AAG GAT TC 3'

env (rev): 5' CCA TCT CCA CAA GTG CTG 3'

#19967936: 5' CTG CCA GGT AAG TAT 3' (anti-U1 snRNA oligo)

HIV-1 exon 3 ESE:

T7 promotor sequence complementary to the T7 sense primer is underlined

#2503: 5' TAA TAC GAC TCA CTA TAG G 3' (T7 sense primer)

#3133: 5' GTT ATG TCC TGC TTG ATA TTC ACA CCT ATA GTG AGT CGT
ATT A 3' (ESE)

#3134: 5' GTT ATG TCC TGC TAG ATA TTA ACA CCT ATA GTG AGT CGT
ATT A 3' (ESE -24G>T,-17A>T)

II.1.5 Recombinant plasmids

Sequences of all generated recombinant plasmids were verified by DNA sequencing of the respective target regions.

II.1.5.1 SV-env/eGFP reporter plasmids

The parental HIV-1 glycoprotein/eGFP expression plasmid (SV-env/eGFP) was kindly provided by Dr. M. Freund (161). The subgenomic SV-env/eGFP splicing reporter contains the coding sequence for the viral glycoprotein (*env*). The eGFP (enhanced green fluorescent protein) coding sequence was cloned into the plasmid by substitution of the 3'-terminal *Bam*HI/*Xho*I fragment of SV-env (72, 161) for a PCR product amplified using primer #1848 and #1849 and pEF eGFP-neo (kindly provided by Dr. Dirk Lindemann) as a template. Thereby the cytoplasmatic domain of the gp41 subunit of the viral glycoprotein was partially removed because it is dispensable for fusogenicity assays and syncytia formation in the context of Hela-T4⁺ cells stably expressing the viral entry receptor CD4. In generally, all upstream elements were cloned into the reporter plasmid by substitution of the *Avr*II/*Nde*I fragment of SV 2MS2 D1-env/eGFP or the *Eco*RI/*Nde*I fragment of SV IAS-1 D1 env-eGFP with an amplified PCR-product using a specific forward primer combined with the reverse primer #640. Whereas all downstream elements and 5'ss variations were introduced by substitution of the *Sac*I/*Nde*I fragment of SV GAR⁻ D1-env/eGFP with a PCR-product generated with a specific forward primer and the reverse primer #640. Positive clones were identified by changes in fragment sizes performing an *Eco*RI/*Nde*I restriction analyses or by loss of an *Hind*III restriction site in an *Hind*III/*Nhe*I restriction analyses. **SV GAR⁻ ESE⁻ D1-env/eGFP** was generated by replacing the *Cla*I/*Sac*I fragment of SV 2MS2 D1-env/eGFP with the corresponding fragment of SV GAR⁻ ESE⁻ D4-env/eGFP cloned by K. Schöneweis (466). Positive clones were confirmed in an *Cla*I/*Eco*RI restriction analyses. Plasmids **SV TIA-1(2x) SRSF7(2x) D1-env/eGFP**, **SV neutral(3x) SRSF7(2x) D1-env/eGFP**, **SV SRSF7(2x) TIA-1(1x) D1-env/eGFP** and **SV SRSF7(2x) TIA-1(2x) D1-env/eGFP** were generated by substitution of the *Avr*II/*Nde*I fragment of SV IAS-1(2x) D1-env/eGFP with a PCR-product using a specific forward primer and the reverse primer #640. Positive clones were selected by loss of an *Hind*III restriction site using an

HindIII/NheI restriction analyses. **SV neutral(3x) D1 SRSF7(2x)-env/eGFP**, **SV neutral(3x) D1 TIA-1 (2x)-env/eGFP**, **SV neutral(3x) D1 SRSF2 (2x)-env/eGFP** and **SV neutral(3x) D1 hnRNPF/H (2x)-env/eGFP** were created by substituting the *SacI/NdeI* fragment of SV neutral(3x) D1-env/eGFP with the respective fragment of SV GAR⁻ D1 SRSF7(2x)-env/eGFP, SV GAR⁻ D1 TIA-1(2x)-env/eGFP, SV GAR⁻ D1 SRSF2 (2x)-env/eGFP and SV GAR⁻ D1 hnRNPF/H(2x)-env/eGFP. Positive clones were identified by loss of an *AflII* restriction site performing an *Clal/AflII* restriction analyses. **SV SRSF7(4x) D1-env/eGFP** was cloned by introduction of the *Clal/EcoRI* fragment from SV SRSF7(2x) TIA-1(2x) D1-env/eGFP and of the *EcoRI/NdeI* fragment from SV SRSF7(2x) D1-env/eGFP into the *Clal/NdeI*-digested SV IAS-1(2x) D1-env/eGFP plasmid. **SV SRSF7(2x) neutral(5x) D1-env/eGFP** was created by introduction of the *Clal/EcoRI* fragment from SV SRSF7(2x) TIA-1(2x) D1-env/eGFP and of the *EcoRI/NdeI* fragment from SV neutral(5x) D1-env/eGFP into the *Clal/NdeI*-digested SV IAS-1(2x) D1-env/eGFP plasmid. Successful cloning of the latter plasmids was confirmed by loss of an *HindIII* restriction site in an *HindIII/NheI* restriction analyses. For construction of **SV neutral(3x) D1 TIA-1(4x)-env/eGFP** the *AflII/NdeI* fragment of **SV neutral (3x) D1 TIA-1(2x)-env/eGFP** was replaced by an amplified PCR-product using a specific forward primer and the reverse primer #640. Clones **SV SRSF7(2x) D1 SRSF7(2x)-env/eGFP** and **SV TIA-1 (2x) D1 SRSF7(2x)-env/eGFP** as well as **SV SRSF7(2x) D1 TIA-1(2x)-env/eGFP** and **SV TIA-1(2x) D1 TIA-1(2x)-env/eGFP** were obtained by introduction of the *SacI/NdeI* fragments of SV GAR⁻ D1 SRSF7(2x)-env/eGFP and SV GAR⁻ D1 TIA-1(2x)-env/eGFP, respectively, into SV SRSF7(2x) D1-env/eGFP and SV TIA-1(2x) D1-env/eGFP, respectively. **SV SRSF7(2x) L1 TIA-1(2x) env/eGFP**, **SV SRSF7(2x) L2 TIA-1(2x) env/eGFP** and **SV SRSF7(2x) L3 TIA-1(2x) env/eGFP** were cloned through substitution of the *SacI/AflII* fragment of SV SRSF7(2x) D1-env/eGFP with a PCR-product generated with a specific forward primer and the reverse primer #640. **SV SRSF7(4x) L2 TIA-1(4x)-env/eGFP** and **SV SRSF7(4x) L3 TIA-1(4x)-env/eGFP** were generated by cloning of the *EcoRI/AflII* fragment of SV SRSF7(2x) L2 TIA-1(2x)-env/eGFP or SV SRSF7(2x) L2 TIA-1(2x)-env/eGFP, respectively, in combination with the *AflII/NdeI* fragment of SV neutral(3x) D1 TIA-1(4x)-env/eGFP into the *EcoRI/NdeI*-digested plasmid SV SRSF7(4x) D1-env/eGFP. Positive clones in all the latter cases were verified by larger fragments in an *EcoRI/NdeI* restriction analyses. For cloning of **SV neutral(3x) D1-env/eGFP HA**, the HA tag sequence was fused to the 3' end of the coding

sequence for eGFP via *BsrFI* and *XhoI* restriction sites and introduction of a PCR-product amplified with primer pair #1848/#3611. **SV SRSF7(2x) D1-env/eGFP HA**, **SV neutral(3x) D1 SRSF7(2x)-env/eGFP HA**, **SV SRSF7(2x) D1-env/eGFP HA**, **SV TIA-1(2x) D1-env/eGFP HA** and **SV neutral(3x) D1 TIA-1(2x)-env/eGFP HA** were subsequently generated by substitution with the *NheI/XhoI*-fragment of SV neutral(3x) D1-env/eGFP HA. Positive clones containing the HA epitope sequence were identified through introduction of a *BamHI* restriction site using *BamHI* restriction analyses. **SV GAR⁻ D1 2MS2-env/eGFP** was cloned by substitution of the *SacI/NdeI* fragment of SV GAR⁻ D1-env/eGFP with an amplified PCR product using a specific forward primer and the reverse primer #640. Successful insertion was confirmed by a larger fragment in an *EcoRI/NdeI* restriction analyses. Plasmid **SV 2MS2 D1-env/eGFP** was generated by insertion of the *Clal/SacI* fragment of SV 2MS2 D4-env, cloned by Dr. M. Freund (72, 161), into *Clal/SacI* digested SV GARH D1-env/eGFP, cloned by I. Meyer. Positive clones were selected by a smaller fragment in an *EcoRI/NdeI* restriction analyses. Plasmids **SV D1 2MS2 D1-env/eGFP** and **SV 2MS2 D1 2PP7-env/eGFP** have been cloned by S. Rosin (442). All SV-env/eGFP plasmids contained a “D36G” amino acid substitution within the gp120 open reading frame, which was shown to enhance membrane fusogenicity (268).

Material and Methods

Cloned construct	Primer	Template
SV GAR ⁻ D1-env/eGFP	#3164	SV 2MS2 D1-env/eGFP
SV neutral(3x) D1-env/eGFP	#3409	SV 2MS2 D1-env/eGFP
SV neutral(5x) D1-env/eGFP	#3591	SV IAS-1(2x) D1-env/eGFP
SV neutral(3x) D1 neutral(3x)-env/eGFP	#3410	SV 2MS2 D1-env/eGFP
SV TIA-1(2x) D1-env/eGFP	#3180	SV 2MS2 D1-env/eGFP
SV GAR ⁻ D1 TIA-1(2x)-env/eGFP	#3181	SV 2MS2 D1-env/eGFP
SV SRSF7(2x) D1-env/eGFP	#3182	SV 2MS2 D1-env/eGFP
SV GAR ⁻ D1 SRSF7(2x)-env/eGFP	#3183	SV 2MS2 D1-env/eGFP
SV IAS-1(2x) D1-env/eGFP	#3188	SV 2MS2 D1-env/eGFP
SV GAR ⁻ D1 IAS-1(2x)-env/eGFP	#3189	SV 2MS2 D1-env/eGFP
SV SRSF2(2x) D1-env/eGFP	#3192	SV 2MS2 D1-env/eGFP
SV GAR ⁻ D1 SRSF2(2x)-env/eGFP	#3193	SV 2MS2 D1-env/eGFP
SV hnRNPF/H(2x) D1-env/eGFP	#3194	SV IAS-1(2x) D1-env/eGFP
SV GAR ⁻ D1 hnRNPF/H(2x)-env/eGFP	#3195	SV IAS-1(2x) D1-env/eGFP
SV hnRNPA1(2x) D1-env/eGFP	#3403	SV IAS-1(2x) D1-env/eGFP
SV GAR ⁻ D1 hnRNPA1(2x)-env/eGFP	#3404	SV IAS-1(2x) D1-env/eGFP
SV Fox1/2(2x) D1-env/eGFP	#3405	SV IAS-1(2x) D1-env/eGFP
SV GAR ⁻ D1 Fox1/2(2x)-env/eGFP	#3406	SV IAS-1(2x) D1-env/eGFP
SV PTB(2x) D1-env/eGFP	#3407	SV IAS-1(2x) D1-env/eGFP
SV GAR ⁻ D1 PTB(2x)-env/eGFP	#3408	SV IAS-1(2x) D1-env/eGFP
SV SRSF7(1x) D1-env/eGFP	#3423	SV IAS-1(2x) D1-env/eGFP
SV neutral(3x) D1 TIA-1(1x)-env/eGFP	#3428	SV IAS-1(2x) D1-env/eGFP
SV neutral(3x) D1 TIA-1(4x)-env/eGFP	#3419	SV IAS-1(2x) D1-env/eGFP
SV SRSF7(2x) TIA-1(2x) D1-env/eGFP	#3265	SV TIA-1(2x) D1-env/eGFP
SV TIA-1(2x) SRSF7(2x) D1-env/eGFP	#3265	SV SRSF7(2x) D1-env/eGFP
SV neutral(3x) SRSF7(2x) D1-env/eGFP	#3418	SV SRSF7(2x) D1-env/eGFP
SV SRSF7(2x) TIA-1(1x) D1-env/eGFP	#3429	SV TIA-1(2x) D1-env/eGFP
SV neutral(3x) D1 neutral(5x)-env/eGFP	#3592	SV SRSF7(2x) D1-env/eGFP
SV SRSF7(2x) L1 TIA-1(2x)-env/eGFP	#3540	SV IAS-1(2x) D1-env/eGFP
SV SRSF7(2x) L2 TIA-1(2x)-env/eGFP	#3542	SV IAS-1(2x) D1-env/eGFP
SV SRSF7(2x) L3 TIA-1(2x)-env/eGFP	#3544	SV IAS-1(2x) D1-env/eGFP
SV TIA-1(2x) L1 SRSF7(2x)-env/eGFP	#3541	SV IAS-1(2x) D1-env/eGFP
SV TIA-1(2x) L2 SRSF7(2x)-env/eGFP	#3543	SV IAS-1(2x) D1-env/eGFP
SV TIA-1(2x) L3 SRSF7(2x)-env/eGFP	#3545	SV IAS-1(2x) D1-env/eGFP
SV neutral(3x) D1-env/eGFP-HA	#3611 (rev)	SV neutral(3x) D1-env/eGFP
SV GAR ⁻ D1 2MS2-env/eGFP	#3163	SV GAR D1-env/eGFP
SV ESE D4-env/eGFP-HA	#3072	SV IAS-1(2x) D1-env/eGFP
SV ESE -24G>T D4-env/eGFP-HA	#3073	SV IAS-1(2x) D1-env/eGFP

SV ESE -17A>T D4-env/eGFP-HA	#3074	SV IAS-1(2x) D1-env/eGFP
SV ESE -24G>T, -17A>T D4-env/eGFP-HA	#3075	SV IAS-1(2x) D1-env/eGFP

Tab.II-1: Specific primers and PCR templates used for introduction of the respective sequence elements into SV-env/eGFP

II.1.5.2 Bacteriophage coat fusion protein expression plasmids

Construction of the expression plasmid for the “single chain” MS2 Δ FG coat protein variant (scMS2 Δ FG) has been described previously (499). For cloning of plasmids expressing scMS2 Δ FG coat fusions to splicing proteins (or domains) of interest or the respective N-terminally HA-tagged variants, respectively, the *Bam*HI/*Xho*I fragment of SV SD4/SA7 scMS2 Δ FG SRSF1(RS), was substituted with PCR-amplified fragments using appropriate forward and reverse PCR primers containing *Bam*HI and *Xho*I restriction sites. Plasmids **SV SD4/SA7 scMS2 Δ FG SRSF2(RS)**, **SV SD4/SA7 scMS2 Δ FG SRSF6(RS)** and **SV SD4/SA7 scMS2 Δ FG SRSF7(RS)** were generated by introduction of the respective *Bam*HI/*Xho*I fragment of SV SD4/SA7 NLS-MS2 SRSF1(RS), SV SD4/SA7 NLS-MS2 SRSF2(RS) and SV SD4/SA7 NLS-MS2 SRSF7(RS), kindly provided by C. Konermann (279). Positive clones were identified by restriction analyses. Expression plasmids for the “single chain” PP7 coat fusion proteins were cloned by S. Rosin (442).

Material and Methods

Cloned construct	Primer	Template
SVSD4/SA7 scMS2 ΔFG TIA-1	#2869/#2870	pGFP-TIA-1
SVSD4/SA7 scMS2 ΔFG Fox2α	#2697/#2695	pFox2 α
SVSD4/SA7 scMS2 ΔFG hnRNP F	#2980/#2981	pFlag-hnRNP F
SVSD4/SA7 scMS2 ΔFG hnRNP A1	#2954/#2955	hnRNPA1-MS2
SVSD4/SA7 scMS2 ΔFG PTB	#2954/#2955 (I) #2956/#2862 (II)	PTB-MS2
SVSD4/SA7 scMS2 ΔFG SRSF1 (RS)-HA	#1838/#3732	SVSD4/SA7 scMS2 ΔFG SRSF1 (RS)
SVSD4/SA7 scNLS-MS2 ΔFG SRSF2 (RS)-HA	#1781/#3733	SVSD4/SA7 scMS2 ΔFG SRSF2 (RS)
SVSD4/SA7 scNLS-MS2 ΔFG SRSF6 (RS)-HA	#1839/#3734	SVSD4/SA7 scMS2 ΔFG SRSF6 (RS)
SVSD4/SA7 scNLS-MS2 ΔFG SRSF7 (RS)-HA	#2850/#3735	SVSD4/SA7 scMS2 ΔFG SRSF7 (RS)
SVSD4/SA7 scNLS-MS2 ΔFG TIA-1-HA	#2869/#3736	SVSD4/SA7 scMS2 ΔFG TIA-1
SVSD4/SA7 scNLS-MS2 ΔFG Fox2α-HA	#2697/#3737	SVSD4/SA7 scMS2 ΔFG Fox2α
SVSD4/SA7 scNLS-MS2 ΔFG hnRNP F-HA	#2980/#3738	SVSD4/SA7 scMS2 ΔFG hnRNP F
SVSD4/SA7 scNLS-MS2 ΔFG hnRNP A1-HA	#2954/#3739	SVSD4/SA7 scMS2 ΔFG hnRNP A1
SVSD4/SA7 scNLS-MS2 ΔFG PTB-HA	#2954/#3740	SVSD4/SA7 scMS2 ΔFG PTB

Tab.II-2: Specific primers and PCR templates used for generation of the coat fusion expressing plasmids

II.1.5.3 LTR ex2 ex3 reporter plasmids

The minigene “LTR SD SA2Ex2SD2 komplett SA3Ex3SD3 SA7” was provided by M. Otte (400) and contains HIV-1 exon 3 sequences of pNLA1 (523), which is a cDNA derivate of HIV-1 NL4-3 (GenBank Accession No. [M19921](#)). For construction of the parental plasmid **LTR ex2 ex3 (pNLA1)** the *BssHII/EcoRI* fragment of “LTR SD SA2Ex2SD2 komplett SA3Ex3SD3 SA7” was substituted with a 5' ss D1 containing amplicon obtained from a PCR reaction with the primer pair #2346/2347. 3' ss A3 was then introduced by two PCR-amplified fragments (#2386/2381 and #2384/#1766) using *NdeI* and *XmaI* restriction sites. The resultant minigene construct contained the two small non-coding leader exon 2 and 3 and the 5'-part of the tat exon 1, interspersed by the authentic full-length sequences for the introns 2 and 3. Exon 3 mutants were constructed by PCR mutagenesis with appropriate forward primers. For construction the *A/wNI/SpeI* fragment of LTR ex2 ex3 was replaced by the respective PCR products using an appropriate forward PCR primer and the revers primer #2588, respectively. Since an *AvrII* restriction site was lost during mutation of the ESSV, positive clones could be identified using *SacI/AvrII* restriction analysis.

Cloned construct	Primer	Template
LTR ex2 ex3 (pNLA1)		
LTR ex2 ex3 ESSV ⁻	#2949	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -24G>T	#2951	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -24G>A	#2986	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -17A>T	#2950	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -17A>C	#2987	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -17A>G	#3016	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -15G>T,-17A>T	#3017	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -24G>T,-17A>T	#2985	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -24G>T,-17A>C	#3015	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -15G>T	#3202	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -15G>C	#3203	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -26G>T	#3204	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -26G>C	#3205	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -21T>G	#3206	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -21T>A	#3207	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -9C>A	#3208	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ ESE -9C>G	#3209	LTR ex2 ex3

Tab.II-3: Specific forward primers and PCR templates used for site-directed mutagenesis in LTR ex2 ex3 derived from pNLA1

The sequence of pNL4-3 revealed few changes of single residues within exon 3 to that of the derivate cDNA clone pNLA-1. Although none of the examined sequence elements were affected, an LTR ex2 ex3 plasmid was generated entirely based on pNL4-3 sequence (582) allowing to enable the mutations within the ESE sequence in the context of infectious pNL4-3 plasmid. Silent exon 3 mutations were introduced by PCR mutagenesis with specific forward primers. To create **LTR ex2 ex3 ESSV⁻** the *A/wNI/SpeI* fragment of LTR ex2 ex3 (pNL4-3) was substituted with an respective PCR product containing *A/wNI* and *SpeI* restriction sites. Positive clones were confirmed by loss of an *AvrII* restriction site performing *SacI/AvrII* restriction analysis. **LTR ex2 ex3 ESE -25T>C, LTR ex2 ex3 ESE -16A>G and LTR ex2 ex3 ESE -25T>C -16A>G** were cloned by insertion of an *A/wNI/SpeI* digested PCR product using a specific forward primer and the reverse primer #2588 into LTR ex2 ex3 ESSV⁻, while **LTR ex2 ex3 ESSV⁻ ESE -25T>C, LTR ex2 ex3 ESSV⁻ ESE -16A>G and LTR ex2 ex3 ESSV⁻ ESE -25T>C -16A>G** were generated by replacing the *A/wNI/SpeI*

fragment of LTR ex2 ex3 against an amplicon obtained by PCR with an appropriate *forward* primer and the *revers* primer #2588. This allowed to screen for positive clones by either gain or loss of an *AvrII* restriction site in *AvrII*/*SacI* restriction analyses. The varying 5'ss mutants were generated using the same strategy.

Cloned construct	Primer	Template
LTR ex2 ex3 (pNL4-3)		
LTR ex2 ex3 ESSV ⁻ (2)	#3297	LTR ex2 ex3
LTR ex2 ex3 ESE -25T>C	#3299	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -16A>G	#3303	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -25T>C, -16A>G	#3301	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C	#3299	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25A>G	#3300	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C, -16A>G	#3298	LTR ex2 ex3
LTR ex2 ex3 D3down	#3817	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -25T>C, -16A>G D3down	#3818	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESSV ⁻ (2) D3down	#3819	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C, -16A>G D3down	#3820	LTR ex2 ex3
LTR ex2 ex3 D3up	#3813	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -25T>C, -16A>G D3up	#3814	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESSV ⁻ (2) D3up	#3815	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C, -16A>G D3up	#3816	LTR ex2 ex3
LTR ex2 ex3 D3 1stG>C	#3887	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -25T>C, -16A>G D3 1stG>C	#3888	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESSV ⁻ (2) D3 1stG>C	#3889	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C, -16A>G D3 1stG>C	#3890	LTR ex2 ex3
LTR ex2 ex3 GTV	#3883	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESE -25T>C, -16A>G GTV	#3884	LTR ex2 ex3 ESSV ⁻
LTR ex2 ex3 ESSV ⁻ (2) GTV	#3885	LTR ex2 ex3
LTR ex2 ex3 ESSV ⁻ (2) ESE -25T>C, -16A>G GTV	#3886	LTR ex2 ex3

Tab.II-4: Specific forward primers and PCR templates used for site-directed mutagenesis in LTR ex2 ex3 derived from pNL4-3

II.1.5.4 Infectious HIV-1 plasmids

pNL4-3 mutants were constructed by replacing the region between *Pf*/MI and *Eco*RI with mutated LTR ex2 ex3 (pNL4-3) fragments.

II.1.5.5 U1 snRNA expression plasmids

pUCB U1 α D3 and **pUCB U1 α D3 1stG>C** were constructed by insertion of an PCR product amplified with the primer pairs #3924/#3926 and #3925/#3926, respectively, containing *Bgl*II and *Xho*I restriction sites at the template pUCB U1 into pUCB Δ U1 cloned by M. Freund(161) .

II.1.5.6 Protein expression plasmids

SVtat expresses the viral regulatory Tat protein from a cDNA, which is derived from the HIV-1 subgenomic construct pNLA-1(523). SVtat was cotransfected to activate transcription of LTR promoter-driven minigenes (460).

SVcrev expresses the viral regulatory Rev protein from a cDNA, which is derived from the HIV-1 subgenomic construct pNLA-1 (523). Transcription is under control of the SV40 early promoter. The plasmid was generated by substitution of the *Eco*RI/*Xho*I-fragment from pSVT7 with the *Eco*RI/*Xho*I-fragment from pUGcrev (460). The plasmids **pCMV myc SF2/ASF** and **pCMV myc SC35** for the SR proteins SRSF1 (SF2/ASF) and SRSF2 (SC35) were kindly provided by Dr. A. Cochrane. The expression plasmids **pFLAG-hnRNP F** and **pFLAG-hnRNP H** were obtained from Dr. M. Caputi. The mFox2 α expression plasmid was kindly provided by Dr. J. Conboy. The expression plasmid for TIA-1 was a kind gift of Dr. J. Valcarel. Expression plasmids for PTB-MS2 and hnRNP A1-MS2 were kindly provided by Dr. Ch. Smith.

II.1.5.7 Control plasmids

pSVT7 and **pcDNA3.1 (+)** were used to equalize the amounts of DNA between the samples for transient transfection experiments.

pXGH5 (472) was cotransfected to monitor transfection efficiency in quantitative and semi-quantitative RT PCR analyses. The plasmid expresses the human growth hormone 1 (hGH1) under control of the mouse metallothionein-1 promoter.

pCL-dTOM was cotransfected to detect transfection efficiency of each sample in fluorescence microscopical and flow cytometrical analyses. The plasmid expresses the fluorescent protein Tomato and was kindly provided by Dr. H. Harnenberg.

II.1.6 Antibodies

II.1.6.1 Primary Antibodies

Target subgroup	Protein	Provider	Cat.-No	Host	clone	Working dilution
HIV-1	p24	Biochrom AG	D7320	Sh	polyclonal	1:2000
	Vpr	NIH	3951	Rb	polyclonal	1:500
	Vif	NIH	2221	Rb	polyclonal	1:500
Splicing factors	SR proteins (phospho-epitope)	Invitrogen	33-9400	Ms	(1H4)	1:3000
	SRSF1	Zymed	32-4500	Ms	(96)	1:2000
	TDP-43	Prof. Dr. E. Burratti	-	Rb	polyclonal	1:1000
	hnRNP A1	Santa Cruz Biotechnology	sc-56700	Ms	9H10	1:500
U1-specific proteins	U1-C	Sigma-Aldrich	SAB4200188	Rat	4H12	1:1000
	U1-70K	Santa Cruz Biotechnology	sc-9571	Go	C-18	1:500
	U1-A	Sigma-Aldrich	SAB1100605	Rb	(4-15)	1:1000
Control	β -Actin	Sigma-Aldrich	A5316	Ms	AC-74	1:40.000
Others	MS2	Tetracore	TC-7004	Rb	polyclonal	1:2000
	HA	Sigma-Aldrich	H6908	Rb	polyclonal	1:2500

Tab.II-5: Primary antibodies and working dilutions used in western blot analysis

II.1.6.2 Secondary antibodies

Epitope Ig species	Provider	Cat.-No	Host	conjugate	Working dilution
Rb	Sigma-Aldrich	A6154	Go	HRP	1:2500
Go	Santa Cruz Biotechnology	Sc-2020	Do	HRP	1:2500
Rat	Sigma-Aldrich	A9037	Go	HRP	1:2500
Ms	Sigma-Aldrich	A9917	Go	HRP	1:3000
Sh	Jackson ImmunoResearch Laboratories	713-035-003	Do	HRP	1:2000

Tab.II-6: Secondary antibodies and working dilutions used in western blot analysis

II.1.7 Algorithms and databases

II.1.7.1 Algorithms to evaluate intrinsic 5'ss strength

The intrinsic strength of candidate 5'ss was predicted by use of the splice finder algorithm (http://www.uniduesseldorf.de/rna/html/hbond_score.php) and the MaxEntScore algorithm (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) (602).

II.1.7.2 Hexamer algorithm

The Hexamer algorithm was developed in collaboration with S. Theiss. From a dataset of 43,464 constitutively spliced canonical annotated human exons, 100 exonic and intronic nucleotides up- and downstream of 10,407 strong 5'ss ($HBS > 17.0$) and 10,359 weak 5'ss ($HBS \leq 13.5$) were selected, excluding the proper splice site sequence (3+8=11 nt). From these, differential hexamer frequencies for all 4,096 hexamers in four disjoint datasets were obtained: exonic and intronic sequences in the vicinity of strong and weak 5'ss. These hexamer frequencies were transformed into two standard normal Z-scores Z_{EI} and Z_{WS} , measuring the difference in hexamer occurrence in exonic versus intronic sequences, and in exons with strong versus weak 5'ss. A hexamer with large Z_{EI} , e.g., occurred significantly more often upstream than downstream of 5'ss. In the next step, the nucleotide based

scores Z_{EI} and Z_{WS} were calculated from these hexamer scores . Each single nucleotide is part of six neighboring hexamers – at either one of the six different hexamer positions. The index nucleotide \boxed{A} in the sequence TGTGA \boxed{A} TATCA, e.g., is located at the last position in hexamer TGTGA \boxed{A} , at the third position in hexamer GA \boxed{A} TAT and at the first position in \boxed{A} TATCA. For each (index) nucleotide, the two scores Z_{EI} and Z_{WS} were assigned as averages of the respective scores for each of the six hexamers containing the index nucleotide.

II.2 Methods

II.2.1 Cloning

II.2.1.1 Polymerase Chain Reaction (PCR)

Specific DNA fragments used for cloning were amplified from plasmid DNAs in a total volume of 50µl. Each reaction was prepared using the following ingredients:

5µl	10x Pwo buffer (10mM Tris-HCl, pH 8.85, 25mM KCl, 5mM (NH ₄) ₂ SO ₄ , 2mM MgSO ₄)
1µl	<i>fwd</i> -primer (10pmol/µl)
1µl	<i>rev</i> -primer (10pmol/µl)
4µl	dNTPs (each 2.5mM, Qiagen)
1µl	plasmid DNA (~10ng/µl)
0.5µl	Pwo DNA Polymerase (2,5U/µl, Roche)
ddH ₂ O till 50µl	

For amplification of large DNA fragments (>1.2kb) the expand High Fidelity PCR system (Roche) was used. PCR reactions were performed in a Robocycler Gradient 96 (Stratagene) under primer- and amplicon size-dependent conditions based on a standard PCR programm as follows:

94°C	3min	denaturation
94°C	30sec	denaturation
T _A °C	1min	primer-annealing
72°C	1min/kb	DNA synthesis
72°C	10min	final elongation

PCR reactions using High Fidelity Polymerase were performed at an elongation temperature of 68°C. PCR products were purified by addition of 1 vol. phenol (Roth) and 1 vol. chloroform/isoamyl alcohol (24:1). After the samples were mixed up phases were separated by centrifugation (12.000rpm, 10-15min, Eppendorf microcentrifuge) and the aqueous phase into which DNA portions was purified from traces of phenol by addition of 1 vol. chloroform/isoamyl alcohol (24:1), vortexing and

centrifugation at 12.000rpm for 5-10min (Eppendorf microcentrifuge). The supernatant with the DNA was removed and precipitated with 0.1 vol. 4M LiCl and 2.5 vol. ethanol (96°C) on dry ice for 20-30min. After centrifugation (maximum speed, 15-30min, Eppendorf centrifuge), the DNA was washed with 120µl ethanol (70% (v/v)) (maximum speed, 10min, Eppendorf centrifuge), air-dried and resuspended in 15µl ddH₂O.

II.2.1.2 Ligation

For ligation plasmids and PCR products were digested using the respective restriction enzymes and corresponding buffers provided by the manufacturer in a total reaction volume of 20µl. The expected fragment sizes were controlled by segregation in 1% agarose gels (LE agarose, Biozym) using TBE buffer (89mM Tris borate, pH 8, 2mM EDTA). Remaining plasmid DNA was loaded on a 0.8% low melting agarose gel (Sieve GP agarose, Biozym) using TB_{1/10}E buffer (89mM Tris borate, pH 8, 0.2mM EDTA) and electrophoresed. Separated DNA fragments were visualized by long wave UV-light (320nm) and subsequently cut out from the gel. Agarose with the isolated DNA fragments was melt at 65°C. Target vectors and PCR product with complementary ends were ligated using T4-DNA ligase (NEB) in a 20µl reaction containing 1µl T4-DNA ligase (400U/ml), 2µl 10x T4 ligase buffer (100mM MgCl₂, 100mM DTT, 10mM ATP, 500mM Tris-HCL, pH 7.5). Relative amounts of fragments were adjusted to an approximately 3:1 insert:vector molar ratio. In generally, ligation reactions were performed over night at 16°C. E.coli cells were afterwards transformed with the ligation samples.

II.2.1.3 Transformation

For amplification of plasmid DNA competent E.coli DH5αF'H (Invitrogen) cells were mixed with either 5µl from the ligation reaction or 0.5ng of highly pure Plasmid DNA, respectively, and incubated for 20min on ice. After heat-shocking for 90sec at 42°C, the transformed bacteria cells were chilled on ice and 800µl LB medium were added prior to incubation for 1h at 37°C and 220rpm. Subsequently, 200µl of the samples

were streaked on ampicillin-containing LB agar plates (100µg/ml) and grown overnight at 37°C.

II.2.1.4 Isolation of plasmid DNA from bacteria for restriction analysis.

Single colonies were picked from an agar plate and transferred into 5ml LB medium plus ampicillin (100µg/ml). Bacteria cells were incubated overnight at 37°C meanwhile continuously rotating at 220rpm. After centrifugation for 1min at maximum speed (Eppendorf micocentrifuge), cell pellets were resuspended in 300µl buffer 1 (50mM Tris-HCL pH7.5, 10mM EDTA, 400µg/ml RNase A) and lysed at RT for 5min by addition of buffer 2 (0.2M NaOH, 1% (w/v) SDS). Addition of buffer 3 (3M KAc, pH 5.5) neutralised the lysates and precipitated the proteins and bacterial debris, while the plasmid DNA remained in solution. Following centrifugation for 15min at 12.000rpm and 4°C the supernatant was removed and precipitated with 0.7 vol. isopropanol. Plasmid DNA was obtained by centrifugation for 15-30min at full speed and RT, washing with 120µl ethanol (70% (v/v)) and resuspension in TE buffer (pH 8.0). Positive clones were confirmed by restriction analyses with respective DNA endonucleases and specific digestion patterns resolved by gel electrophoresis using 1% agarose (LE, Biozym).

II.2.1.5 Isolation of plasmid DNA from bacteria for preparation.

150ml of ampicillin-containing LB medium (100µg/ml) were inoculated with 200µl of a 5ml bacterial pre-culture (see II.2.4) or a cryo-stock (prepared from 700µl bacteria culture and 300µl glycerol, stored at -80°C) and grown overnight at 37°C constantly shaking at 220rpm. Bacteria cells were pelleted by centrifugation for 10min at 5.000 rpm (Beckmann JS-21, JA14 Rotor) and resuspended in 4ml buffer 1 (50mM Tris-HCL pH7.5, 10mM EDTA, 400µg/ml RNase A). For alkaline lysis 4 ml buffer 2 (0.2M NaOH, 1% (w/v) SDS) was added, followed by incubation for 5min at RT. Addition of 4ml buffer 3 (3M KAc, pH 5.5) neutralised the lysate and after 15min incubation on ice, proteins and bacterial debris was pelleted via centrifugation for 30min at 10.000rpm and 4°C (Beckmann JS-21, JA14 rotor). Supernatant was cleared through

a folded filter (Schleicher & Schüll, 5951/2, Ø150 mm) and loaded on silica-based anion-exchange-columns (Plasmid DNA Midi Kit, Qiagen) for plasmid DNA purification according to the manufacturer's instructions. Plasmid DNA eluted from the columns was precipitated with 0.7 vol. isopropanol and centrifuged for 30min at 10.000rpm and RT (Beckman JS-21, JS13.1). After washing with 70% ethanol, plasmid-DNA was air-dried and resuspended in 50-300µl TE buffer (pH 8). DNA concentrations were quantitated by spectral photometry at 260nm and 280nm (NanoDrop) and adjusted to ~1µg/µl with TE buffer (pH 8). Positive clones were controlled by sequencing reactions performed in the Analytical Core Facility of the Biological-Medical Research Centre (BMFZ, HHUD).

II.2.2 Eukaryotic cell culture

II.2.2.1 Maintenance of human cell lines

Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% of heat-inactivated fetal calf serum (FCS, GIBCO), 100U/ml penicillin and 100µg/ml streptomycin (PenStrep, GIBCO) was used for culturing of HeLa and human embryonic kidney (HEK) 293T cells. HeLa-T4⁺ cells were positively selected by addition of 100 µg/ml geneticin in DMEM prepared as described above. Every 3-4 days subconfluent cell monolayer were detached for passaging or transient transfection experiments. Therefore, cells were washed twice with 5ml PBS_{def} (Dulbecco's phosphate buffered saline deficient in Ca²⁺ and Mg²⁺, DPBS, Invitrogen) and trypsinated with 1,5ml Trypsin-EDTA (Gibco) for 5-10min at 37°C. Enzyme was inactivated by resuspension of the cells in 10ml FCS-containing DMEM for sub-culturing or cell counting, respectively. All cell lines were grown at 37°C at humidified conditions (5% CO₂).

II.2.2.2 Transient transfection of human cell lines using FuGENE® 6 or TransIT®-LT1

Living cells were counted by negative staining using trypan blue (0.4% (w/v), Gibco) in a Neubauer chamber. For transient transfection experiments 2.5×10^5 cells were seeded in 6-well plates (TRP) in 2ml DMEM plus 10% FCS in absence of any antibiotics. 24h later cells were transfected with the respective expression plasmids by use of FuGENE® 6 (Roche) or TransIT®-LT1 (Mirus Bio LLC) transfection reagent according to the manual provided by the manufacturer, respectively. For western blot analysis medium was renewed 24h after transfection and transfected cells were collected after 48h. For RNA extraction, cells were harvested 24-30h following transfection. Transfections for transient expression of viral infectious plasmids were performed in the BSL-3 facility under S3 conditions.

II.2.2.3 siRNA assay

HeLa cells were seeded in six-well plates. Prior to transfection with the siRNAs, cell media was removed, cells washed with PBS and serum-free OPTIMEM media without antibiotics was added. Transfection with the small interfering RNA (siRNA) against TDP-43 and a negative siRNA control occurred at a final concentration of 160 pmol using Lipofectamine 2000 (Invitrogen). At day 2 (~ 12-14h later) media was changed against DMEM containing 10% fetal calf serum and 1% Penicillin/Streptomycin (50µg/ml each). Procedure was repeated at day 3 (~ 24h later) and cells were treated with siRNA for a second time. 12h after the second treatment with siRNA, cells were transfected with the HIV-1 infectious plasmids. For gene silencing of human TDP-43 the following siRNA duplex was used: GCAAAGCCAAGAUGAGCCUdTdT. The siRNA against TDP43 was kindly provided by Dr. E. Buratti.

II.2.3 Flow cytometrical analysis of transiently transfected HeLa cells

Cells samples were collected and washed with PBS. After trypsination for 5min at 37°C and several washing steps in FACS buffer (PBS + 3% FCS), samples were resuspended and acquired on a FACS-CANTOII cytometer (Becton Dickinson). To quantify the mean fluorescence intensity data were exported to the FlowJo (Tree Star, Inc.) analysis software.

II.2.4 Western Blot Analysis

II.2.4.1 Preparation of total protein cellular extracts

Adherent cells were harvested from the cell culture dish by scraping into the culture medium 48h post transfection. Cells were transferred into a fresh tube and pelleted by centrifugation for 2-3min at 2.000rpm and 4°C (Eppendorf microcentrifuge) followed by resuspension in 1ml PBS_{def}, washing and lysis in 70 µl RIPA buffer (25mM Tris-HCl pH 7.6, 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail (Roche)). After 10min incubation on ice, protein lysates were cleared from cell debris by centrifugation for 10min at full-speed and 4°C (Eppendorf microcentrifuge). Supernatant was collected and the protein concentration for each of the samples was determined by Bio Rad Protein assay according to the instructions provided by the manufacturer (Bradford et al. 1976). Protein samples were normalized, mixed with 5xSDS sample buffer (60mM Tris-HCL (pH 6.8), 24% glycerol, 2% SDS, 14.4 mM β-Mercaptoethanol, 1% bromphenol blue) and heated to 95°C for 10min.

II.2.4.2 Protein gel electrophoresis (SDS-PAGE)

Proteins samples were separated under denaturing conditions (Laemmli et al. 1970) in discontinuous sodium dodecyl sulphate (SDS) polyacrylamide gels (Rotiphorese Gel 30, Roth) using vertical flat bed gel electrophoresis. Gels were run in 1xLaemmli running buffer (1% SDS, 0.25M Tris-Base, 1.9M glycine) for 2-3h at 30mA and 4°C.

The composition of the gels is given in the following table:

	Separation			stacking
	7%	10%	12%	5%
PAA-30	5,6ml	8,0ml	9,6ml	1,5ml
0,5 M Tris, pH 6,8	-	-	-	1,2ml
2 M Tris, pH 8,8	5ml	5ml	5ml	-
20% SDS	120µl	120µl	120µl	45µl
60% Saccharose	-	-	-	2,1ml
APS	288µl	288µl	288µl	120µl
TEMED	48µl	48µl	48µl	12µl
water	13ml	10,6ml	10ml	4,2ml

Tab.II-7: Recipes for the gels used for protein separation by SDS-PAGE

II.2.4.3 Western blotting

After gel electrophoresis separated proteins were immobilized on a polyvinylidene difluoride (PVDF) membrane (Protran) using a semi-dry blotting chamber (Transblot® SD Semi-Dry, Bio-Rad). Electroblothing was performed at approximately 1mA per cm² of the membrane for 70min using transfer buffer composed as follows: 192mM glycine, 25mM Tris-base, 20% Methanol. Subsequently, transfer of the proteins to the PVDF membrane was controlled by Ponceau-S staining (2% Ponceau S, 30% trichloroacetic acid, 30% sulfosalicylic acid). After destaining in TBS-T buffer (20mM Tris-HCL, pH 7.5, 150mM NaCl, 0,1% Tween 20(v/v)), membranes were blocked (10% non-fat dry milk in TBS-T buffer) for at least 1h at RT and then probed either for 1h at RT or overnight at 4°C with the primary antibody directed against the protein of interest in TBS-T containing 5% dry milk. Blots were washed three times for 10min in TBS-T and then incubated for 45-60min at RT with the respective secondary antibody conjugated to horseradish peroxidase (HRP). Membranes were extensively washed in TBS-T and proteins were detected using ECL reagent (Amersham) that resulting in chemiluminescence signals, which can be captured by exposure to photosensitive ECL hyperfilms (Amersham) or Lumi-Imager F1 (Roche). For reprobing of the membranes, bound antibodies were removed through incubation in stripping solution (Chemicon) for 10-20min.

II.2.5 Immunoprecipitation (IP)

Transfected HEK 293T cells were cultured in six-well plate. 48h post transfection cells from one plate were scraped into the supernatant, pooled and washed three-times with PBS. After the final wash, cells were resuspended in buffer A containing 20mM HEPES-KOH (pH 7.9), 10mM $MgCl_2$, 150mM NaCl, 10mM KCL, 100mM EDTA, 1mM DTT, 10% glycerol (v/v), 0,5 μ M Pepstatin, 0,1mM PMSF, 0,1mM natrium-vanadate, 0,5% Nonidet P-40 (IGEPAL) and complete protease inhibitor mixture (Roche). Cell were lysed for 30min at 4°C and centrifuged for 30min at 4°C and 13.500rpm (Eppendorf microcentrifuge). The supernatant was collected, mixed with 30-40 μ l of protein-G-sepharose (PGS) equilibrated in buffer A and incubated for 1h at 4°C – constantly rotating – to remove cellular proteins unspecifically binding to the beads. After centrifugation, supernatant was used as input material for immunoprecipitation. 1-2 μ g of antibody were added to the samples, followed by incubation for 1-2h at 4°C on a rotator. Subsequently, input material was incubated with PGS for 30-45min at 4°C. Following six washes (A-B-C-C-B-A; buffer A: 150mM NaCl, buffer B: 200mM NaCl and buffer C: 400mM) protein complexes were eluted from the beads with 2xSDS sample buffer and analyzed by SDS-PAGE and western blotting.

II.2.6 Reverse transcriptase (RT)-PCR analysis

II.2.6.1 RNA extraction

Total cellular RNA from transfected cells was isolated using an acid phenol/chloroform-guanidium thiocyanate based extraction protocol [(92); reviewed in (93)]. Cells grown in 6-well culture plates were washed two times with 2ml ice-cold PBS and lysed in 500 μ l Solution D (4M guanidium isothiocyanate, 25mM sodium citrate, pH 7, 0.5% (v/v) sodium N-laurylsarcose, 0.1M β -mercaptoethanol). Lysates were collected from the wells using a cell scraper (Nunc) and transferred into 1.5ml reaction tubes (Eppendorf). RNA was extracted by addition of 7.2 μ l β -mercaptoethanol, 0.1 vol. 2M sodium acetate (pH4), 1 vol. phenol (pH 4, Roth) and 0.2 vol. chloroform/isoamyl alcohol (24:1). After shaking vigorously for 15sec, the

samples were chilled on ice for 15min and centrifuged for 20min at 10.000rpm and 4°C (Eppendorf microcentrifuge). RNA partitions into the upper aqueous phase, which was therefore transferred into a fresh 1.5ml reaction tube and mixed up with 1 vol. isopropanol for overnight precipitation at -20°C. Samples were pelleted by centrifugation for 20min at 12.000rpm and 4°C (Eppendorf microcentrifuge), washed twice with 120µl ethanol (70% (v/v)) each, air-dried and resolved in 10µl RNase-free ddH₂O. RNA samples could be stored at -80°C.

II.2.6.2 Reverse Transcription

For reverse transcription 4µg of total-RNA were subjected to DNA digestion with 10U of recombinant RNase-free DNase I (Roche) in a total reaction volume of 10µl at 37°C for 20min. During this step contaminating traces of plasmid DNA were removed. DNase I was heat-inactivated at 70°C for 5min and 4,5µl of the DNase, treated RNA were used for subsequent cDNA synthesis. 1µl of oligo-d(T)₁₅ (7.5M, Roche), 4µl of deoxyribonucleoside triphosphates (2.5mM each, Qiagen) and 3.5µl of RNase-free water were added to the DNase-digested RNA and incubated for 5min at 65°C. After cooling down on ice for 1min, 4µl of 5xfirst strand buffer (FSB, Invitrogen), 1µl 0.1M DTT, 1µl (40U) RNasin (Promega) and 1µl (200U) of SuperScriptTM III RNase H-deficient reverse transcriptase were added to the samples. Reverse transcription was performed for 1h at 50°C and 15min at 72°C. As a control for plasmid DNA contaminations an identical reaction without reverse transcriptase was prepared for each sample.

II.2.6.3 Semi-quantitative PCR analysis

3µl of the prepared cDNA were used as a template for specific PCR reactions. Amplification was performed in a total reaction volume of 50µl containing 5µl 10x PCR buffer (10mM Tris-HCL, pH 8.3, 50mM KCL, 0.15mM MgCl₂, 0.001% (w/v) gelatin), 4µl dNTPs (2.5 mM each, Qiagen), 1µl of each primer (10pmol/µl) and 0.25µl AmpiTaq DNA polymerase (5U/µl, Applied Biosystems) in a Robocycler

Gradient 96 PCR cycler (Stratagene). The exponential phase of each of the PCR reactions was determined in preceding experiments.

Target RNA	Primer	amplicon size (in bp)	Annealing temperature (T _A in °C)	Amplification cycles
SV-env/eGFP ¹	#3210/#3211 (s)	165 – 235	53	24-26
	SD1up	138	53	30
	SD1down	193-221	53	
LTR ex2 ex3 ²	#1544/#2588	267 – 1059	53	28-30
pNL4-3 ³	#1544/#3632 (E1/E4)	186-1050	53	35
	#1544/#3392 (E1/E7)	212-867	53	35
	#1544/#640 (E1/I4)	602-1230	53	35
GH1 (exon 3-5)	#1224/#1225	187	56	29

Tab.II-8: Amplicon sizes and semi-quantitative PCR conditions

¹ see Appendix Fig.VII-1.1; ² see Appendix Fig.VII-1.4; ³ see Appendix Fig.VII-1.6

II.2.6.4 Quantitative real-time PCR analysis

Quantitative SYBR green PCR (qPCR) analysis of the reporter mRNAs was carried out utilizing respective primer pairs, which are given in the table below. Samples were prepared corresponding to the instructions obtained from the manufacturer. cDNA samples were diluted 1:10 with PCR-grade ddH₂O. 2µl of cDNA were added to a prepared reaction mix within pre-cooled glass 20µl capillaries (Roche) for the Light Cycler[®] carousel-based detection system and contained the following components: 2µl ready-to-use 10x concentrated SYBR green I DNA master mix (SYBR green I dye, Taq DNA polymerase, dNTP mix, 10mM MgCl₂), 1µl of each primer (5pmol/µl, HPCL-grade), 1,2µl MgCl₂ (25mM MgCl₂) and 13,8µl PCR-grade ddH₂O. Each Capillary was sealed with a stopper and carefully loaded into the Light Cycler[®] sample carousel.

Fluorescence emission was read by a LightCycler1.5 (Roche) using the following standard program:

Analysis mode	Cycles	Hold Time	Target Temperature	Target Temperature
None	1	30 sec	95°C	Denaturation
Quantification	45 cycles	0 sec	95°C	Denaturation
		4-6 sec	T _A °C	primer annealing
		(amplicon (bp)/25)sec	72°C	Extension
Melting Curve	1	0 sec	95°C	Denaturation
		15 sec	65°C	Annealing
		Ramp rate = 0.1°C/sec	95°C	Melting
None		30 sec	40°C	Cooling

Tab.II-9: Standard program for quantitative real-time PCR analysis

Parameters were specifically programmed for the respective cDNA target and primers, which were used for amplification. Fluorescence acquisition occurred at each cycle during the extension step. Analysis of the melting curve allowed the identification of the PCR product by the unique temperature it is denaturing and the fluorescence is set free. Additionally, qPCR products were controlled by native gel electrophoresis (see chapter II.2.3.5.3). Relative quantities of the target cDNA were calculated using the $2^{-\Delta\Delta CT}$ method, where CT (cycle threshold) represents the value at which fluorescence surpasses a set baseline (315).

Target RNA	Primer	amplicon size (in bp)	Annealing temperature (T _A in °C)	Extension time
SV-env/eGFP ¹ (spliced)	#3239/#3240	165 – 230*	56	6
	#3832/#3240	104	56	
Viral mRNAs				
All viral RNAs	#3323/#3324	122	56	8
	#3387/#3388	153	56	10
Tat1	#3631/#3632	104	54	6
Tat3	#3397/#3633	95	54	6
Gag/pol (us)	#3389/#3390	143	53	9
Multiply spliced	#3391/#3392	133	56	8
Vif	#3395/#3396	181	54	10
Vpr	#3397/#3398	161	54	10
Nef4/Env8	#3397/#3636	93	53	6
Nef2/Env1	#3629/#3637	67	53	5
Controls				
GH1 (exon 3-5)	#1224/#1225	187	56	10

Tab.II-10: Amplicon sizes and quantitative PCR conditions

¹ see Appendix Fig.VII-1.1

II.2.6.5 Native gel electrophoresis and EtBr staining to visualize RT-PCR products

10µl of the RT-PCR products were separated on 6-10% non-denaturing polyacrylamide (PAA) gels (Rotiphorese Gel 30, Roth) using 1xTBE running buffer. Gels were run at 200V dependent on the percentage for 1h up to 2h, stained with ethidium bromide (EtBr, 4µg/ml in 1xTBE) for 5-10min and exposed to UV light excitation in the Lumi-Imager F1 (Roche).

II.2.6.6 Purification of RT-PCR products from native polyacrylamide gels (PAA)

RT-PCR products visualized by EtBr-staining were cut out from the gels using long wave UV light (320nm) and diced into small pieces before transfer into a 1.5ml reaction tube. DNA was eluted from the gel by overnight incubation at 37°C in PAA elution buffer (0.5M sodium acetate, 0.1% (w/v) SDS, 1mM EDTA). After centrifugation for 1min at full-speed and 4°C (Eppendorf microcentrifuge) supernatant was removed into a new 1.5ml reaction tube and gel pieces once more mixed with 0.5 vol. PAA elution buffer. After centrifugation, both supernatants were pooled and purified from gel leftovers through filtration using glass fibre filters (GF/C filter, Whatman). DNA was precipitated by addition of 2 vol. ethanol (96%) on ice for 30min followed by centrifugation for 10min at full-speed and 4°C (Eppendorf microcentrifuge). Pellets were resuspended in 200µl TE (pH 8) and 25µl 3M sodium acetate (pH 5.2). Precipitation with 2 vol. ethanol (96%) was repeated and after a new centrifugation DNA pellets were washed with 120µl ethanol (70% (v/v)), air-dried and resuspended in 10µl ddH₂O. DNA was used as template for re-amplification by a PCR reaction using proofreading Pwo DNA polymerase (Roche) and purified by phenol/chloroform extraction. After concentration was photometrically measured, 10 up to 50ng of DNA were given away for sequencing to confirm the individual splicing events (II.2.1.5).

II.2.7 RNA affinity chromatography

II.2.7.1 *In vitro* transcription

II.2.7.1.1 *In vitro* transcription of small size templates (< 60 nt)

100µM of DNA Oligonucleotides containing the T7 promotor sequence in the 5'-end followed by the target sequence in reverse complementary orientation, respectively, were purified from degradation products by separation in 15% polyacrylamide gels using 8M urea containing loading buffer to denature secondary structures. After electrophoresis, DNA was detected by UV shadowing (320nm) and fractions of full-length Oligonucleotides were cut out from the gel. Diced pieces were transferred into

1.5ml reaction tubes and incubated in 600µl PAA elution buffer (0.5M sodium acetate, 0.1% (w/v) SDS, 1mM EDTA) overnight at 4°C, constantly rotating. After addition of 0.1 vol. 3M sodium acetate (pH 5), 1 vol. phenol and 0.2 vol. chloroform/isoamyl alcohol (24:1), samples were centrifuged for 5min at full speed and 4°C (Eppendorf microcentrifuge) and DNA in the aqueous phase precipitated with 1ml ethanol (96%) on dry ice for 5min. Oligonucleotides were sedimented by centrifugation (full speed, 30min at 4°C, Eppendorf microcentrifuge), air-dried and resolved in 50µl RNase-free ddH₂O. Concentration of the DNA was photometrically quantitated. For *in vitro* transcription 500pmol of the specific oligonucleotides and the T7 primer were adjusted to a total volume of 500µl with RNase-free ddH₂O. Samples were denaturated for 5min at 90°C to melt secondary structures followed by annealing at RT for 5min. For *in vitro* transcription an 1ml reaction with 500pmol of pre-annealed oligonucleotides, 50mM Tris-HCL (pH 7.5), 15mM MgCl₂, 5mM DTT, 5mM rNTPs (containing dUTP instead of dTTP, pH 8, Sigma), 2mM spermidine and 70µl T7 RNA polymerase were separated into 500µl aliquots and incubated for 4-6h at 37°C. RNA was precipitated by addition of 1ml ethanol (96%) at -80°C for 5min. Subsequently, RNA was pelleted by centrifugation for 10min at full-speed and 4°C (Eppendorf microcentrifuge). For purification of the full-length *in vitro* transcribed RNAs, pellets were resolved in 200µl 8M urea/bromophenolblue (BB) loading buffer and gelelectrophoresed in 15% polyacrylamide gels (300V, 2-3h). Upper bands representing full-length, non-degraded RNA were cut out from the gel and transferred into a 15ml reaction tube for overnight elution at 4°C in 3 ml PAA elution buffer (0.5M sodium acetate, 0.1% (w/v) SDS, 1mM EDTA). On the next day, RNA was purified by addition of 0.1 vol. 3M sodium acetate (pH 5), 1 vol. acetic phenol (pH 4) and 0.2 vol. chloroform/isoamyl alcohol (24:1). After centrifugation for 10 min at 4.000 rpm and 4°C (Eppendorf 5810 R), supernatant with the RNA was precipitated with 6 ml ethanol (96%) at -80°C for 5min and sedimented (4.000rpm, 4°C, 45min, Eppendorf 5810R). RNA pellets were air-dried, resolved in 50µl RNase-free ddH₂O and subsequently, the concentration was measured by photometry. RNAs were stored at -80°C or immediately used for RNA affinity chromatography.

II.2.7.1.2 *In vitro* transcription of larger size templates (< 60 nt)

Larger templates for *in vitro* transcription were amplified from plasmid DNA by standard PCR using proofreading Pwo DNA polymerase (Roche). T7 promotor sequence and MS2 RNA hairpin sequence were placed into the 5'-end of the forward primer. By use of specific primers the desired sequences were amplified at high cycle numbers to produce large amounts of DNA template. PCR products were purified by phenol/chloroform extraction. For RNA synthesis a RiboMaxTM large scale RNA production system (P1300, Promega) was applied following the instructions provided by the manufacturer. 1-2µg of DNA template were mixed with 4 µl 5xT7 transcription buffer (400mM HEPES-KOH, pH 7.5, 120mM MgCl₂, 10mM spermidine, 200mM DTT), 6µl rNTPs (25mM each) and 2µl T7 enzyme mix and adjusted to a final reaction volume of 20µl with nuclease-free ddH₂O. After incubation at 37°C for 2-4h DNA template was removed by DNase digestion (1U RQ1 RNase-free DNase per µg DNA template for 15min at 37°). *In vitro* transcribed RNA was extracted using 0.5 vol. phenol (pH 4.7) and 0.5 vol. chloroform/isoamyl alcohol (24:1). After addition samples were vigorously vortexed for at least 1min and centrifuged at full-speed for 2min (Eppendorf microcentrifuge). The upper, aqueous phase was transferred to a fresh reaction tube and mixed with 1 vol. chloroform/isoamyl alcohol (24:1). Samples were again vortexed for 1min and centrifuged at full-speed for 2min (Eppendorf microcentrifuge). The supernatant was pipetted into a fresh reaction tube and 0.1 vol. 3M sodium acetate (pH 5.2) and 2.5 vol. ethanol (96%) were added. Transcripts were precipitated for 2-5min on ice before centrifugation at full-speed for 10min. Finally, pellets were washed with 200µl of ethanol (70% (v/v)), dried and suspended 30µl in RNase-free ddH₂O. RNA quantities were determined by photometry and the quality of the preparations was checked by gel electrophoresis in 1% agarose gels (LE, Biozym).

II.2.7.2 Protein isolation by RNA affinity chromatography

500-2000pmol of *in vitro* transcribed RNA were adjusted to a total volume of 340µl with RNase-free ddH₂O and chemically activated for 1h at RT in the dark using a 20µl of a saturated sodium-meta-periodate solution and 40µl 1M sodium acetate (pH

5). Subsequently, RNAs were precipitated with 0.2 vol. sodium acetate (pH 5) and 2.5 vol. ethanol (96%) at -80°C for exactly 5min and pelleted by centrifugation at full-speed and 4°C for 30min. Meanwhile 125µl of an adipic acid-dihydrazide-agarose bead suspension (Sigma) for each sample were washed four times with 0.1M sodium acetate (pH5) (300rpm, 4°C, 3min, Eppendorf 5810R). After the final wash, bead suspension was adjusted to 1ml per sample with 0.1M sodium acetate (pH 5) and added to the RNA for immobilization at 4°C overnight on rotator. Uncoupled RNA was removed by two washing steps with 1ml 2M sodium chloride (800rpm, 2-3min, Eppendorf microcentrifuge). Subsequently, immobilized RNA was equilibrated to the nuclear salt conditions by washing three times with 1ml buffer D (20mM HEPES-KOH, pH7.9, 5% (v/v) glycerol, 0.1M KCL, 0.2mM EDTA, 0.5mM (DTT). HeLa nuclear extract (Cilbiotech) was diluted to concentration varying from 5% to 50% with buffer D and where appropriate supplemented with 1.2µg of recombinant M2 coat protein (kindly provided by Dr. K. Hertel) per sample. 200µl of diluted nuclear extract were added to the immobilized RNA and incubated for 10-30min at 30°C on a rotator. Proteases and phosphates were inhibited through addition of complete protease inhibitors (Roche) and okadaic acid (OA), respectively. To get rid of unbound proteins, samples were washed five times with 1ml buffer D containing 4mM magnesium chloride (800rpm, 2-3min, Eppendorf microcentrifuge). After final wash bead pellet was resuspended in equal volume of 2xprotein sample buffer (0.75M Tris-HCL, pH 6.8, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol, 4% (w/v) SDS) and proteins were eluted from the RNA by heating to 95°C for 10min. Agarose beads were pelleted by centrifugation and the supernatant either immediately loaded on a SDS polyacrylamide gel (II.2.3.2) or transferred into a fresh 1.5ml reaction tube for storage at -20°C.

II.2.7.3 Oligo-directed RNase H depletion of U1 snRNA 5'-ends

To remove the 5'-ends of the U1 snRNA 185µl of HeLa nuclear extract (Cilbiotech) were incubated for 1h at 30°C in a reaction containing 3µl 100mM ATP, 3µl 500mM creatine phosphate (CP), 4µl 80mM magnesium chloride, 2µl RNase H (NEB), 4µl 2mM anti-U1 DNA oligo (#19967936). As control, HeLa nuclear extract was incubated under similar conditions with exception of the anti-U1 DNA oligo.

II.2.7.4 Protein sequencing by mass spectrometry

II.2.7.4.1 Coomassie Staining

Gels containing the electrophoresed protein samples from RNA affinity chromatography assays were either subjected to western blot analysis with specific antibodies or directly stained with Coomassie blue to identify differences within the binding profile of nuclear proteins to the target sequences. Gels were incubated at least for 2-3h in coomassie blue staining solution (0,02% Coomassie Brilliant Blue G250, 2% (w/v) phosphoric acid, 5% aluminium sulphate, 10% ethanol) and then destained with ddH₂O until the background became clear.

II.2.7.4.2 In gel digestion and sample preparation

Bands containing proteins, which should be identified, were cut out from the Coomassie stained gels, reduced to smaller pieces and given into a fresh 0.5ml reaction tube (Protein-Low-Bind, Eppendorf). Samples were either directly sent to the bioanalytic core unit of the center for molecular medicine (CMMC, University of Cologne) for mass spectrometrical (MS) analysis or further processed for MS analysis at the Analytical Core Facility of the Biological-Medical Research Centre (BMFZ, HHU). Latter samples were mixed four times in 100µl 25mM ammonium hydrogen carbonate buffer/50% acetonitrile and incubated for 30min at RT to remove salts. Subsequently gel pieces were entirely dehydrated by incubation in acetonitrile (100%) for 30min and after carefully removing the liquid dried in a vacuum centrifuge (DNA110 SpeedVac®, Thermo Scientific). Gel pieces were rehydrated in trypsin solution (0.1µg/µl (Sigma) in 25mM ammonium carbonate buffer, pH 8), liquid not soaked into the gel was promptly removed and 25mM ammonium carbonate buffer (pH 8) added. Proteins were in gel digested for 12-16h at 37°C before the supernatant was transferred into a new 0.5ml reaction tube (Protein-Low-Bind, Eppendorf). Gel pieces were shaken in 2 vol. ddH₂O for 5min and sonicated before the supernatant was pooled with the supernatant from the preceding step. Gel pieces were mixed with 1 vol. elution buffer (50% acetonitrile, 5% formic acid) and incubated for 30min at RT. This step was repeated two times, and each supernatant was given

to the supernatants combined before. Finally, gel pieces were treated with 1 vol. acetonitrile (100%) for 30min and the resulting eluate was also pooled with the supernatant from the prior steps. Subsequently, the eluted proteins were lyophilised in a vacuum centrifuge (DNA110 SpeedVac[®], Thermo Scientific) and sent to mass spectrometrical (MS) analysis, which was performed by Dr. W. Bouschen using an ESIQuadrupole-TOF (QSTAR XL, Applied Biosystems) at the Analytical Core Facility of the Biological-Medical Research Centre (BMFZ, HHUD).

II.2.8 *In vitro* splicing

In vitro splicing experiments were performed in 30% RAC-depleted HeLa nuclear extracts with T7 polymerase synthesized ³²P-labeled β -globin derived transcripts as a template, 1mM ATP, 20mM creatine phosphate and 3.2mM MgCl₂. Reaction mixtures were incubated at 30°C. Range of reaction time points is defined by the numbers indicated above the gel. After splicing reaction, samples were digested with proteinase K to remove associated protein complexes from the substrate, RNA was phenol-chloroform extracted, precipitated by use of ethanol (96%) and resolved in urea/polyacrylamide gels. Radioactivity was read by PhosphorImager.

III. Results

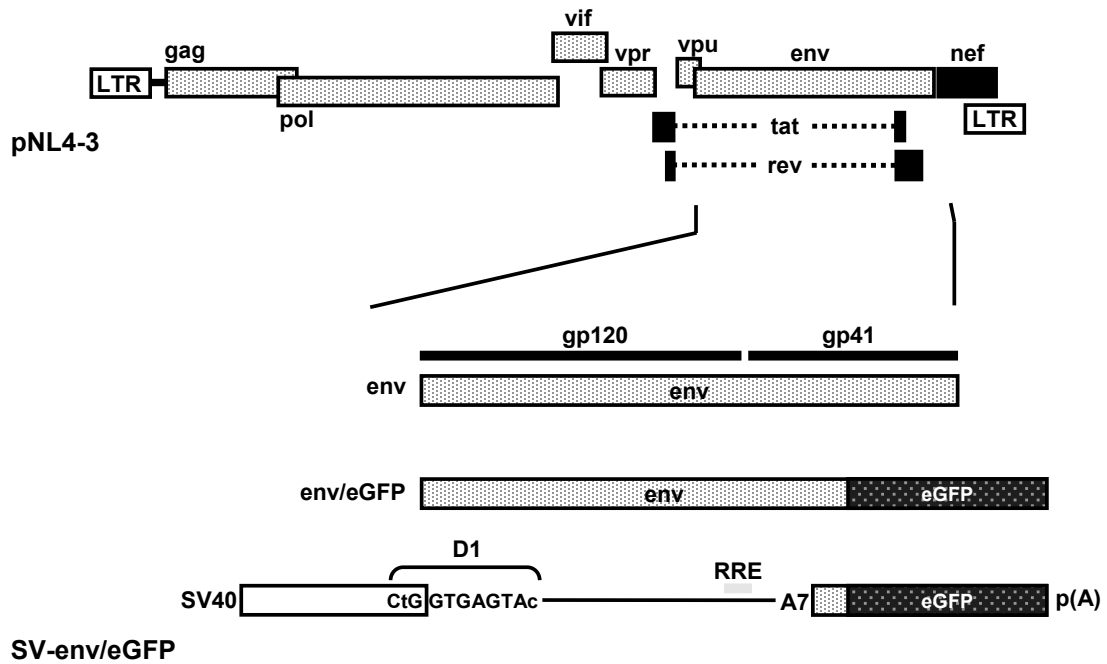
III.1 “Don’t buy the house, buy the neighbourhood”: relative positioning determines the phenotype of a splicing regulatory protein

Splicing regulatory elements (SREs) are extensively analyzed “landing sites” for splicing regulatory proteins, which can either enhance or repress recruitment of general splicing components to a proximal splice site [reviewed in (310, 561)]. Studies have well established the role of SR protein binding to exonic SREs for enhancement of U1 snRNP interactions with a neighbouring 5’ss (90, 164, 276, 589). However, in a few studies SR proteins bound to intronic positions were shown to inhibit splicing (58, 120, 213, 229, 256, 426, 483). Whereas the modes of activation through which exonic bound SR proteins act to facilitate U1 snRNP binding are quite well elucidated (see I.3.5.1 SR proteins), the mechanisms underlying splicing repression by intronic positioned SR proteins are yet poorly understood. However, in contrast to SR proteins, intronic binding of hnRNP-like proteins has been shown to assist recruitment of U1 snRNP to the 5’ss (see I.3.5.2 hnRNP proteins and I.3.5.3 hnRNP-like proteins). This is exemplified by the splicing factor TIA-1 interacting with the U1-specific protein C to activate 5’ss, which are otherwise rarely recognized (121, 157). Much interest has been generated to understand how SREs can activate splicing in one case, and repress it in another. Recently defined “RNA splicing maps” could relate the RNA binding positions of individual splicing factors across the genome to their positive or negative outcome for splice site selection [reviewed in (585)]. Other comprehensive studies made similar observations for the other splicing factors such as the hnRNP protein PTB (316) or members of the Fox1/2 protein family (614) and supported the notion that the relative location of an SRE to a 5’ss may decide about its mode of action. Based on these findings, systematic analyses were carried out to address the question of whether the position of an SRE relative to the 5’ss determines the splicing outcome.

III.1.1 The position of splicing regulatory elements relative to a 5'ss determines the splicing outcome

To determine if SREs act in a position-dependent manner, an HIV-1 based splicing reporter was used, which derives from the 3' part of the proviral genome of the viral isolate NL4-3 (GenBank Accession No. [M19921](#)) and by this means contains the coding sequences for the viral glycoprotein (Env) (Fig.III-1A). The transcriptional unit is under the control of the simian virus 40 (SV40) early promoter. An efficient transcription termination is provided by the early polyadenylation signal of SV40. Because eGFP (*enhanced green fluorescent protein*) coding sequences have been placed downstream of the splice acceptor A7, it was possible to detect eGFP either from a spliced mRNA [Fig.III-1B, (a) and (b)] or – in the presence of the HIV-1 regulatory protein Rev – as part of an Env/eGFP fusion protein from an unspliced U1 snRNA-dependent mRNA [Fig.III-1B, (c) and (d)]. Binding of the U1 snRNP to the 5' splice site (5'ss) D1 - independent of its splicing function – protects this unstable reporter mRNA from degradation (253). Additionally, recruitment of the U1 snRNP to D1 has been shown to be dependent on a purine-rich exonic splicing element, termed GAR (11, 72, 253). This SRE-dependent recognition of D1 allowed to substitute the GAR element with a so-called “neutral” sequence and then place SREs either into the exon or intron, respectively, and measure if they act on splicing in a position-dependent manner. In the first set of experiments high affinity binding sites for the SR protein SRSF7 and hnRNP-like protein TIA-1 were tested on each side of 5'ss D1 for splicing activation (Fig.III-2A). As controls, two splicing almost neutral sequences with slightly different residual activities referred to as “neutral 1” (N1) [GAR⁻ESE⁻ (11)] and “neutral 2” (N2) [GAR⁻ (253)] were placed into the upstream exon of 5'ss D1 (Fig. III-2A). The SRE activities of these “neutral” sequences have also been compared in a third study by K. Schöneweis (466). SREs analyzed for their ability to promote 5'ss D1 use from an intronic position contained control sequence “N2” within the upstream exon. HeLa-T4⁺ cells were used for transient transfections with each of the constructs and 24h later the cells were analyzed under the fluorescence microscope to determine their level of eGFP expression as an read-out for the SRE activity.

A



B

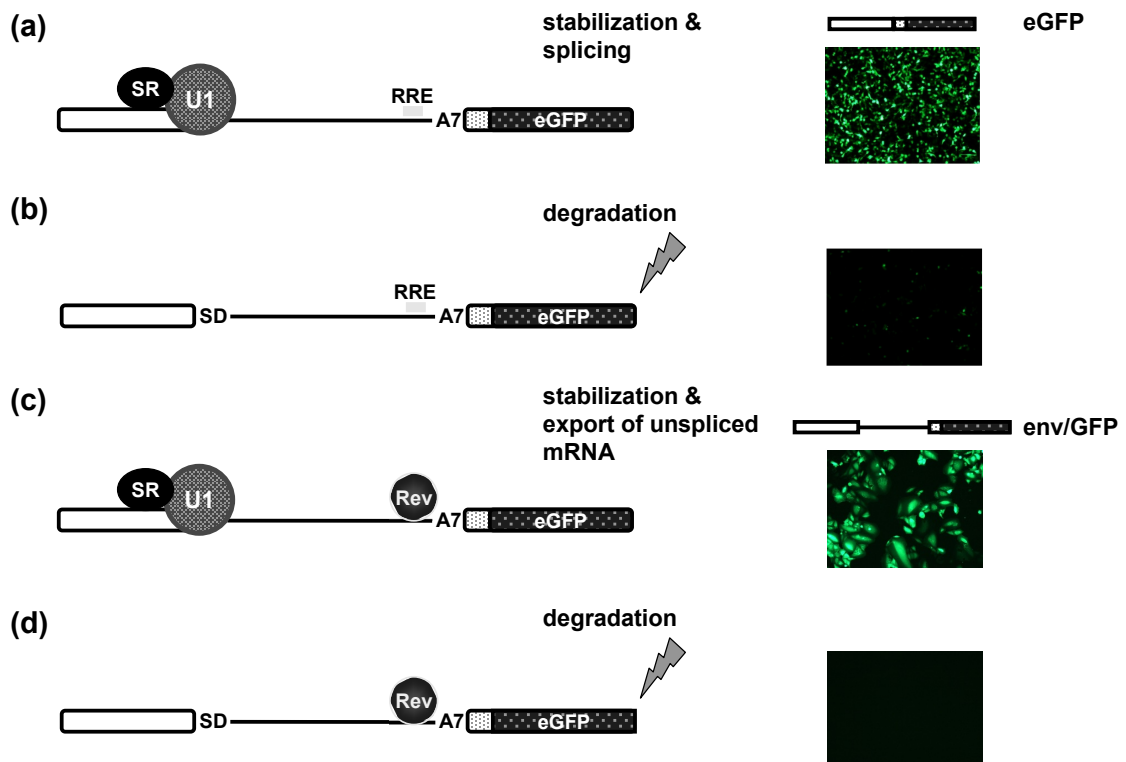


Fig.III-1: The subgenomic HIV-1 based one-intron splicing reporter to analyse the position-dependence of splicing regulatory elements on 5'ss use. [continued on next page]

Surprisingly, efficient eGFP expression was found to be totally dependent on the positioning of the tested SREs relative to the 5'ss. SRSF7 enhanced splicing exclusively when its binding sites were placed within the upstream exon, but failed to activate D1 after relocation to the opposite side (Fig.III-2B), while TIA-1 showed a “mirror-inverted” splicing phenotype, requiring location within the downstream intron. As expected, none of the control “neutral” sequences induced efficient eGFP expression. To confirm these position-dependent results and to (semi-)quantify the SRE activities the transfected cells were analyzed by flow cytometry (Fig.III-2C). Constitutively coexpressed dTomato was used to monitor the transfection efficiencies. Each histogram summarizes the eGFP expression of the various constructs as a bold blue line relative to the controls “neutral 1” (N1) (grey area) and “neutral 2” (N2) (bold black line). In agreement with microscopic inspection of the cells, SRSF7 enhanced eGFP expression exclusively from the exonic side, while TIA-1 required the opposite position within the downstream intron to induce efficient eGFP expression (Fig.III-2C).

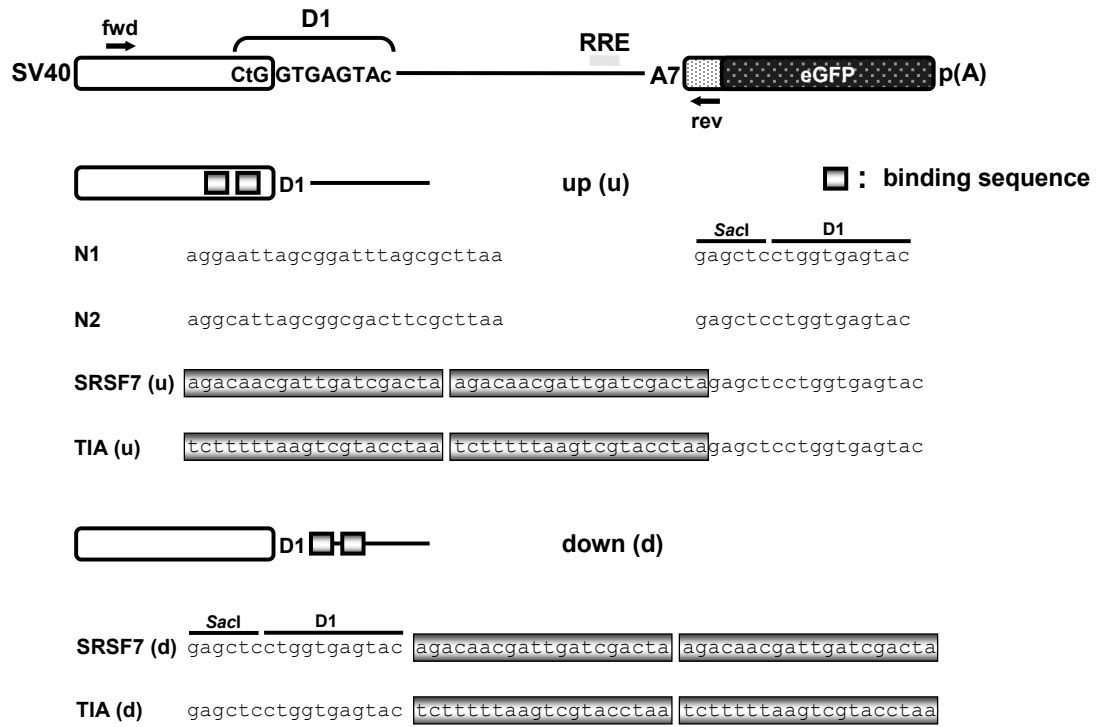
Fig.III-1: continued

(A) Schematic diagram of the RNA splicing reporter SV-env/eGFP. The HIV-1 derived construct contains the coding sequence for the viral glycoprotein (Env) derived from the viral laboratory clone pNL4-3 (GenBank Accession No. [M19921](#)). The *Bam*HI-*Xho*I 3'-fragment encoding the cytoplasmatic moiety of the gp41 subunit was replaced by the open reading frame (ORF) for eGFP generating an env/eGFP fusion construct. The *env* overlapping *vpu* open reading frame has been deleted to allow exclusive expression of eGFP or env/eGFP fusion proteins. Which one of both is expressed depends on splicing of the internal *env*-intron that is defined by the viral 5'ss D1 and the 3'ss A7. Expression of the transcriptional unit is driven by a simian virus 40 early promotor (SV40). [LTR: Long Terminal Repeat; SV40: Simian Virus 40, gp120: glykoprotein 120kDa subunit; gp41: glykoprotein 41kDa subunit ; RRE: Rev Responsive Element, p(A): Polyadenylation site].

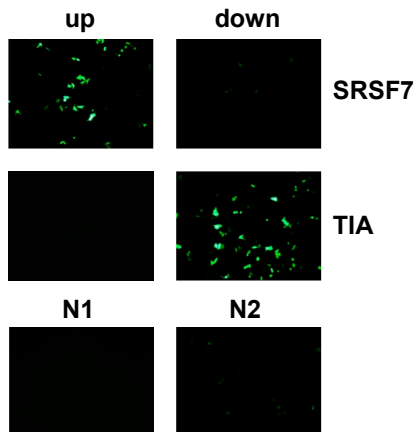
(B) Enhancer-mediated U1 snRNP recruitment to the 5'ss D1 protects the inherent unstable reporter mRNAs from degradation. Dependent on the absence or presence of the viral Rev protein recognizing the RRE within the intron, binding of the U1 snRNP either induces the expression of eGFP from spliced mRNA (a) or an env/eGFP fusion protein (c) from unspliced mRNA. The latter is expressed on the cell surface and can be indirectly detected by formation of fluorescent fusion cells (syncytia) in the context of a CD4-receptor expressing HeLa cell line (HeLa-T4⁺).

Results

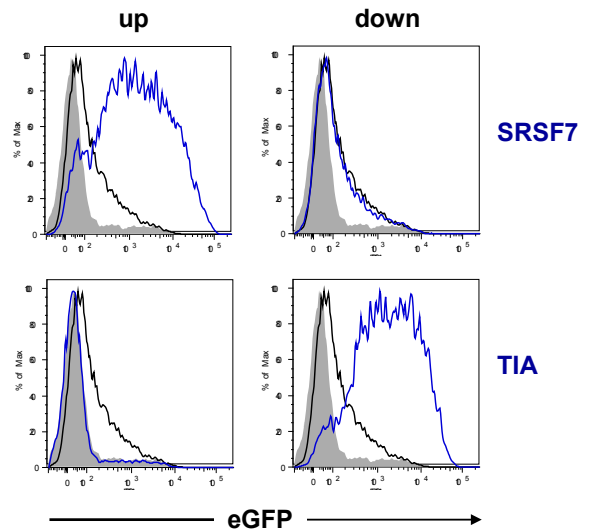
A



B



C



D

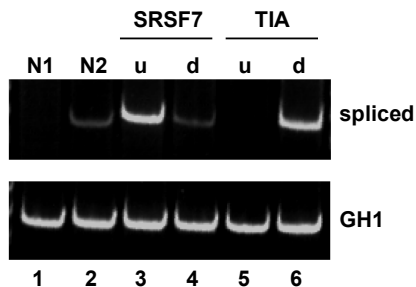


Fig.III-2: The position of high affinity binding sites for SRSF7 and TIA-1 relative to 5'ss D1 determines whether it is used or not. [continued on next page]

Semi-quantitative RT-PCR analysis (Fig.III-2D) further confirmed the results obtained from the flow cytometry and revealed that both SREs activated D1, but only from one side relative to the 5'ss and in a “mirror-inverted” manner (Fig.III-2D, cf. lanes 3-6). For normalization of the transfection efficiency pXGH5 was cotransfected and the level of constitutively spliced GH1 mRNA detected for each sample.

To substantiate these findings, and due to the difficulty to define a sequence as “neutral” (see Discussion) a further control sequence [“neutral 3” (N3); Tab.III-1, (616)] was inserted immediately upstream or downstream of D1 (Fig.III-3A). Transient transfection experiments of HeLa-T4⁺ cells, a cell line stably expressing the human T cell CD4 receptor on its surface, allowed a microscopic read out for fluorescent syncytia formed by interaction between the env/eGFP fusion protein and the CD4 receptor of the neighbouring cells. Thus, HeLa-T4⁺ cells were transiently transfected with each one of the splicing reporters together with a Rev-expressing vector (SVcrev).

Fig.III-2: continued

(A) Schematic drawing of the RNA splicing reporter SV-env/eGFP (upper panel). Primers used for RT-PCR analyses are indicated by arrows. Splicing almost neutral sequences [N1 and N2, (11)] and a two copies of a high affinity binding site for the SR protein SRSF7 (79) and the hnRNP-like protein TIA-1 (123) were each placed immediately upstream of 5'ss D1 (middle panel). Additionally, binding sites were inserted into the intron – immediately downstream of exonic sequence N2 and 5'ss D1 (lower panel). [(u): upstream; (d): downstream; N1: neutral 1; N2: neutral 2. **(B)** 2.5 x 10⁵ HeLa-T4⁺ cells were transiently transfected with 1µg of each splicing reporter and 1µg of pCL-dTom expressing fluorescent dTomato to monitor transfection efficiency. 24h following transfection, eGFP expression was detected by fluorescence microscopy. **(C)** Cells from **(B)** were subsequently analyzed by flow cytometrical analyses. The mean fluorescence intensity was calculated by FlowJo software. **(D)** 2.5 x 10⁵ HeLa-T4⁺ cells were cotransfected with 1µg of each construct and 1µg of pXGH5 expressing constitutively spliced GH1 mRNA to monitor both transfection efficiency and viability of the cellular splicing apparatus. 30h later total-RNA samples were collected from the transfected cells. After DNase I digestion to remove traces of plasmids, cDNA was synthesized using Oligo d(T) primer and SuperScriptTMIII revers transcriptase. Subsequently, PCR reactions were performed to specifically amplify spliced reporter mRNAs (#3210/#3211) and GH1-mRNA (#1224/#1225). PCR products were separated in 8% non-denaturing polyacrylamide gels and visualized by ethidium bromide staining. [u: upstream; d: downstream; GH1: human growth hormone 1]

In agreement with the preceding experiments, fluorescence microscopy revealed that green syncytia could only be observed when SRSF7 was placed upstream or when TIA-1 was placed downstream of D1 (Fig.III-3B). Again, the control sequence (N3) did not promote splicing regardless of its exonic or intronic position. The experiment was repeated in absence of Rev and the transiently transfected HeLa-T4⁺ cells were analyzed by semi-quantitative RT-PCR and immunoblotting (Fig.III-3C). Results obtained from the semi-quantitative RT-PCR analysis were in agreement with the results obtained by fluorescence microscopy and efficient splicing required either positioning of SRSF7 within the upstream exon or that of TIA-1 within the downstream intron (Fig.III-3C). Relocation of the SREs from their activating positions to their opposing sides abrogated their activating splicing phenotype. Immunoblot analyses of proteins extracted from the transiently transfected HeLa-T4⁺ cells reiterated this finding with considerable amounts of eGFP protein only detected for SRSF7 upstream and TIA-1 downstream of the 5'ss (Fig.III-3C, lanes 3 and 6). Finally, realtime PCR assays for quantitation of the relative levels of spliced eGFP-mRNA revealed an up to 100-fold increase in splicing for the SREs in their activating positions (Fig.III-3D, lanes 3 and 6). In addition, qPCRs indicated that both SREs were not only splicing neutral outside of their activating positions, but even tend to reduce 5'ss recognition (Fig.III-3D, cf. lanes 1 and 4 or lanes 2 and 5).

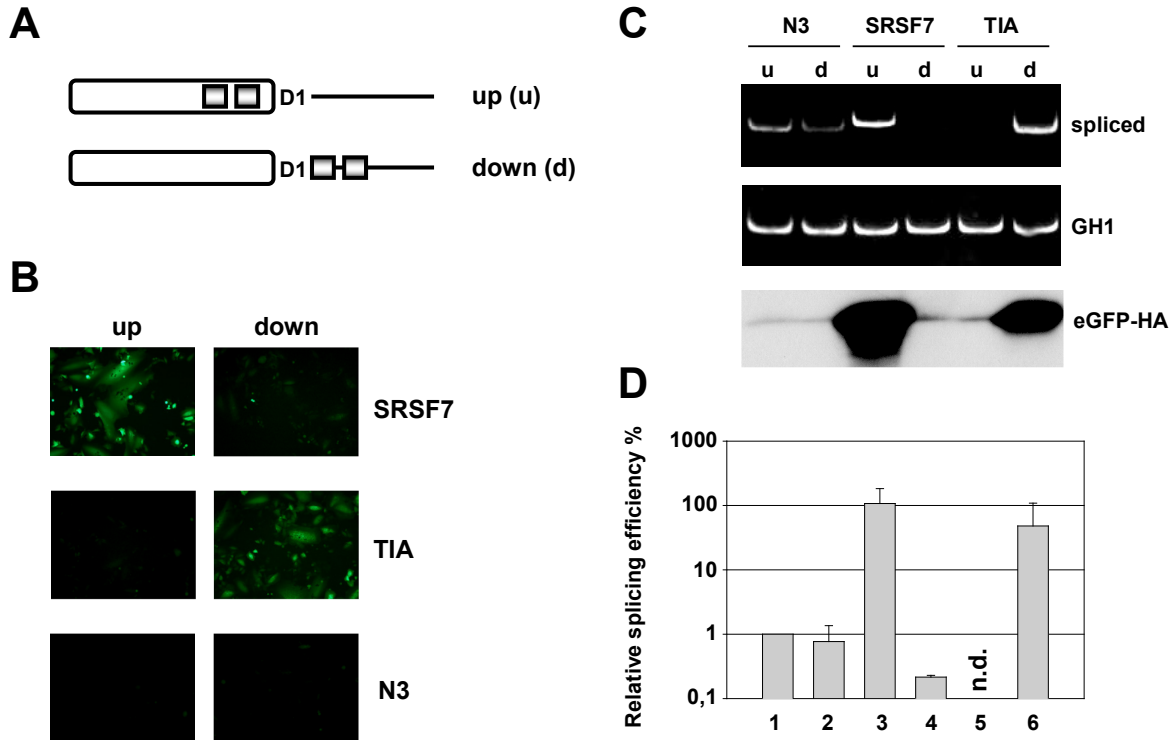


Fig.III-3: The position of SRSF7 and TIA-1 high affinity RNA binding sites determines their outcome for 5'ss use.

(A) Two repeats of binding sites for SRSF7 and TIA-1 were inserted upstream or downstream of 5'ss D1 into SV-env/eGFP-HA. As a control another splicing so-called neutral sequence "N3" [for sequence see Table I; (616)] was placed into the exon or exon and intron, respectively. Downstream binding sites were tested within the context of N3 within the exon. **(B)** 2.5×10^5 HeLa-T4⁺ cells were transiently cotransfected with splicing reporters, SVcrev expressing viral Rev protein, and pCL-dTom to control for equal transfection efficiencies. Env/eGFP-HA fusion protein expression and syncytium formation were analyzed 48h post transfection using fluorescence microscopy. **(C)** 2.5×10^5 HeLa-T4⁺ cells were transiently transfected with each of the constructs and pXGH5 to monitor transfection efficiency (GH1). Total-RNA was isolated 30h after transfection and prior to reverse transcription samples were DNase I digested to remove plasmid contaminations. Resultant cDNAs were used as templates for specific PCR reactions detecting spliced eGFP mRNA (#3210/#3211) and GH1 (#1224/#1225). Amplificates were resolved in 8% polyacrylamide gels and stained with ethidium bromide. Additionally, protein lysates were collected and separated in a 10% denaturing SDS polyacrylamide gel. After electroblotting nitrocellulose membrane was probed with a specific antibody directed against the HA epitope to detect eGFP expression. **(D)** cDNA samples were prepared as in (C) and used in realtime PCR assays with HPLC-purified primers to specifically quantitate levels of spliced eGFP mRNA (#3239/#3240). For normalization primer pair #3242 and #3243 was used to measure the amount of GH1 mRNA for each sample. Fluorescence emission was read by a LightCycler 1.5 (Roche). The values of spliced eGFP mRNA are shown in logarithmic scale, levels for SRSF7 upstream of 5'ss D1 (lane 2) were set as 100%. Bar graph represents mean values \pm standard deviation obtained by three independent transfection experiments. [(u): upstream; (d): downstream; N3: neutral sequence 3; GH1: human Growth Hormone 1; HA: Human Influenza hemagglutinin]

Results

Given this striking observation, that these two arbitrarily chosen SREs could only promote splicing from one side of the 5'ss, the analysis was extended by testing additional high affinity binding sites for SRSF2, hnRNP F/H, Fox1/2, hnRNP I/PTB, hnRNP A1 and TIA-1 (Tab.III-1). Fluorescence microscopy analysis was used to detect the splicing-dependent eGFP expression within the cells (Fig.III-4A). Importantly, all tested target sequences (except the hnRNP A1 binding site), irrespectively whether they were classically categorized as a *cis*-acting splicing enhancer or silencer element, could be classified as a position-dependent *cis*-acting splicing enhancer element (PDE), exclusively activating the 5'ss from either an upstream exonic (uPDE) or a downstream intronic position (dPDE). Whereas the target sequences for the SR proteins SRSF2 and SRSF7 needed to be strictly located within the upstream exonic position to upregulate use of the 5'ss others such as hnRNP F/H, TIA-1, IAS-1 and PTB activated splicing when they were localized only within a downstream intronic position (Fig.III-4A).

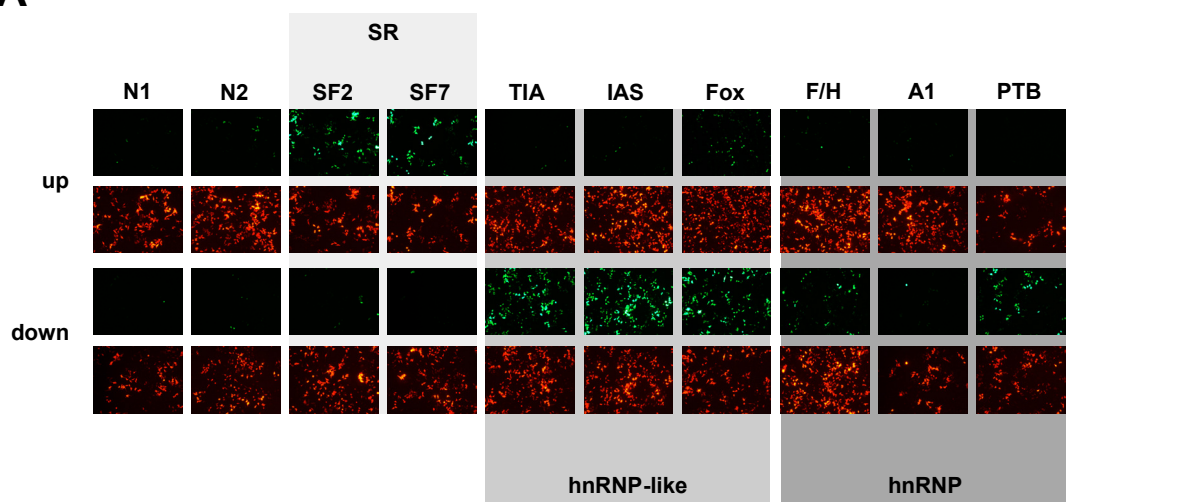
Abbrev.	SRE	Sequence	Ref.
N1	GAR ⁻ ESE ⁻	AGGAAUUAGCGGAUUUAGCGCUUAA	(11)
N2	GAR ⁻	AGGCAUUAGCGGCGACUUCGCUUAA	(11)
N3	neutral	CCAAACAA	(616)
SRSF2	SC35-1	GUUCCAGAUAAAGUCCAGCC	(79)
SRSF7	9G8-66a	AGACAACGAUUGAUCGACUA	(79)
TIA	TIA-1	UCUUUUUAAGUCGUACCUA	(123)
IAS	TIA-1 (IAS-1)	GCUUUCAUUUUUGUCUUUUUUUUAA	(121)
Fox	Fox1/2	UGCAUG	(307)
F/H	hnRNP F/H	AGGGA	(463)
A1	hnRNP A1	UUAGGG	(196)
PTB	PTB	CUCUCU	(83, 84)

Tab. III-1: Neutral control sequences and high affinity binding sites for splicing proteins

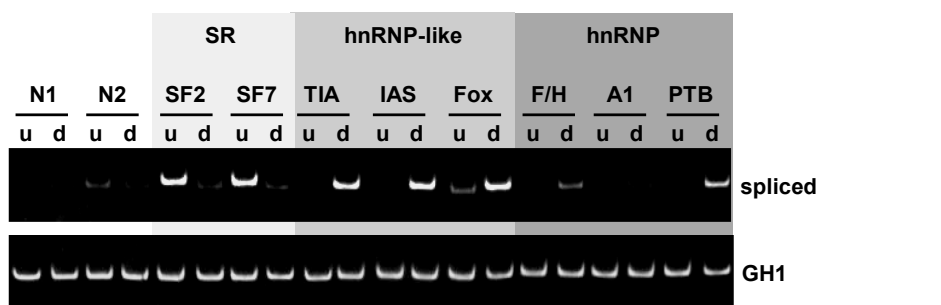
The binding sequences for Fox1/2 protein family were the sole exception, because they also enhanced splicing from the upstream exonic position. However, activation from the downstream intronic position was significantly stronger, recurring the general trend of position-dependent splice site activation. These results were also confirmed by semi-quantitative (Fig.III-4B) and quantitative real-time RT-PCR analyses (Fig.III-4C) of RNAs isolated from transfected HeLa-T4⁺ cells to specifically detect the level of spliced reporter mRNAs. For the normalization of the transfection efficiency constitutively spliced GH1 mRNA expressed from the cotransfected reporter pXGH5 was determined. RT-PCR analyses demonstrated that in all cases splicing was critically dependent on the position of the SREs relative to the 5'ss.

SR proteins facilitated spliced eGFP mRNA expression exclusively when they were placed within the upstream exon (Fig.III-4B, cf. lanes 3 and 4), while hnRNP and hnRNP-like proteins had to be located within the downstream intron to accumulate spliced products (Fig.III-4B, cf. lanes 5, 6, 7, 8 and 10). Quantitative realtime PCR assays showed an increase in expression of spliced eGFP mRNA by activating PDEs up to more than 100-fold relative to the control (Fig.III-4C). To summarize these findings, the position of SREs relative to the regulated 5'ss determine their outcome for splicing. Based on this overall consistent observations, SREs can be functionally categorized according to their activating location in upstream and downstream position dependent enhancer (PDE), emphasizing the general idea of a “*splicing code*” defined by both 5'ss and its sequence environment.

A



B



C



GH1

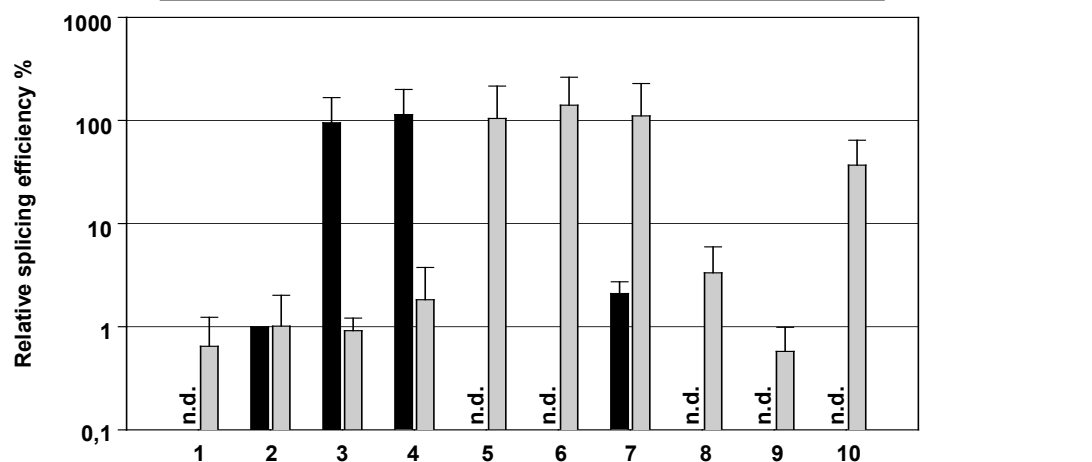


Fig.III-4: SREs activate 5'ss use in a strict position-dependent manner. [continued on next page]

(A) Neutral sequences N1 and N2 and high affinity binding sites for the SR proteins SRSF2 and SRSF7 and the hnRNP (-like) proteins TIA-1, Fox1/2, hnRNP F/H, hnRNP A1 and PTB (sequences listed in Table 1) were placed either into the upstream exon or downstream intron of 5'ss D1 into SV-env/eGFP. All binding sites tested downstream of 5'ss D1 contained neutral sequence N2 in the upstream exon. 2.5×10^5 HeLa-T4⁺ cells were transiently transfected with 1µg of the different splicing reporters and 1µg of pCL-dTom to monitor transfection efficiency for each of the samples. 30h post transfection fluorescence pictures were taken.

III.1.2 The number of binding sites positively correlates with 5'ss use

To examine whether the number of high affinity RNA binding sites correlates with the efficiency of 5'ss use, splicing was determined, supported by one, two or four copies of the target site for SRSF7 or TIA-1 (Fig.III-5A). HeLa-T4⁺ cells were transiently transfected and subsequently analyzed for eGFP expression under the fluorescence microscope (Fig.III-5B) or by semi-quantitative and quantitative RT-PCR (Fig.III-5C-D). Fluorescence images were taken and they revealed for both splicing proteins that already a single high affinity RNA binding site was sufficient to produce considerable amounts of eGFP within the cells (Fig.III-5B). Moreover, the fluorescence intensity could be elevated by multiplying the number of binding sites up to four. This is likely due to an increase of the RNA binding affinity of the splicing proteins or their number within the enhancer complex. RNA samples were then prepared from these cells and subjected to semi-quantitative and quantitative RT-PCR analysis (Fig.III-5C-D). As expected from the observed fluorescence intensities the level of spliced eGFP mRNA increased gradually with the number of enhancer binding sites neighbouring the 5'ss (Fig.III-5C-D). For instance, quantitative realtime PCR analyses could detect an increase in the level of spliced eGFP mRNA up to 7-fold by duplicating the SRSF7 binding site from two to four copies (Fig.III-5D, cf. lanes 3 and 4).

Fig.III-4: continued

(B) Total RNA was isolated from 2.5×10^5 HeLa-T4⁺ cells transiently cotransfected with 1µg of each of the splicing reporters and 1µg of pXGH5 that served as a transfection control (GH1). After DNase I digestion and cDNA synthesis, PCR reactions were carried out with primer pairs #3210/#3211 and #1224/#1225 to specifically detect levels of spliced eGFP mRNA and GH1 mRNA, respectively. Separation of the RT-PCR products was obtained in 8% non-denaturing polyacrylamide gels. **(C)** cDNA samples were prepared as in (B) and used as template DNA for realtime PCR assays with SYBR green and primer pair #3239/#3240 (spliced eGFP mRNA) or #3242/#3243 (GH1 mRNA). Each sample was normalized to the amount of GH1-mRNA. Spliced eGFP mRNA expression was shown relative to N2 upstream of 5'ss D1 (lane 2). Values are shown in logarithmic scale with each division representing a magnitude of 10^2 ($1=10^2$). For quantitation of the fluorescence the emission LightCycler 1.5 (Roche) was used. Data are presented as the average of three independent RT-PCR experiments. [u: upstream; d: downstream; GH1: human Growth Hormone 1; n.d.: not detectable]

These results were in line with previous observations showing that the number of SR proteins positively correlate with the efficiency of the 3'ss use (186). Although splicing at the 5'ss could be upregulated by increased numbers of downstream TIA-1 binding sites, no difference was found between one or two TIA-1 high affinity binding sites in the level of detectable mRNA and merely a 2-fold increase by placing four repeats downstream of the 5'ss (Fig.III-5D, cf. lanes 5-7). However, from these results, it was concluded that the splicing efficiency can be increased by providing multiple enhancer protein binding sites in the proximity of the 5'ss.

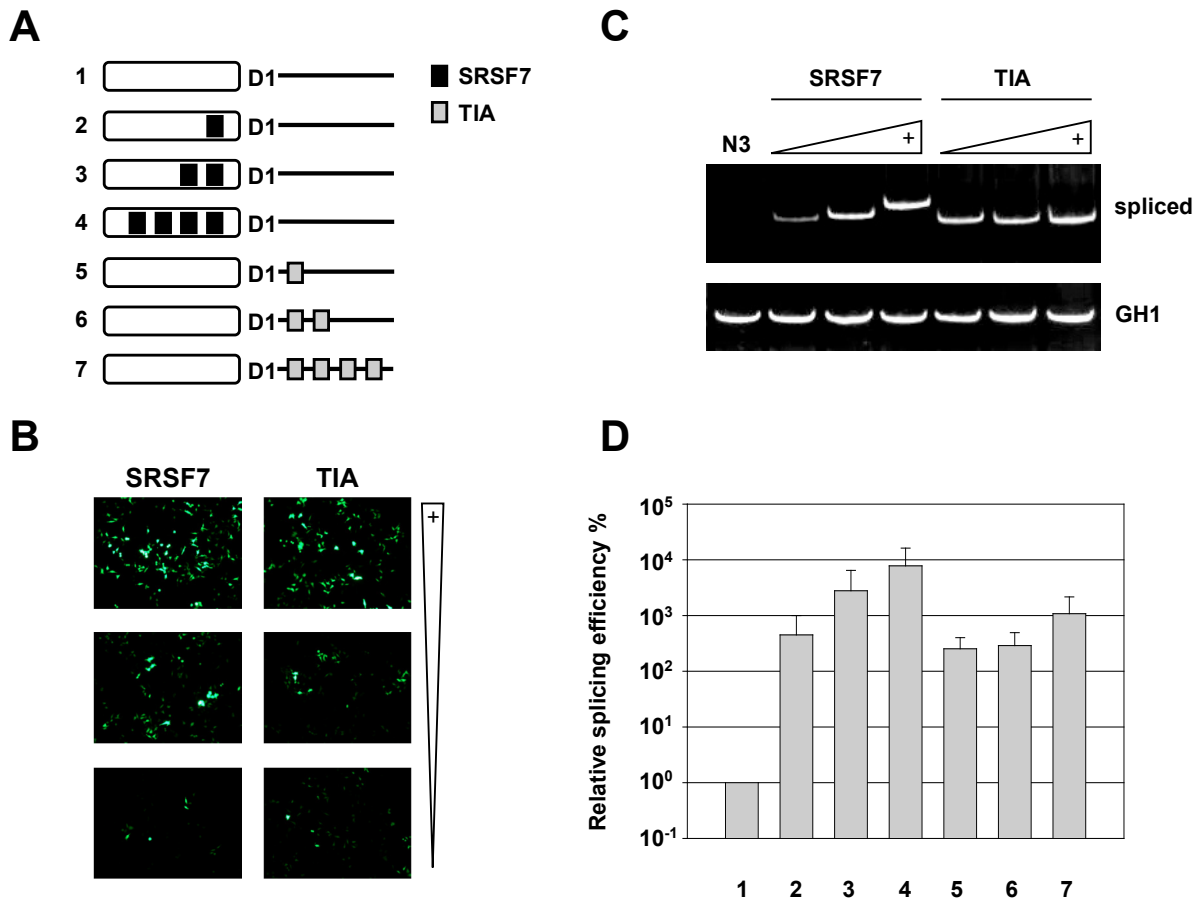


Fig.III-5: Splicing activity can be titrated in through the number of RNA binding site copies.

(A) Schematic drawing of the SV-env/eGFP splicing reporters used for transfection experiments. One, two or four copies of high affinity RNA binding sites for SRSF7 and TIA-1 were placed into their activating position. As a control the neutral sequence N3 was inserted in the upstream exon and downstream intron. Furthermore, N3 was present within the upstream exon of the downstream tested binding sites for TIA-1. **(B)** 2.5×10^5 HeLa-T4⁺ were transiently transfected with each of the construct and pCL-dTom to control transfection. 30h later fluorescence was assayed by microscopy. **(C)** Total-RNA was harvested from transfected HeLa-T4⁺ cells and after revers transcription subjected to RT-PCR analysis. The levels of spliced mRNA were detected with primer pair #3210 and #3211. To ensure equal transfection efficiencies GH1 mRNA was amplified with primers #1225 and #1225. Following amplification, PCR-products were resolved in 8% non-denaturing polyacrylamide gels and stained with ethidium bromide. **(D)** cDNA samples were generated as in (C) and used as template DNA for realtime PCR reactions with SYBR green in LightCycler1.5 (Roche). Spliced eGFP mRNA was detected with primers #3239 and #3240. For normalization of the samples, levels of GH1 were determined (#3242/#3243). The value of two copies of an SRSF7 binding site in the upstream exon were set as 100% (lane 3). Data represents the average of three-independent transfection experiments. [GH1: human Growth Hormone 1; N3: neutral sequence 3]

III.1.3 Multiple enhancing PDEs compensate for weak intrinsic strength of a 5'ss

Because it was demonstrated that the number of enhancer elements positively correlated with use of 5'ss D1, the question was addressed of whether a surrounding rich in activating PDEs can compensate for weak interactions between the 5'ss and the U1 snRNA. For this purpose, site-directed mutations were introduced into D1 to stepwisely weaken its complementarity to the U1 snRNA. The 5'ss strength was calculated for each sequence using both the Maximun Entropy model and the HBond Score algorithm (Fig.III-6A). A panel of four 5'ss sequences of variable strength was tested ranging from strong (D1) to very weak (L3) based on the computational predictions concordantly obtained by both algorithms. To determine whether splicing at weak 5'ss can be rescued by its neighbourhood, the 5'ss sequences were placed between either four or eight activating PDE binding motifs (Fig.III-6B). HeLa-T4⁺ cells were transiently cotransfected with each of the constructs and pXGH5 to monitor the transfection efficiency. RNA was isolated from the cells and analyzed by semi-quantitative and quantitative RT-PCR analyses (Fig.III-6C-D). As expected for those PDEs outside of their activating positions (TIA-1 in an exonic and SRSF7 in an intronic position), independent on their intrinsic strength, no spliced eGFP-mRNA could be detected when 5'ss D1 was flanked by two affinity binding sites at each side (Fig.III-6C-D, lane 1). However, splicing at D1 was efficiently restored by permutation of both PDEs into their activating locations (Fig.III-6C-D, lane 2).

As predicted, weakening D1 through L2 gradually decreased the relative levels of spliced eGFP-mRNA (Fig.III-6C-D, lane 2-4). Finally, the complementarity of 5'ss L3 to the U1 snRNA (HBS 10.8, Maxent 7.59) was too low to be rescued by an overall number of four activating PDE binding sites in the vicinity and no spliced eGFP-mRNA could be detected (Fig.III-6C-D, lane 5). However, even this poor 5'ss L3 with an intrinsic strength rarely represented in the dataset of annotated human 5'ss could be activated by further increasing the number of activating PDE elements (Fig.III-6C-D, cf. lane 5 and 7). These results reiterate that use of the 5'ss is the net result of U1 snRNA complementarity and the strength of PDEs in its vicinity.

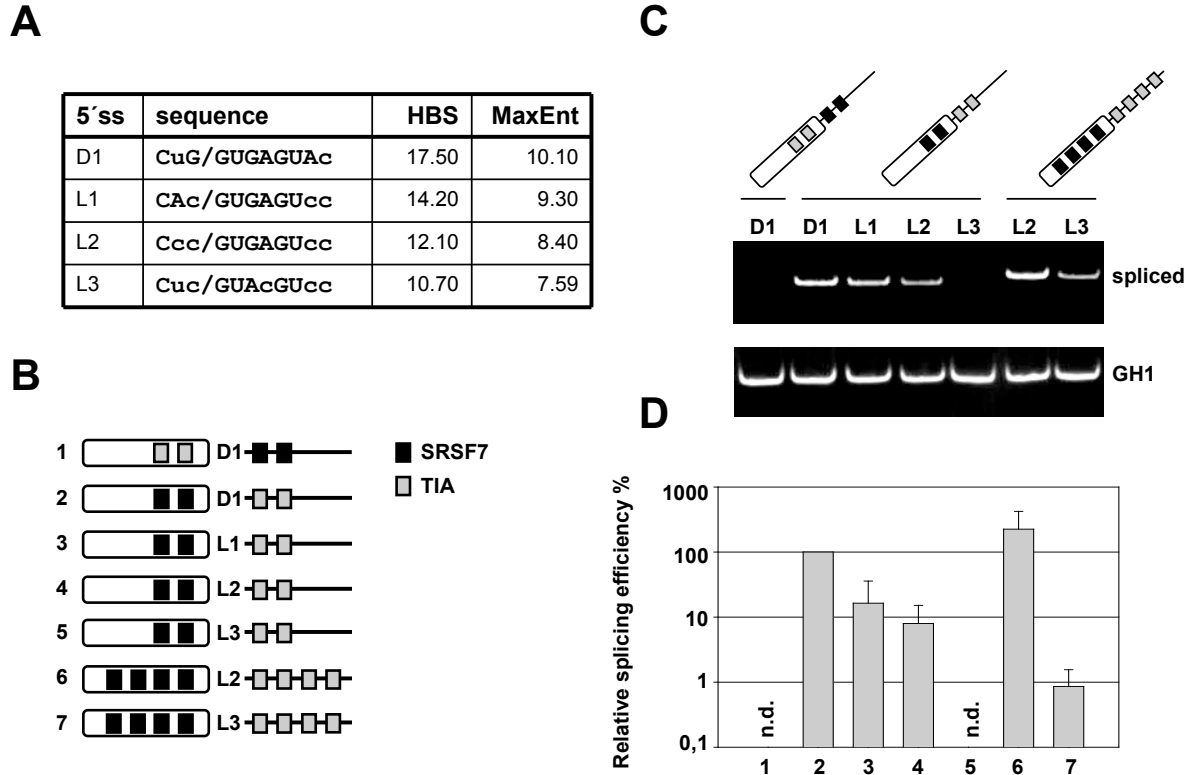


Fig.III-6: Multiple activating PDEs can rescue an intrinsic weak 5'ss.

(A) Mutations were inserted into 5'ss D1 to gradually weaken its complementarity to the U1 snRNA. Sequences and predicted intrinsic strength by splice finder (HBS) and MaxEnt score are listed. (B) Schematic diagram of the SV-env/eGFP constructs. Splice site variations were tested with two or four copies of an activating PDE within each flank. As a control PDEs were tested outside their activating positions in combination with 5'ss D1. (C) 2.5×10^5 HeLa-T4⁺ cells were cotransfected with each of the constructs and pXGH5 to monitor equal transfection efficiency. Afterwards total RNA was isolated from the cells and revers transcribed. Resultant cDNA served as template DNA to specifically amplify spliced eGFP mRNA (#3210/#3211) and GH1 mRNA (#1224/#1225) for control. PCR products were separated in 8% denaturing polyacrylamide gels and visualized by ethidium bromide staining. (D) cDNA were prepared as in (C) and analyzed by realtime PCR. For quantitation of spliced eGFP mRNAs primers #3239 and #3240 were used. Relative expression was normalized to the amount of GH1 mRNA (#3242/#3243) for each sample. Fluorescence was monitored by LightCycler1.5 (Roche). The values of spliced eGFP mRNA are shown in logarithmic scale, levels for two activating PDEs at each side of 5'ss D1 (lane 2) were set as 100%. Bar graph represents mean values \pm standard deviation obtained by three independent RT-PCR experiments. [HBS: H-Bond Score; L1-L3: Low 1-3; GH1: human Growth Hormone 1]

III.1.4 MS2-tethering of splicing regulatory proteins mimics their position-dependence for 5'ss activation

To investigate the position-dependent splicing outcome independently from the varying RNA affinities of the candidate splicing proteins to their natural binding sites, splicing regulatory proteins were selectively directed to target sites upstream or downstream of the 5'ss by use of an MS2-tethering system (Fig.III-7A). The bacteriophage MS2 coat protein dimer specifically recognizes a 19nt sized RNA hairpin, which can be placed into reporter mRNAs (Fig.III-7B). Coexpression of fusion in-frame coat proteins (Fig.III-7C) allows the site-directed tethering of any splicing protein (or a splicing protein-domain) to the reporter mRNA whose splicing fate then can easily be analyzed e.g. by RT-PCR. To functionally mimic the upstream or downstream positioning of the splicing proteins near the 5'ss two copies of the RNA hairpin were placed either into the upstream exon (SV 2MS2 D1-env/eGFP) or the downstream intron (SV GAR⁻ D1 2MS2-env/eGFP) of the 5'ss D1 (Fig.III-7A). Data obtained from transient transfections of HeLa-T4⁺ cells with these reporter constructs and the expression vectors for the MS2 coat fusion proteins followed by semi-quantitative RT-PCR analysis were consistent with what was found in the native PDE experiments. As a negative control, the level of splicing for each MS2 splicing reporter in the absence of a coexpressed fusion protein as well as in the presence of the MS2 coat protein alone was determined (Fig.III-8D, lanes 1 and 2). The controls demonstrated that neither the RNA hairpins itself supported splicing (the higher basal activity of the "SV GAR⁻ D1 2MS2-env/eGFP" reporter can be attributed to the weak enhancing activity of the "neutral 2" (N2) sequence located within the upstream exon) – nor does the MS2 coat protein alone positively or negatively act on use of the 5'ss. Positioning of SR proteins-derived MS2 coat fusion proteins immediately upstream of the 5'ss enhanced splicing, whereas no increased levels of spliced eGFP-mRNA were detected following fusion protein recruitment to the downstream intronic position (Fig.III-8D, lanes 3-6). On the other hand, TIA-1, Fox2 α , hnRNP F and PTB only enhanced splicing when they were tethered immediately downstream of the 5'ss (Fig.III-8D, lanes 7-11). Outside of their activating positions fusion proteins rather repressed the production of spliced eGFP-mRNA. Proper expression of all MS2 fusion proteins was confirmed by western blot analysis (Fig.III-7E).

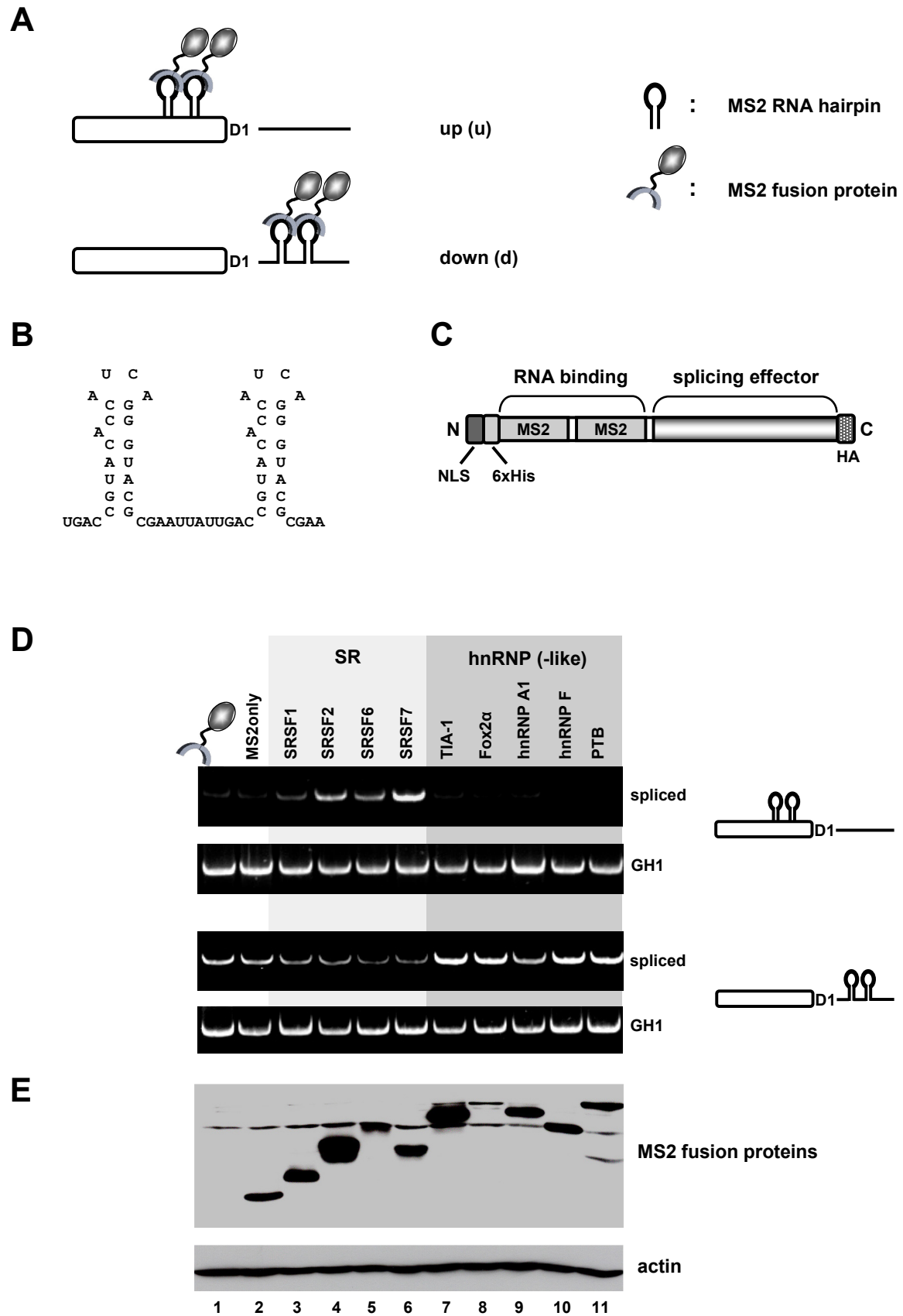


Fig.III-7: Tethering of splicing regulatory proteins reiterates position-dependent 5'ss activation.
[continued on next page]

The results obtained from the tethering experiments revisit the conclusion that the activity of splicing regulatory proteins is highly position-dependent.

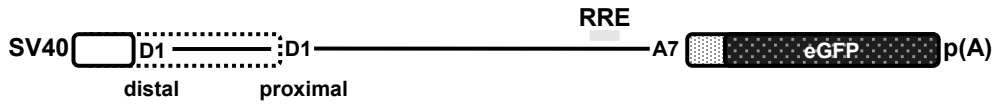
Fig.III-7: continued

(A) Schematic drawing of the SV-env/eGFP derived splicing reporters either containing two copies of a bacteriophage MS2 RNA hairpin within the upstream exon or the downstream intron of 5'ss D1 (latter with N2 within the exon). MS2 fusion proteins can be specifically targeted to the MS2 RNA hairpin to assay their outcome on 5'ss use. **(B)** Secondary structure of the folded MS2 RNA hairpins within the reporter pre-mRNA. **(C)** Schematic diagram of the MS2 fusion proteins coexpressed to recruit splicing regulatory proteins (or domains) to the reporter mRNAs. **(D)** 2.5×10^5 HeLa-T4⁺ cells were transfected with the one of the two splicing reporters together with each of the MS2 fusion protein expressing plasmids and the internal control pXGH5. Total-RNA samples were collected 30h post transfection and cDNA was synthesized following DNase I digestion. PCR reactions were performed using primer pairs #3210/#3211 and #1224/#1225 to specifically detect spliced eGFP- and GH1-expression. Amplificates were electrophoresed using 8% non-denaturing polyacrylamide gels and stained with ethidium bromide. **(E)** Expression of all MS2 fusion proteins was controlled by isolating protein lysates from the transfected cells, 10% SDS-PAGE, electroblotting on a nitrocellulose membrane and specific probing with a primary antibody against the HA epitope. Equal amounts of protein were detected by western blot with a specific antibody directed against β -actin. [(u): upstream; (d): downstream; NLS: Nuclear Localization Signal; HIS: Histidin epitope; HA: Human Influenza hemagglutinin; GH1: human Growth Hormone 1].

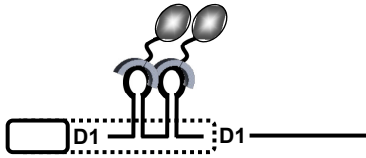
III.1.5 Splicing regulatory proteins modulate the competition of two 5'ss

To determine the effects of splicing regulatory proteins on the selection of two alternative 5'ss, a modified splicing reporter was generated, containing a duplicated 5'ss D1 (Fig.III-8A). The 5'ss were defined as "proximal" and "distal" according to their distance relative to the 3'ss A7. A tandem MS2 RNA hairpin was inserted between the two identical 5'ss to test the influence of varying fusion proteins on 5'ss choice (Fig.III-8B). HeLa-T4⁺ cells were transiently cotransfected with the reporter construct (SV D1 2MS2 D1-env/eGFP) and each one of the fusion protein expression plasmids. Total-RNA was harvested and analyzed by semi-quantitative RT-PCR (Fig.III-8B). Unexpectedly, the proximal 5'ss was chosen above the distal 5'ss when either MS2 alone or no fusion protein was expressed, although both 5'ss were of identical intrinsic strength (Fig.III-8B, lanes 1 and 2). However, in agreement with the results obtained with the splicing reporter harbouring a single test 5'ss, all SR-fusion proteins promoted proximal splicing (Fig.III-8B, lanes 3-6), whereas all other tested fusion proteins preferentially directed spliceosome assembly to the distal 5'ss, indicated by a smaller sized RT-PCR product resulting from excision of a larger intron (Fig.III-8B, lanes 7-11). Competing 5'ss choice was also tested in the context of native PDEs replacing the two MS2 RNA hairpins (Fig.III-8C). Consistent with the results from the tethering experiment, neutral sequences and RNA high affinity binding sites for SRSF7 each sandwiched by two identical 5'ss exclusively activated proximal splicing (Fig.III-8C, lanes 1-4), whereas binding sites for TIA-1 led to predominant detection of a smaller-sized splicing product formed by use of the distal 5'ss (Fig.III-8C, lane 5). Differences in size of the proximal RT-PCR products resulted from size variations among the tested sequences. As a recurrent theme, splicing proteins showed an opposite splicing phenotype: hnRNP and hnRNP-like proteins shifted the spliceosome assembly to the distal 5'ss, while SR proteins enhanced proximal splicing confirming a position-dependent splice site activation by the tested splicing factors.

A



B



C

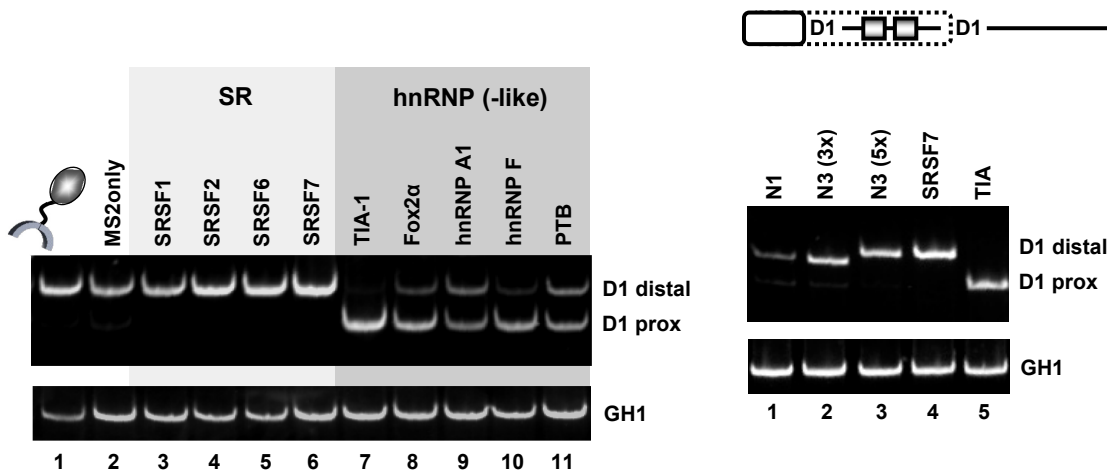


Fig.III-8: Splicing regulatory proteins modulate the selection between two 5' ss.

(A) Schematic diagram of the SV-env/eGFP derived splicing reporter containing a duplicated 5' ss. 5' ss were defined by their position relative to 3' ss A7 (distal and proximal). **(B)** Two copies of the MS2 RNA hairpin were inserted between the two 5' ss. 2.5×10^5 HeLa-T4⁺ cells were transfected with the splicing reporter, each of the MS2 fusion protein expression vectors and pXGH5 to control equal transfection efficiencies. 30h after transfection, RNA was isolated from the cells and subjected to RT-PCR analysis. Primer pair #3210/#3211 was used in a specific PCR reaction to detect alternative 5' ss choice. To monitor transfection efficiencies a separate PCR reaction was performed with primer pair #1224/#1225 (GH1). PCR products were electrophoresed using 8% non-denaturing polyacrylamide gels and visualized by ethidium bromide staining. **(C)** Neutral sequences N1 and N3 (in two different lengths) and two repeats of high affinity binding sites for SRSF7 and TIA-1 were placed between the 5' ss. 5' ss splice site choice was assayed as described in (B). [RRE: Rev Responsive Element; prox: proximal; GH1: human Growth Hormone 1; (3x): three repeats; (5x): five repeats;]

III.1.6 Position-dependent enhancers (PDEs) can act as silencers of splicing

Given the unifying feature of splicing regulatory proteins to enhance splicing in a strict position dependent manner, the combinatorial outcome of multiple PDEs for 5'ss activation was evaluated (Fig.III-9A). Since the observation was made, that PDEs could have a functionally opposite splicing phenotype outside their activating positions, it was tested whether a downstream PDE (dPDE) competes with splicing after repositioning it between an activating upstream PDE and the 5'ss (Fig.III-9A). Therefore, two RNA high affinity binding sites for SRSF7 and D1 were separated by either one or two copies of the target sequence for TIA-1 (Fig.III-9A, sketches 2 and 4). Effects on splicing related to increased spacing between the binding sites for SRSF7 and the 5'ss were controlled by insertion of splicing neutral sequences ("neutral 3", N3) of comparable length (Fig.III-9A, sketches 1 and 3). In addition, positions of the binding sites for SRSF7 and TIA-1 (or the neutral sequence) were permuted within the exon (Fig.III-9A, sketches 5 and 6). The relative levels of spliced eGFP-mRNA were determined by semi-quantitative and quantitative RT-PCR analysis (Fig.III-9B and C). Inserting the TIA-1 binding sites between the SRSF7 PDE and the 5'ss inhibited splicing (Fig.III-9B and C). However, this could not be explained by the greater distance of the SRSF7 uPDE from the 5'ss because substitution of the TIA-1 binding site(s) by a splicing neutral sequence of similar length still allowed efficient splicing (Fig.III-9B and C, lanes 1-4). Changing the order so that the TIA-1 uPDE came first followed by the SRSF7 dPDE within the exon restored splicing (Fig.III-9B and C, lane 6). However, to further exclude a decreased splicing efficiency as the result of a greater distance from the 5'ss, the same experiment was repeated using another control sequence ("neutral 1", N1). Semi-quantitative RT-PCR analysis revealed that while spacing of the SRSF7 uPDE and the 5'ss by the TIA-1 dPDE repressed splicing, the presence of the splicing neutral control sequence was tolerated (see Appendix, Fig.VII-1.2, cf. lanes 7 and 8).

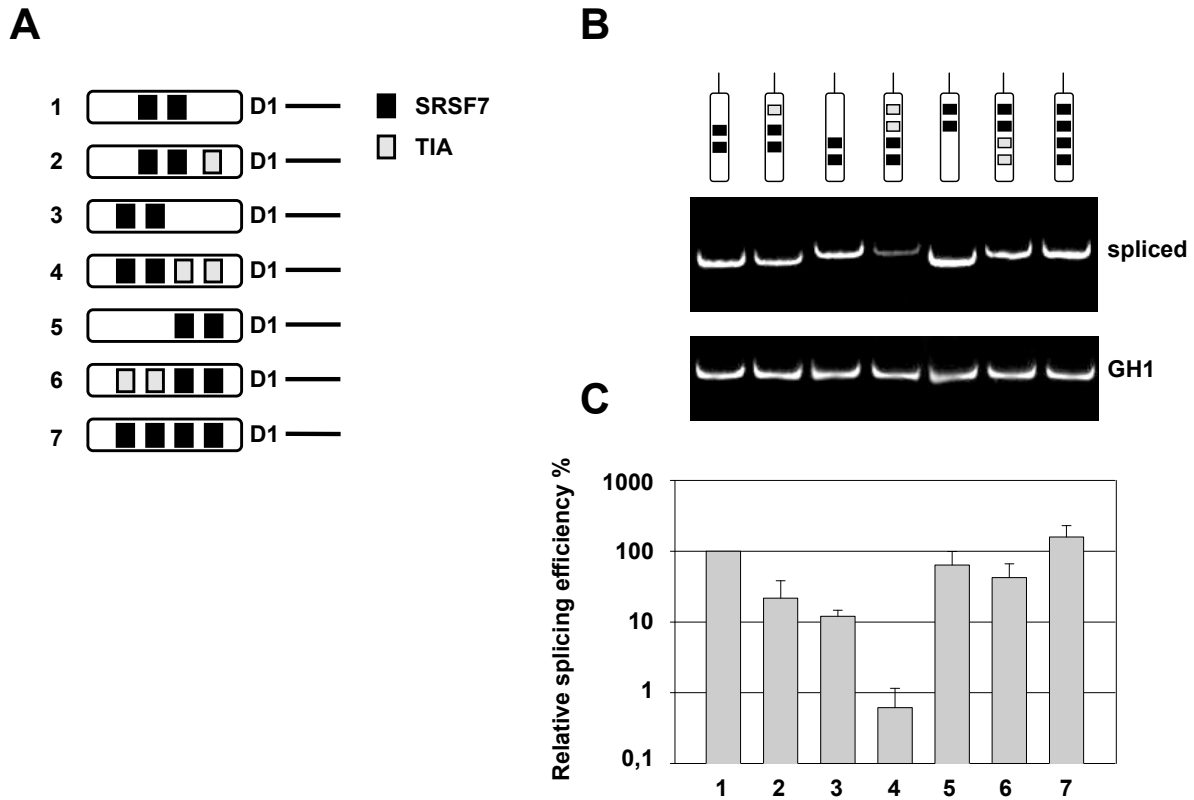


Fig.III-9: Combinatorial outcome of multiple PDEs for 5'ss use reveals that PDEs act as silencers of splicing outside their activating position.

(A) Schematic drawing of SV-env/eGFP splicing reporter containing different SRSF7 and TIA-1 permutations within an exonic location. Neutral sequence N3 was used as a distance control for spacing between the PDE and the 5'ss. **(B)** 2.5×10^5 HeLa-T4⁺ cells were transiently transfected with each of the constructs and pXGH5 to determine transfection efficiency. Total-RNA was harvested from the transfected cells and revers transcribed. The obtained cDNA samples were used in a PCR reaction detecting the levels of spliced eGFP mRNA (#3210/#3211). To monitor equal transfection efficiencies a separate PCR was performed with primer pair #1224/#1225 detecting GH1 mRNA for each sample. PCR-products were resolved by polyacrylamide gelelectrophoresis and stained with ethidium bromide. **(C)** Realtime PCR assays were carried out using cDNA samples prepared as described in (B). Spliced eGFP mRNA levels (#3239/#3240) were normalized to the transfection efficiencies (GH1, #3242/#3243) and depicted as average value obtained from three independent RT-PCR experiments relative to sample 1 as baseline. [GH1: human Growth Hormone 1]

To extend these observations, it was determined whether splicing activated from an exonic PDE (SRSF7) or an intronic PDE (TIA-1) could be repressed when a second PDE of the same type is placed onto the opposite side of the neighbouring 5'ss. Transiently transfected HeLa-T4⁺ cells were flow-cytometrically analyzed to determine the splicing-dependent eGFP expression of each of the constructs (Fig.III-10). Constitutively coexpressed dTomato was used to monitor the transfection efficiencies. The splicing activity measured for each construct was shown as bold blue line relative to splicing neutral control – “neutral 2” (N2) (solid grey area) - exhibiting only residual eGFP expression (Fig.III-10). As expected, inside of their activating positions the PDEs efficiently upregulated splicing, whereas eGFP was poorly detected when they were relocated to the opposite side of the 5'ss (Fig.III-10, cf. histograms 1, 2, 5 and 6). Compared to the activity of a single uPDE or dPDE at an enhancing position flanking the 5'ss at both sides with the same PDE led to drastically lowered levels of eGFP (Fig.III-10, cf. histograms 3 and 7), indicating repressive effects that can even dominate activation (Fig.III-10, cf. histogram 7).

Results

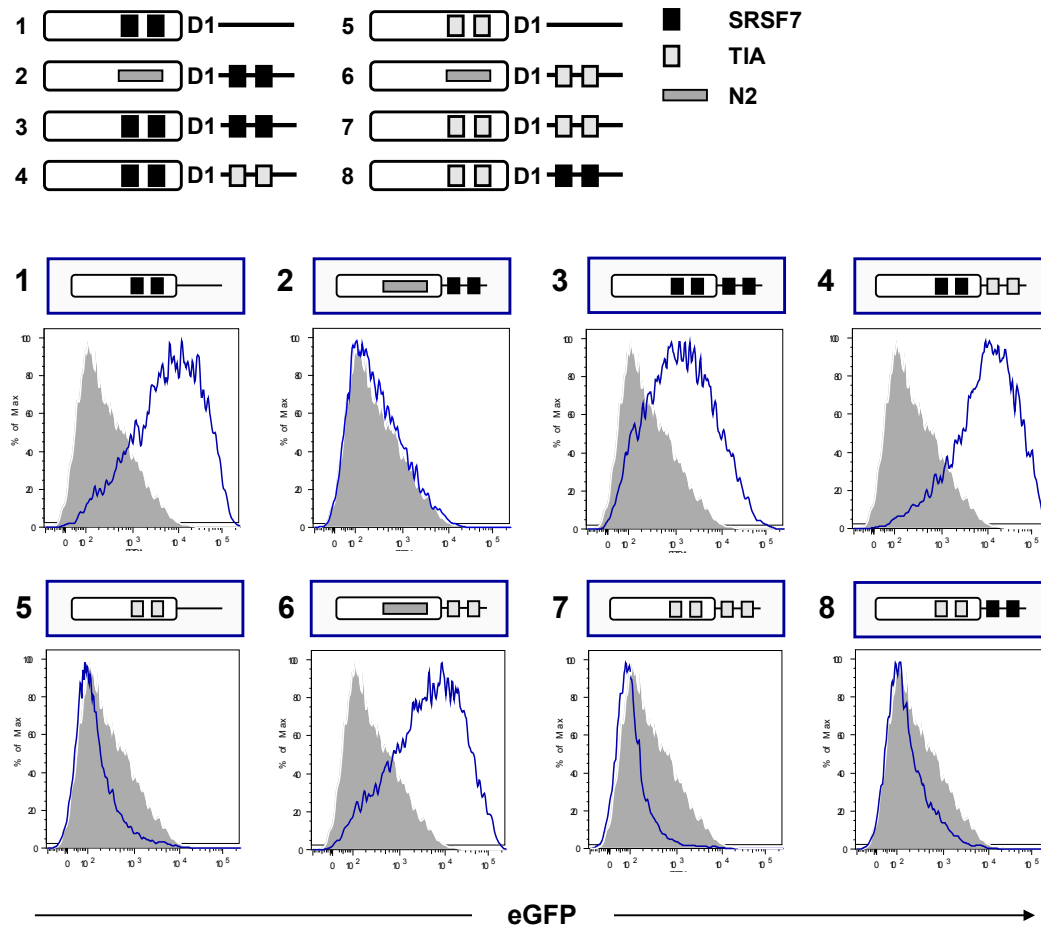


Fig.III-10: Splicing regulatory proteins compete and cooperate in a position-dependent manner.

Schematic drawing of SV-env/eGFP derived splicing reporters containing either a single PDE at an enhancing position or PDEs at both sides of 5'ss D1 (upper panel). 2.5×10^5 HeLa-T4⁺ cells were transiently transfected with each of the constructs or SV-env/eGFP containing neutral sequence N1 or N2 within an upstream exonic location as well as pCL-dTom constitutively expressing fluorescent dTomato as internal control for transfection efficiency. 30h post transfection cells were harvested and analyzed for eGFP expression by flow cytometry. For separation of living and dead cells, samples were stained with DAPI. The mean fluorescence intensity was calculated by FlowJo software. [N2: neutral sequence 2; DAPI: 4',6-Diamidin-2-phenylindol].

Simultaneous tethering of splicing regulatory proteins to the upstream exonic and downstream intronic position led to a splicing repressive phenotype (Fig.III-12). SR protein derived PP7 fusion proteins decreased the levels of spliced eGFP-mRNA, whereas hnRNP-like proteins such as TIA-1 or Fox2 α reduced 5'ss use following recruitment to the upstream exonic position (Fig.III-12). These results further supported the notion that splicing enhancer proteins act as repressors of splicing outside of their activating positions.

Results

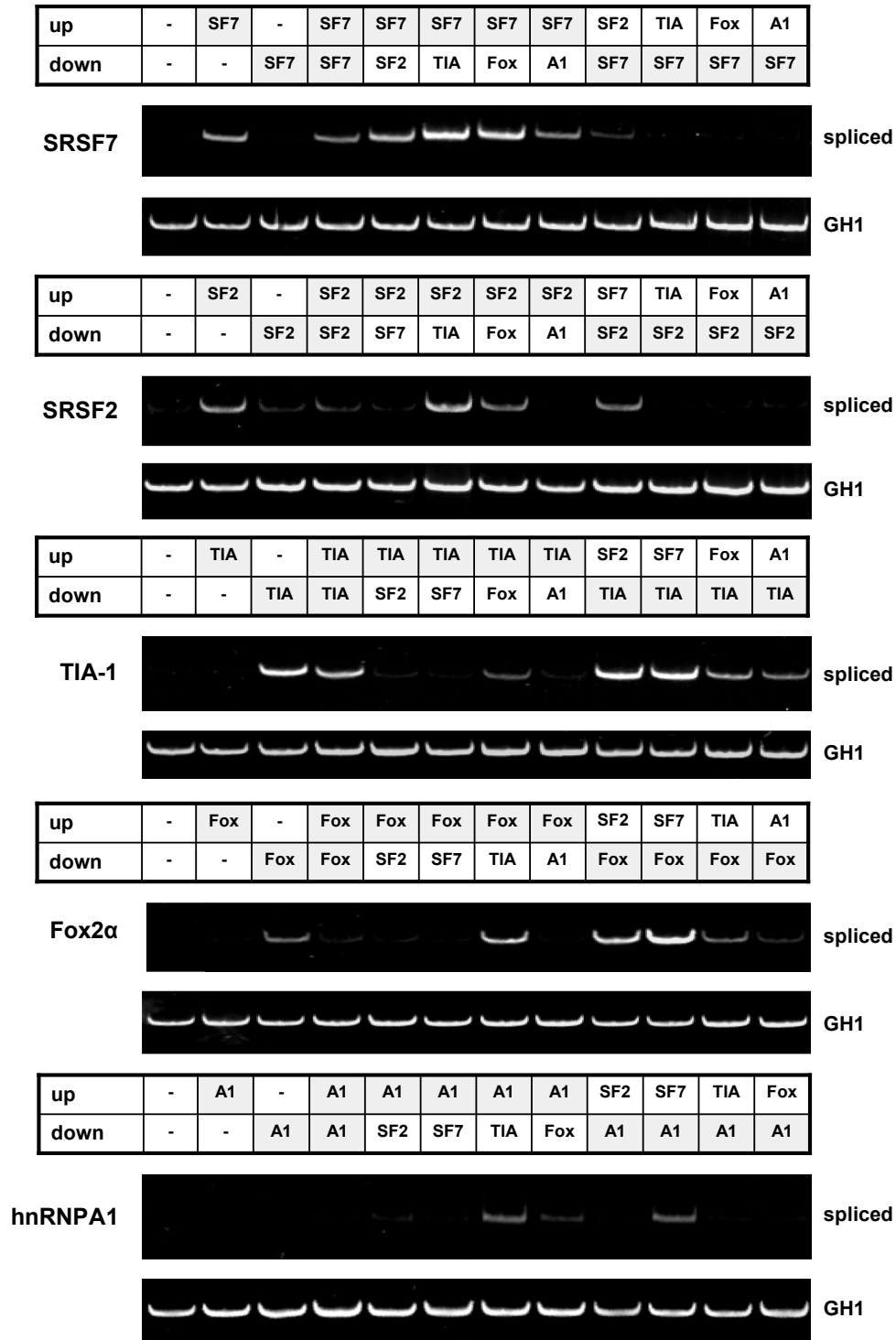


Fig.III-12: Simultaneous tethering of splicing regulatory proteins upstream and downstream of the 5'ss to analyse their combined outcome for 5'ss use.

Total-RNA was purified from HeLa-T4⁺ cells cotransfected with the splicing reporter, the bacteriophage fusion protein expression vectors and pXGH5 for transfection control and subjected to RT-PCR analysis with primers #3210 and #3211 (spliced eGFP-mRNA) or #1224 and 1225 (GH1 mRNA). PCR products were separated by 8% non-denaturing polyacrylamide gel-electrophoresis and stained with ethidium bromide. Relative expression of spliced eGFP-mRNA is depicted below in panel in %. [up: upstream; down: downstream; GH1: human Growth Hormone 1].

III.1.7 The progression into a functional spliceosome is determined by the side from which U1 snRNP is recruited to the 5'ss

Because of the recurrent findings that the splicing outcome of SREs was highly position-dependent and an activating PDE could be converted into a silencing one by permutation of its location relative to the 5'ss, possible mechanisms for position-dependent repression were addressed. Studies performed in collaboration with the Hertel lab (Dr. Klemens J. Hertel group, UCI) provided evidence that SRSF7 localized to the repressive intronic position did not interfere with spliceosomal formation at an early E complex stage, indicating that repression does not underlie a failure to efficiently recruit the U1 snRNP to the 5'ss (145). However, progression into later stages of spliceosome assembly and formation of higher order complexes (A, B and C complex) was arrested, suggesting an accumulation of “frozen” dead-end assembly intermediates containing the U1 snRNP. To determine whether the U1 snRNP is efficiently recruited by splicing regulatory proteins to the pre-mRNA regardless of their position, the presence of U1 snRNP specific proteins upon immobilization of various *in vitro* transcribed RNAs containing a 5'ss and PDEs in upstream exonic or downstream intronic locations was analyzed. RNA affinity purification experiments were performed with short RNA substrates, including either the SRSF7 uPDE or the TIA-1 dPDE each placed immediately upstream or downstream relative to the 5'ss D1 (Fig.III-13A). As a non-binding control the splicing neutral sequences “neutral 3” (N3) was used, which poorly promoted splicing in the functional assays and therefore was expected to recruit if at all only minor levels of U1 snRNP-associated proteins. *In vitro* transcribed RNAs were immobilized on agarose beads and incubated in HeLa nuclear extract under conditions allowing E complex formation. Bound fractions were eluted from the RNA substrates and analyzed for the presence of the U1 snRNP by immunoblotting. As expected, the control pulled down only small amounts of U1 snRNP specific proteins (Fig.III-13B, lanes 3). However, enhanced levels for both U1-specific proteins were detectable with RNAs containing the SRSF7 PDE regardless of its upstream or downstream location relative to the 5'ss (Fig.III-13B, lanes 4 and 5).

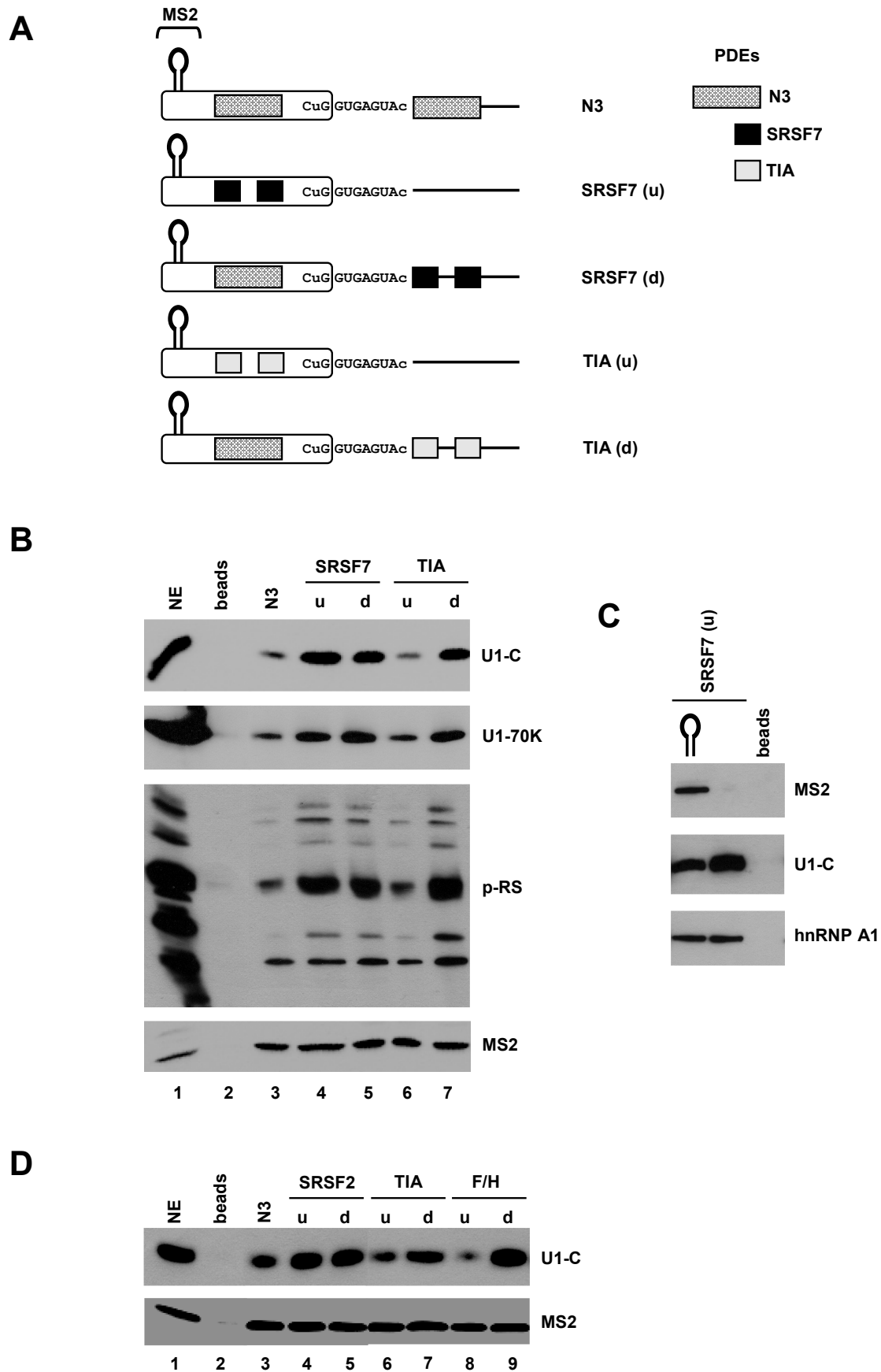


Fig.III-13: Position determines the recruitment of a „dead-end“ U1 snRNP to the 5'ss.
[continued on next page]

This was consistent with the spliceosomal complex formation experiments and indicated that the repressive splicing phenotype by SRSF7 was not the result of a failure to recruit the U1 snRNP to the 5'ss. For TIA-1 the detected levels of U1-70K and U1-C were more affected by the side on which the PDE was placed. While an intronic location of the TIA-1 PDE, as expected, enhanced recruitment of U1 snRNP, an RNA substrate containing the PDE within the upstream exonic location precipitated minor levels of U1-70K and U1-C albeit to the same extent than the control (Fig.III-13B, cf. lanes 3, 6 and 7). This argued against an interference of U1 snRNP recruitment when TIA-1 was located in its repressive position.

Fig.III-13: continued

(A) Schematic diagram of the T7-polymerase generated RNA substrates used in the RNA affinity chromatography (RAC) assays. High affinity RNA binding sites for SRSF7 and TIA-1 were each placed immediately upstream or downstream of 5'ss D1. The 5'-end of all transcripts contained a single MS2 RNA hairpin to control for equal RNA precipitation efficiencies. **(B)** PCR reactions were performed to specifically amplify the regions containing the respective PDE permutations. Resultant amplicates were used as templates for *in vitro* transcription with T7 polymerase. *In vitro* transcribed RNAs were immobilized to agarose beads and incubated in 20% HeLa nuclear extract saturated with recombinant MS2 coat protein at 30°C for 20min. Bound proteins were eluted with SDS sample buffer and heated to 95°C prior to resolution by 12% SDS-PAGE. Separated proteins were electroblotted to a nitrocellulose membrane and probed with specific antibodies directed against U1-C, U1-70K, phosphorylated RS domains and MS2. Appropriate secondary antibodies conjugated to horseradish peroxidase (HRP) were used for detection of signals. **(C)** To ensure specific binding of MS2 coat protein, transcripts were produced either containing or lacking a functional RNA hairpin within the 5'-end. Samples were prepared as in (B) and probed with primary antibodies specifically detecting MS2 coat protein, U1-C or hnRNP A1. **(D)** Further PDEs were tested (SRSF2 and hnRNP F/H) either upstream or downstream of 5'ss D1. Samples were treated as described in (B) and U1 snRNP recruitment was determined by detection of bound U1-C protein levels. [PDE: position dependent enhancer; N1: neutral sequence 1; N3: neutral sequence 3; u: upstream; d: downstream; NE: Nuclear Extract]

An antibody directed against phosphorylated RS domains (p-RS) showed that high levels of U1-70K and U1-C were accompanied by an increased presence of SR- and SR-like proteins on the RNAs, which are known to associate with spliceosomal complexes. To monitor the RNA precipitation efficiency for each reaction, the 5'-end of the RNAs was equipped with a single MS2 hairpin and recombinant MS2 protein was added to the HeLa nuclear extracts. A control experiment confirmed specific binding of the MS2 coat protein to the *in vitro* transcribed RNAs as no MS2 coat protein could be detected from transcripts lacking the MS2 hairpin (Fig.III-13C). To gain further insights into the mechanisms by which SR proteins and hnRNP-like proteins act to repress splicing, the experiment was repeated using other PDEs. This time *in vitro* transcribed RNAs were synthesized containing high RNA affinity sites for SRSF2 and hnRNPF/H immediately upstream or downstream of the 5'ss and E-like complex formation was screened by western blotting with the antibody against the U1 snRNP-specific protein U1-C (Fig.III-13D). Again, SR proteins recruited U1 snRNP to the RNA regardless of the position of the PDE (Fig.III-13D, lanes 4 and 5), whereas detection of U1-C in the case of hnRNP and hnRNP-like proteins was significantly influenced by the location relative to the 5'ss (Fig.III-13D, lanes 6-9). Further support for efficient recruitment of the U1 snRNP by SR proteins independently of the PDEs position was obtained from a titration-experiment, using increasing concentrations of HeLa nuclear extracts in presence of a constant quantity of *in vitro* transcribed RNA (Fig.III-14). Here, it was shown that SRSF7 could efficiently precipitate the U1 snRNP-specific proteins U1-70K, U1-C and U1-A from both sides relative to the 5'ss.

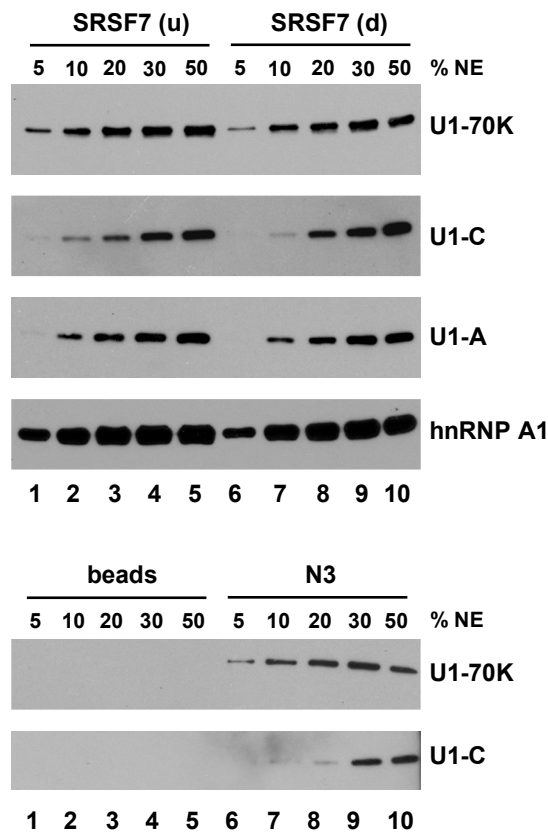


Fig.III-14: The position determines recruitment of a „dead-end“ U1 snRNP to the 5'ss.

The SRSF7 PDE was tested at each side of 5'ss D1 for U1 snRNP specific protein recruitment using increasing amounts of HeLa nuclear extract. In addition, recruitment by neutral sequence N3 was examined. *In vitro* transcribed RNAs were immobilized to agarose beads and incubated in HeLa nuclear extracts concentrations ranging from 5% up to 50%. After extensive washing to remove unbound proteins, samples were subjected to 12% SDS PAGE and western blot analysis with antibodies specific for U1-C, U1-70K, U1-A and hnRNP A1. [PDE: position dependent enhancer; N3: neutral sequence 3; (u): upstream; (d): downstream; NE: Nuclear Extract]

To elucidate the particular contribution of U1 snRNA interactions with the pre-mRNA to the recruitment of the “dead-end” U1 snRNP complexes, the U1 snRNA binding was disabled by oligonucleotide directed RNase H digestion of the U1 snRNA 5'-end. In principle, levels of U1-C and U1-70K were dramatically reduced in the presence of 5'-end depleted U1 snRNAs irrespective of the PDEs location, demonstrating the importance of RNA duplex formation for U1 snRNP recruitment (Fig.III-15). However, the relatively low level binding of U1-70K by the “TIA-1 (u)” RNA was maintained upon RNase H digestion, rather indicating a weak sequence-independent binding of the U1 snRNP to the RNA (Fig.III-15, 11 and 12). In addition, the “SRSF7 (u)” sequence was less dependent on a functional 5'-end of the U1 snRNA than its counterpart containing the SRSF7 binding sites in the downstream intronic position (Fig.III-15, lanes 7-10), indicating a superior importance of protein-protein interactions for the U1 snRNP recruitment. This observation was further supported by unaltered levels of phosphorylated RS domains in the presence of the truncated U1 snRNA. These results were consistent with a model in which the RNA duplex formation between the U1 snRNA and the 5'ss is critical for U1 snRNP recruitment to

Results

the pre-mRNA, but that position of the PDEs determines the splicing competence of the formed E complexes.

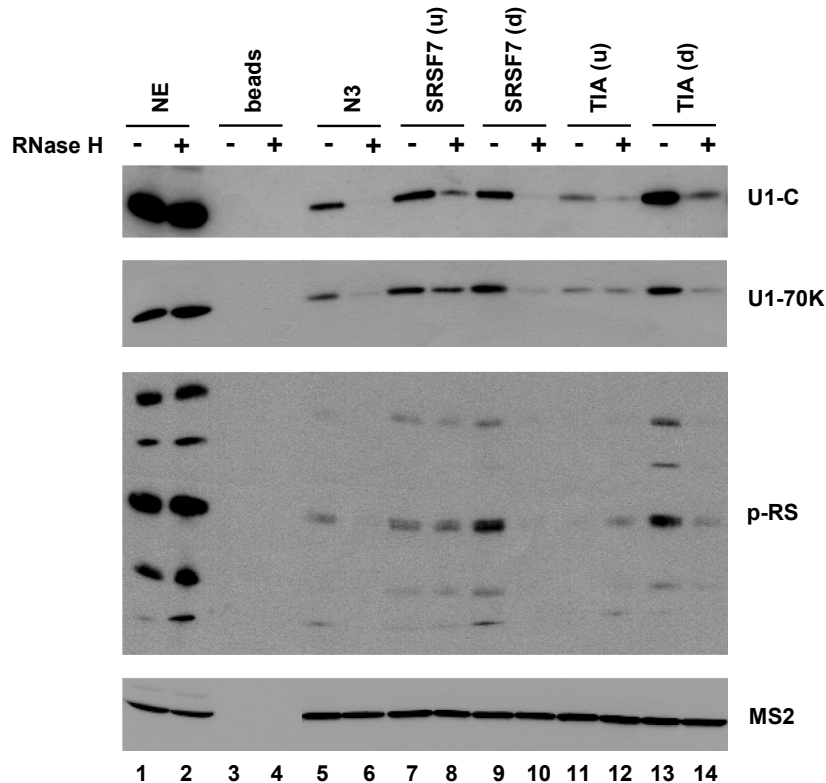


Fig.III-15: Recruitment of U1 snRNP to the 5'ss relies on a functional 5'-end of the U1 snRNA.

RNAs containing the 5'ss and PDEs at various locations were tested for their protein-protein interaction dependent U1 snRNP recruitment. T7 polymerase generated transcripts were coupled to agarose beads. U1 snRNA binding was incapacitated by oligonucleotide (#19967936) directed RNase H digestion of the U1 snRNA 5'-end. Immobilized RNAs were incubated for 20 min at 30°C in untreated and U1 snRNA 5'-end depleted HeLa nuclear extracts. After several washing steps, bound proteins were eluted and separated by 12% SDS-PAGE. Subsequently, proteins were transferred to a nitrocellulose membrane and probed with antibodies specifically detecting U1-C, U1-70K, phosphorylated RS domains and MS2. [PDE: position dependent enhancer; u: upstream; d: downstream, NE: Nuclear Extract]

To determine the integrated outcome of multiple PDEs on U1 snRNP recruitment, RNAs were *in vitro* transcribed containing permutations of the high affinity binding sites for SRSF7 and TIA-1. For both PDEs present in their repressing positions U1 snRNP specific proteins were hardly detected (Fig.III-16A, lane 1), suggesting that a negative impact of the upstream TIA-1 binding sites on U1 snRNP recruitment dominates E-like complex formation mediated by the SRSF7 PDE on the opposite

side relative to the 5'ss. However, switching the PDEs to their activating positions led to efficient precipitation of U1-70K and U1-C (Fig.III-16A, lane 2), which could be further increased by a higher number of high RNA affinity binding sites flanking the 5'ss (Fig.III-16A, lane 3). That the complementarity of the U1 snRNA to the 5'ss sequence also contributes to the U1 snRNP recruitment was once more suggested by severely decreased levels of U1-C, U1-70K and RS-domain containing proteins by testing 5'ss of variable intrinsic strength in the RNA purification experimental set-up (Fig.III-16B).

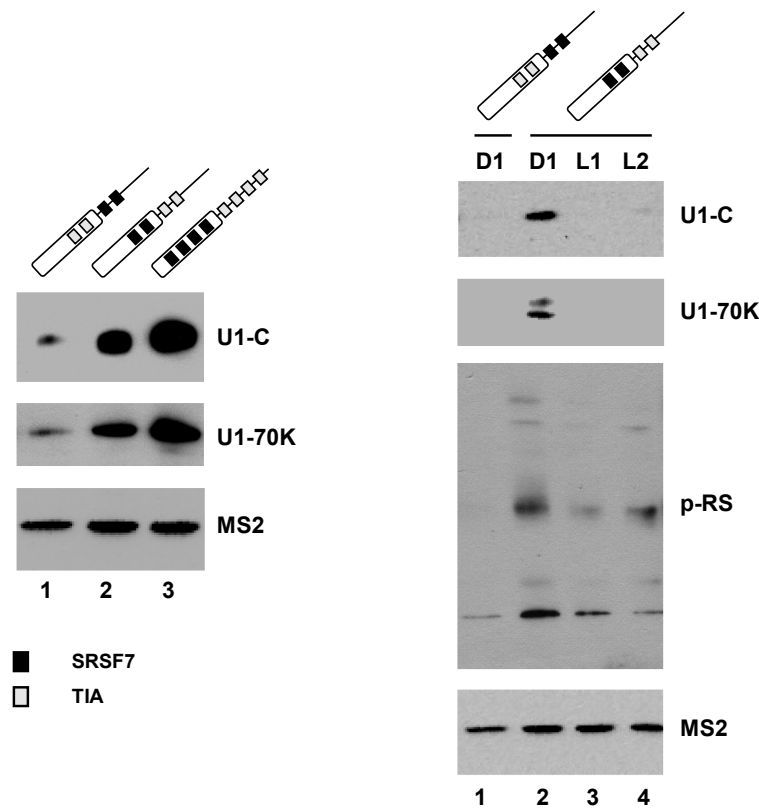


Fig.III-16: Recruitment of U1 snRNP to the 5'ss relies on a functional 5'-end of the U1 snRNA.

(A) *In vitro* transcribed RNAs containing multiple mis-located PDEs or PDEs inside their activating positions were used for RNA affinity chromatography (RAC) to examine the levels of U1 snRNP binding. Transcript were immobilized at agarose beads and incubated in HeLa nuclear extracts for 20 min at 30°C. Samples were washed five times and bound proteins were eluted by SDS sample buffer and heating to 95°C for 10min and loaded on 12% denaturing SDS polyacrylamide gels. After separation, proteins were blotted on nitrocellulose membranes and probed with primary antibodies directed against U1-C, U1-70K and MS2. For detection appropriate secondary antibodies conjugated to horseradish peroxidase (HRP) and ECLTM reagent were used. **(B)** U1 snRNP recruitment by RNAs containing 5'ss variations (L1-L2) embedded within multiple activating PDEs was measured by RAC as described in (A). [L1-2: Low 1-2; N3: neutral sequence 3; PDE: position dependent enhancer]

Given these results, it is still not fully clear whether repression of SR and hnRNP or hnRNP-like proteins relies on a common mechanism. At least it was found that while SR proteins can enhance recruitment of the U1 snRNP to the 5'ss from repressing positions, exonic hnRNP and hnRNP-like proteins only modestly alter the levels of precipitated U1 snRNP-specific proteins. It awaits further investigations to elucidate whether hnRNP and hnRNP-like proteins in their repressing positions perturb U1 snRNP binding to the 5'ss or arrest progression into a functional spliceosome.

Ultimately, the results obtained here are consistent with the model of an integrated splice code in which the use of a 5'ss is determined by its intrinsic strength, but predominantly regulated by PDEs in the immediate vicinity. PDEs do not only up- or downregulate the assembly of early splicing complexes, but also determine the quality of the recruited U1 snRNP. Here, it could be demonstrated that although SR proteins placed to the downstream intron repress splicing, they do not significantly alter the recruitment of the U1 snRNP to the pre-mRNA. Furthermore, these data suggest that hnRNP and hnRNP-like proteins may use two modes of repression, defined by an early inhibition of U1 snRNP binding to the 5'ss and formation of "dead-end" E-complexes.

III.1.8 SRSF7 PDE-depleted nuclear extracts inhibits proximal 5'ss use

To shed light into whether the position-dependent splicing outcome observed, might be due to distinct U1 snRNP complexes recruited to 5'ss D1, HeLa nuclear extracts incubated with either agarose beads alone (mock-depleted) or an immobilized RNA containing the SRSF7 PDE upstream of 5'ss D1 were used in *in vitro* splicing assays. For complementation alternative 5'ss selection of a radiolabeled β -globin RNA substrate was analyzed [kindly provided by Dr. Klemens Hertel, (213)] (Fig.III-17). Kinetic analyses revealed that mock-depleted HeLa nuclear extract enhanced proximal 5'ss splicing after onset of time course and showed that splice site selection gradually shifted from distal to proximal (Fig.III-17, lanes 1-4). However, the shift towards proximal splicing was kinetically delayed for the SRSF7 uPDE-depleted nuclear extract (Fig.III-17, lanes 5-8). Since distal 5'ss use was maintained this could poorly be explained by a reduction in total U1 snRNP content, but rather indicated

Results

that a factor exclusively critical for proximal 5'ss use was depleted from the extract. For instance, the β -globin RNA template may contain a yet unidentified SRSF7 dependent enhancer, which is inefficiently bound in the SRSF7 PDE depleted-extract. Another interesting hypothesis is that distal and proximal 5'ss splicing might be driven by functionally different U1 snRNP complexes. The fact that distal splicing is unaffected may then indicate that the SRSF7 uPDE RNA predominantly depletes "proximal" U1 snRNP complexes from the nuclear extract.

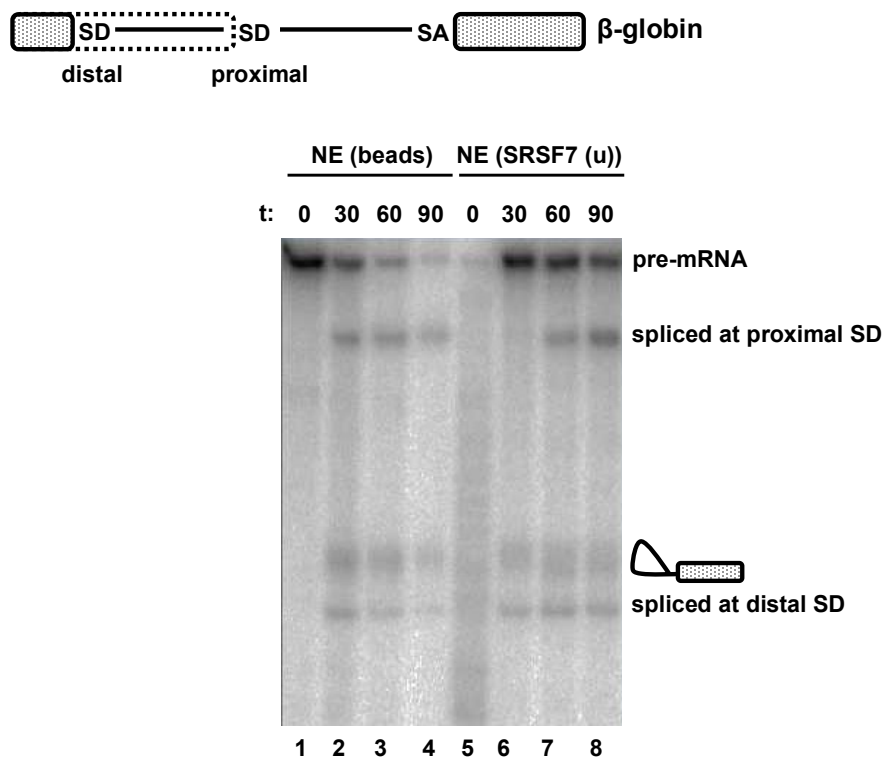


Fig.III-17: HeLa nuclear extract depleted with an RNA containing an activating uPDE kinetically delay *in vitro* splicing at the proximal 5'ss

HeLa nuclear extracts pre-incubated with uncoupled agarose beads or agarose beads conjugated to an *in vitro* transcribed RNA containing an activating SRSF7 PDE upstream of 5'ss D1 were added to a ^{32}P -labeled β -globin derived pre-mRNA harbouring two competing 5'ss. *In vitro* splicing reactions were carried out for defined time points and afterwards RNAs were resolved by PAGE. Radioactivity was read by PhosphorImager (Bio-Rad) and quantified using Quantity One Software. Values were depicted relative to overall amounts of spliced mRNA. [NE: Nuclear Extract; (u): upstream]

III.1.9 Recruitment of U1 snRNP by hnRNP and hnRNP-like proteins might not exclusively be mediated through interactions with U1-C

To determine whether interactions with U1-C represent a unifying pathway by which hnRNP and hnRNP-like proteins recruit the U1 snRNP to pre-mRNA, immunoprecipitation assays were performed. Several splicing protein-derived MS2 fusion proteins were overexpressed within HEK 293T cells, precipitated and analyzed by western blot with an antibody specific for U1-C (Fig.III-18). The well-studied interaction between TIA-1 and U1-C could be confirmed (Fig.III-18, lane 5) (157). However, hnRNP F failed to coprecipitate with U1-C, indicating the protein might use a different mechanism to recruit U1 snRNP to the pre-mRNA (Fig.III-18, lane 6). Unexpectedly, the RS domain of the SR protein SRSF1 bound to U1-C (Fig.III-18, lane 3).

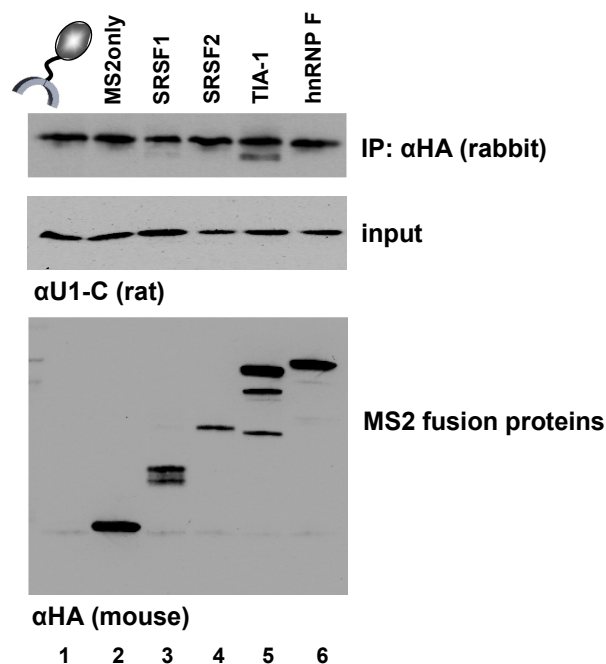


Fig.III-18: U1 snRNP recruitment by hnRNP and hnRNP-like proteins might not be exclusively mediated by interaction with U1-C.

1 x 10⁷ HEK 293T cells were transiently transfected with 18μg of MS2 fusion protein expression plasmids. 48h post transfection cells were harvested and resuspended in lysis buffer. Ten percent of the input material were kept, the remaining samples were precipitated using 2μg of an antibody directed against the HA epitope and protein-G-sepharose (PGS). Subsequently, samples were washed six times, eluted and separated by denaturing 12% SDS-PAGE. Proteins were transferred to nitrocellulose membrane and probed with antibodies specific for U1-C and the HA tag. [IP: immunoprecipitation]

SR proteins have been widely described to interact with U1-70K to enhance binding of U1 snRNP to the 5'ss (90, 164, 276, 589). However, the fact that it was recently described that the hnRNP protein PTB interact with the stem loop IV of the U1 snRNA to stall progression into functional spliceosome assembly (478) and the failure of hnRNP F to coprecipitate U1-C, pointed to an alternative mechanism underlying recruitment of U1 snRNP to the pre-mRNA.

III.2 “S(p)lice up the defence”: How SREs within the HIV-1 exon 3 act in concert to regulate expression of viral infectivity factor Vpr

Splicing at A2 and D3 defines the non-coding leader exon 3 whose inclusion leads to isoforms within both the spliced and spliced, but still intron-containing viral mRNA species. The function of exon 3 is still poorly understood, although it is strongly conserved among the different HIV-1 groups (M, major; N, new; and O; outlier). Since exon 3 alters the 5'-UTR (5'-untranslated region) of viral mRNA species, it was suggested that exon 3 might be implicated in post-splicing mRNA metabolism events. However, a functional link between exon 3 inclusion and changes in mRNA stability (286) or translational efficiency (467) could not be confirmed later on (522). Furthermore, exon 3 inclusion competes with *vpr*-mRNA production, raising the question of the role of splicing at 5'ss D3. During HIV-1 gene expression the pool of viral RNAs shifts from spliced, intronless mRNAs towards spliced, but still intron-retaining ones coding for e.g. the viral infectivity factor Vpr. However, the translational start codon of the Vpr open reading frame (ORF) is downstream of the exon 3 5'ss, making it necessary to render it splicing-incompetent for *vpr*-mRNA expression. Therefore it is not surprising that the splice site pairing between exon 3 5'ss D3 and either one of the downstream 3'ss (A3, A4cab, A5 and A7) basically is inefficient, leading to approximately only 2% *vpr*-mRNAs of all 4kb-mRNAs (521). One of the regulators obviously responsible for the relative weak recognition of exon 3 is the exonic splicing silencer (ESS), termed ESSV, which is localized downstream of 3'ss A2 (35, 131, 329). *In vitro* splicing analyses indicated that ESSV selectively represses 3'ss A2 by interference with the access of the general splicing component U2AF65 to the polypyrimidine tract (131). Furthermore, E complex EMSAs suggested that the U1 snRNP is stably associated with the 5'ss D3 even in presence of an intact ESSV (131). The relative poor match of 5'ss D3 to the consensus sequence on the one hand, but stable binding of the U1 snRNP – even in the presence of the ESSV – combined with previous observations from our lab (400) on the other, allowed us to speculate that 5'ss D3 recognition is supported by an enhancer in its proximity. Thus a working model was formulated that the U1 snRNP recruitment to the 5'ss might be positively regulated by an exonic splicing enhancer (ESE) in the vicinity of 5'ss D3 (Fig.III-19).

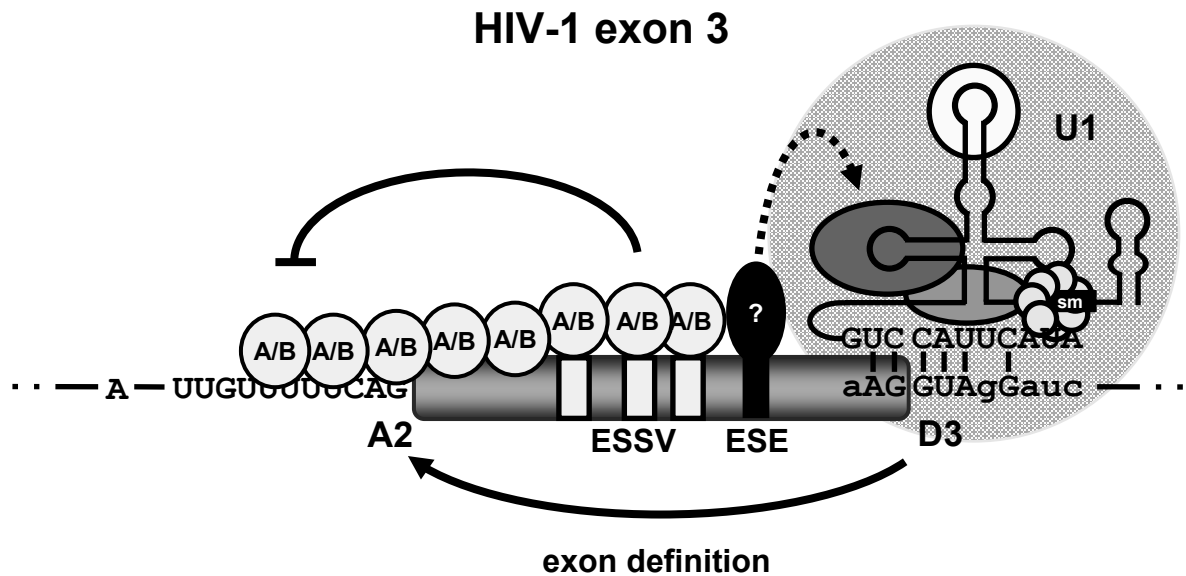


Fig.III-19: Working model of HIV-1 exon 3 splicing regulation

The exonic splicing silencer V (ESSV) contains three (Py)UAG motifs, which are bound by members of the hnRNP A/B protein family. hnRNP A/B proteins spread along the 5'-end of the exon via self-interactions and thereby occlude binding sites for general splicing factors such as U2AF65 (131). As a consequence, HIV-1 exon 3 is weakly recognized and *vpr*-mRNA production is downregulated to appropriate levels in the infected cell. However, U1 snRNP interaction with the 5'ss D3 is not impaired (131), but may be enhanced by a yet unidentified exonic splicing enhancer (ESE) located immediately downstream of the ESSV. The U1 snRNP binding to the 5'ss D3 might drive formation of exon definition complexes (EDCs) that antagonize the ESSV activity.

III.2.1 An exonic splicing enhancer (ESE) acts positively on the inclusion of the exon 3 into reporter mRNAs

In contrast to exonic splicing silencer (ESS) elements that inhibit the assembly of functional splicing complexes at nearby splice sites, exonic splicing enhancers (ESE) activate the splicing of exons in which they are embedded. Recently, an exonic splicing silencer (ESSV), which acts negatively on HIV-1 exon 3 inclusion into viral mRNAs by recruitment of hnRNP A/B proteins, was identified (35, 131). An intact ESSV is important for maintenance of the balance between spliced and unspliced viral mRNAs (329). Disruption of the ESSV by mutagenesis results in a higher level of HIV-1 exon 3 splicing and a decrease in quantities of unspliced viral mRNA

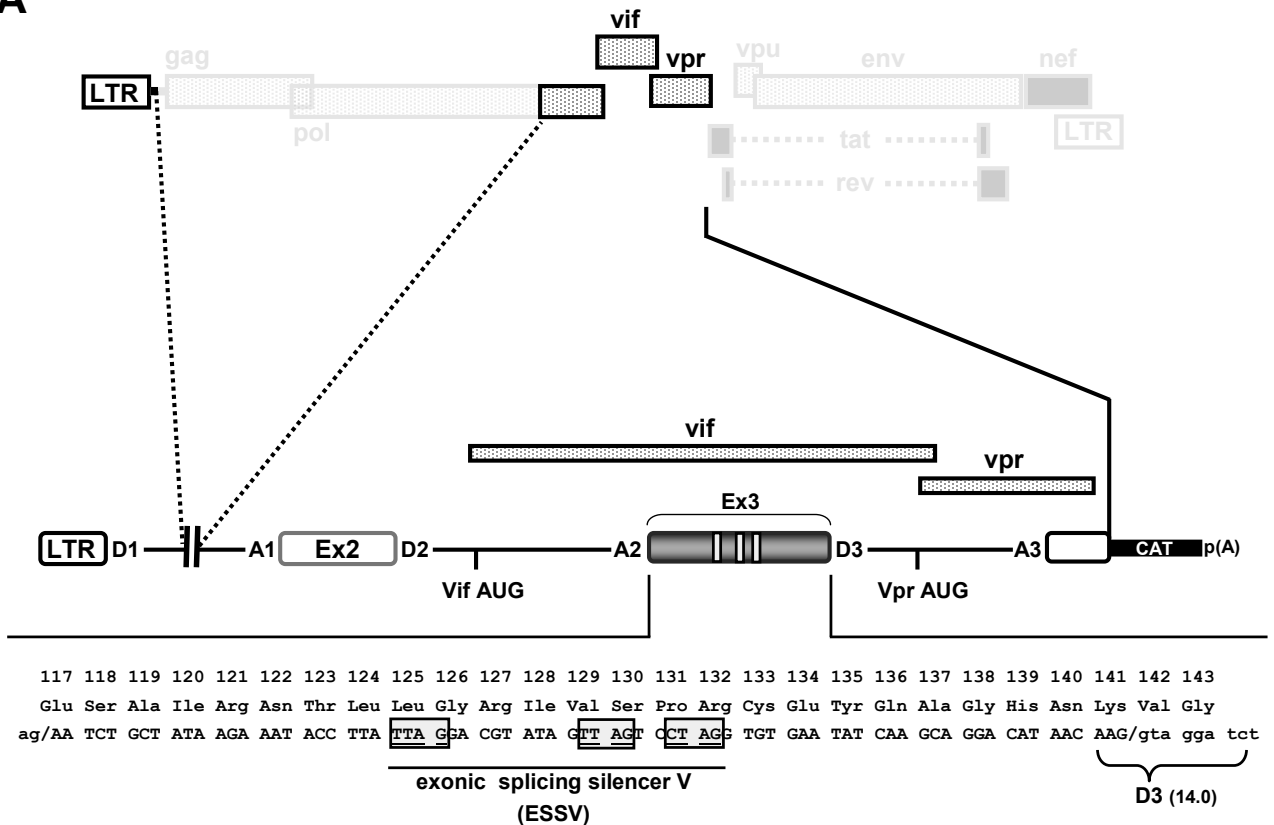
associated with a severe defect in virion production. The observed splicing profile was previously presumed to be entirely due to the removal of the repressing ESSV. However, 5'ss D3 (Hbs: 14.0) was shown to recruit U1 snRNP even in the presence of the dominant negative ESSV (131), suggesting that an exonic splicing enhancer facilitate 5'ss D3 recognition. To determine the extent to which 5'ss D3 recognition and the over-splicing phenotype in absence of a functional ESSV is linked to a putative ESE within HIV-1 exon 3, a subgenomic HIV-1 construct was generated derived from the viral isolate NL4-3 (Fig.III-20A). This HIV-1 based minigene is a shortened version of the proviral 5'-region. A deletion was introduced spanning most of the intron 1 sequences encoding the structural and enzymatic proteins of the provirus. The minigene contains the internal non-coding leader exons 2 and 3, surrounded by the shortened intron 1 and the two full-length introns 2 and 3. Transcription is under control of the viral 5'-LTR promoter and therefore requires Tat for its transactivation. An SV40 early polyadenylation site was placed downstream of the transcription unit to guarantee proper transcription termination and 3'-end processing of the reporter mRNAs (Fig.III-20A).

In order to suggest potential ESE sequences upstream of 5'ss D3, a novel *in silico* hexamer score was developed in collaboration with S. Theiss in an approach analogous to the RESCUE-ESE (147). From a dataset of 43,464 constitutively spliced canonical annotated human exons, 100 exonic and intronic nucleotides up- and downstream of 10,407 strong 5'ss (HBS>17.0) and 10,359 weak 5'ss (HBS≤13.5) were selected, excluding the proper splice site sequence (3+8=11 nt). From these, differential hexamer frequencies for all 4,096 hexamers in four disjoint datasets were obtained: exonic and intronic sequences in the vicinity of strong and weak 5'ss. These hexamer frequencies were transformed into two standard normal Z-scores Z_{EI} and Z_{WS} , measuring the difference in hexamer occurrence in exonic versus intronic sequences, and in exons with strong versus weak 5'ss. A hexamer with large Z_{EI} , e.g., occurred significantly more often upstream than downstream of 5'ss. In the next step, the nucleotide based scores Z_{EI} and Z_{WS} were calculated from these hexamer scores. Each single nucleotide is part of six neighboring hexamers – at either one of the six different hexamer positions. The index nucleotide **A** in the sequence TGTGA**A**TATCA, e.g., is located at the last position in hexamer TGTGA**A**, at the third position in hexamer GA**A**TAT and at the first position in **A**TATCA. For each (index) nucleotide, the two scores Z_{EI} and Z_{WS} were assigned as averages of

the respective scores for each of the six hexamers containing the index nucleotide. Plotting Z_{EI} along the sequence clearly depicted “exon-like” and “intron-like” regions as determined by their hexamer content (Fig.III-20B). The novel scores Z_{EI} and Z_{WS} also permit to *a priori* quantitatively assess point mutation effects on hexamer content. It was hypothesized that point mutations effectively changing “exon-like” to “intron-like” sequences upstream of 5'ss D3 were likely to also inactivate the suspected ESE.

Results

A



B

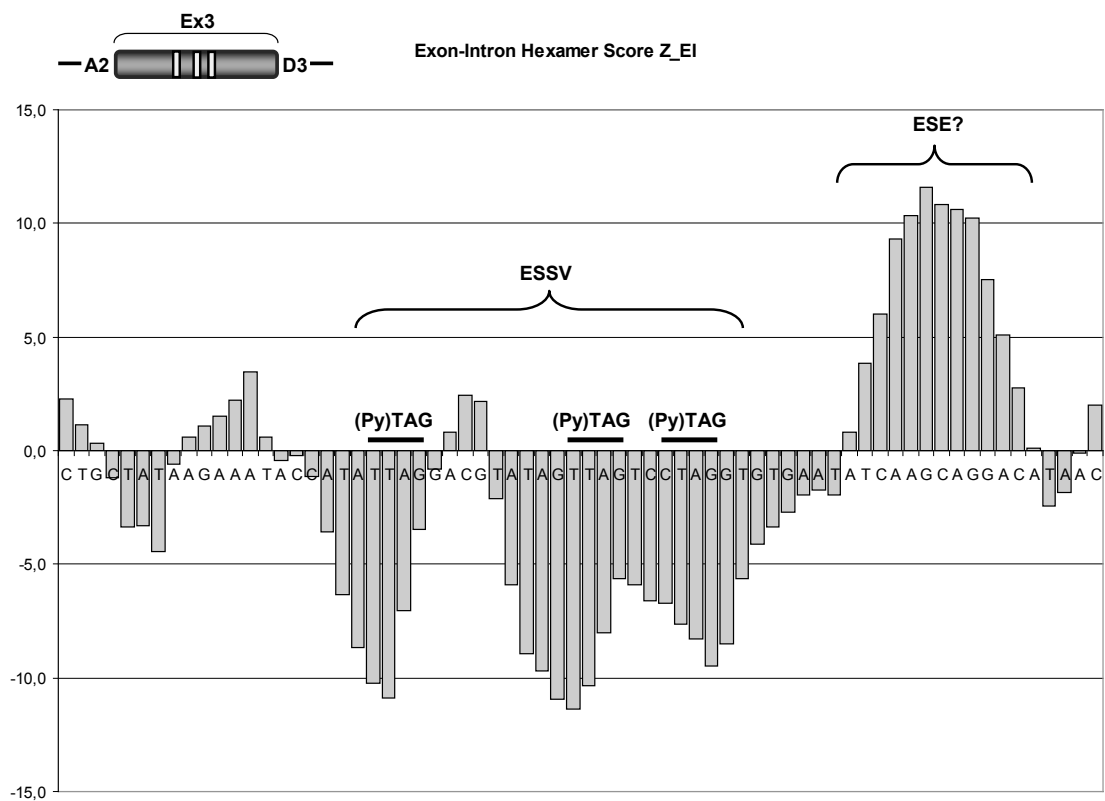


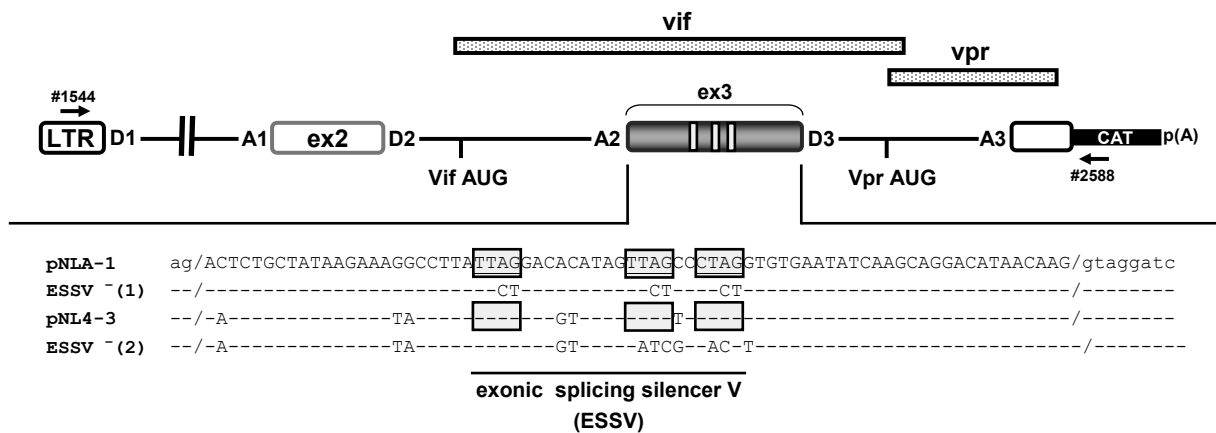
Fig.III-20: Construction of an HIV-1 based 4-exon-3-intron minigene to analyse exon 3 splice site regulation. [continued on next page]

The level of exon 3 inclusion into the transcripts driven by the reporter constructs was determined for each of the samples by semi-quantitative RT-PCR analysis of RNA extracted from transiently transfected HeLa-T4⁺ cells (Fig.III-21B). In the presence of a functional ESSV, exon 3 splicing was rarely detected, emphasizing the dominant-negative role of the ESSV for exon 3 recognition (Fig.III-21B, lane 1 and 2). However, mutations, which inactivate the ESSV function, either led to partial [ESSV⁻⁽¹⁾] or almost exclusive [ESSV⁻⁽²⁾] detection of exon 3-containing mRNAs (Fig.III-21B, lane 3-4).

Fig.III-20: continued

(A) Schematic drawing of the HIV-1 based LTR ex2 ex3 minigene construct: the reporter encompasses the two non-coding leader exons 2 and 3 flanked by pNLA1 sequences. Coding regions for *gag* and *pol* have been deleted and replaced by a short linker sequence. The resultant minigene formally contains the open reading frames for both viral infectivity factors Vif and Vpr. The HIV-1 exon 3 sequence is given below. Overlapping Vif amino acid residues are indicated above the exon 3 nucleotide sequence. The three (Py)UAG motifs of the ESSV are highlighted by grey boxes. The translational startcodons for the Vif and Vpr open reading frames are indicated. The intrinsic strength of 5'ss D3 was assessed by the HBond Score algorithm (HBS) and is indicated by the number in the bracket beside 5'ss D3. **(B)** Z_{EI}-Score plot. The X axis shows the exon 3 nucleotide sequence and the Y axis depicts the differences in score along the sequence. [LTR: long terminal repeat; Py: pyrimidine]

A



B

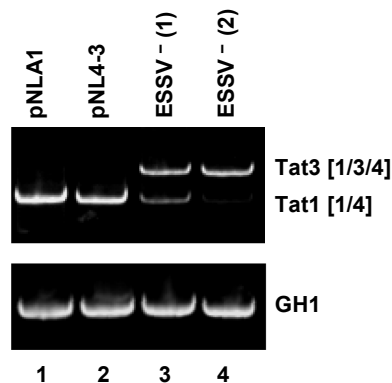


Fig.III-21: Mutational analyses of the ESSV in the context of pNLA1 and pNLA-3 derived exon 3 sequences

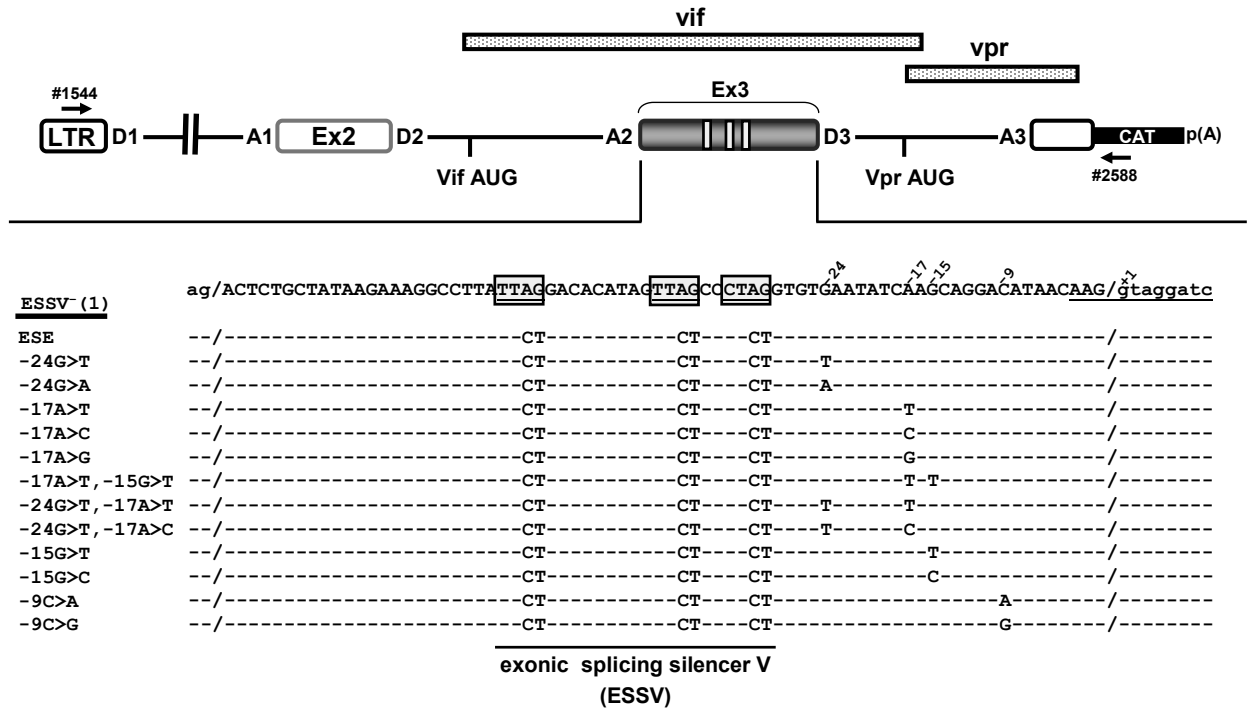
(A) Schematic diagram of the HIV-1 based LTR ex2 ex3 splicing reporter. The truncated pNLA-1 plasmid lacking the *gag* and *pol* genes derives from the laboratory proviral clone NL4-3. Exon 3 sequences are not entirely identical between pNLA-1 and pNLA-3, although the three (Py)UAG motifs are conserved (highlighted by grey boxes).

Mutations previously reported to disrupt ESSV activity (35, 329) were introduced by site-directed mutagenesis. Primers used in RT-PCR analyses are indicated by arrows. (B) 2.5×10^5 HeLa-T4⁺ cells were transiently cotransfected with 1 μ g of each of the LTR ex2 ex3 variants, 0.2 μ g SVctat expressing viral Tat protein to transactivate the LTR promoter and 1 μ g of pXGH5 to monitor equal transfection efficiencies. 30h post transfection total RNAs were isolated from the cells and revers transcribed employing oligo d(T) primer and SuperScriptTMIII. The cDNA obtained was used as template DNA for a PCR reaction with primer pair #1544/#2588 specifically detecting alternatively spliced isoforms by the RNA reporter. A separate PCR reaction was carried out using primer pair #1224/#1225 to monitor the level of constitutively spliced GH1 mRNA for each sample. Amplificates were resolved in 8% non-denaturing polyacrylamide gels and stained by ethidium bromide. Alternatively spliced mRNA isoforms are indicated at the right. [LTR: long terminal repeat; GH1: human Growth Hormone 1]

To systematically scan the HIV-1 exon 3 for a candidate ESE, those residues were mutated that had been predicted by the hexamer algorithm to be part of a putative enhancing sequence (Fig.III-22A). Therefore, the change ΔZ_{EI} induced by mutations at exon 3 positions between -24 and -9 was determined and the correlation between ΔZ_{EI} “prediction” and experimentally determined level of exon 3 inclusion was analyzed. The ESE mutations were tested in combination with ESSV⁻(1), which had an intermediate splicing phenotype (Fig.III-21B) and therefore allowed screening of both, positive and negative effects on exon 3 inclusion. Semi-quantitative RT-PCR of RNA from transiently transfected HeLa-T4⁺ cells was performed to determine the level of exon 3 splicing for each mutated minigene (Fig.III-22B). Transfection efficiency for each sample was monitored by coexpression of the constitutively spliced GH1 mRNA. The extent to which the individual ESE mutations could alter splicing was referred to “ESSV⁻(1)”, defining a baseline for the ratio of exon 3 inclusion (Tat3) to exclusion (Tat1) (Fig.III-22B, lane 2). Regardless of the varying outcome of the single mutations for the level of exon 3 splicing, the observed changes were mostly consistent with the predicted splicing phenotype. All the point-mutations “-24G>T”, “-24G>A”, “-17A>T” and “-15G>T” had in common to reduce the amount of exon 3 inclusion (Fig.III-22B, lanes 3-5 and 11), which was in agreement with their hexamer score distribution across the candidate ESE region. In addition, control mutations either not expected to disrupt the ESE (“-17A>C” and “-17A>G”) or introduced into positions outside of the candidate sequence (“-9C>A” and “-9C>G”) only had negligible impact on exon 3 splicing (Fig.III-22B, lanes 6,7,13 and 14). The reliability of the computational approach was supported by the combinatorial mutations “-17A>T; -15G>T”, “-24G>T; -17A>T” and “-24G>T; -17A>C”, which add up the effects for each single mutation (Fig.III-22B, lanes 9 and 10). Most importantly, the double mutation of “-24G>T” and “-17A>T” nearly abolished splicing of exon 3 (Fig.III-22B, lane 9). Predictions of changes in the hexamer score distribution are exemplified by bar graphs for “-24G>T; -17A>T” and each of the single mutations plotted against the “ESSV⁻(1)” reference in Appendix Fig.VII-1.3.

Results

A



B

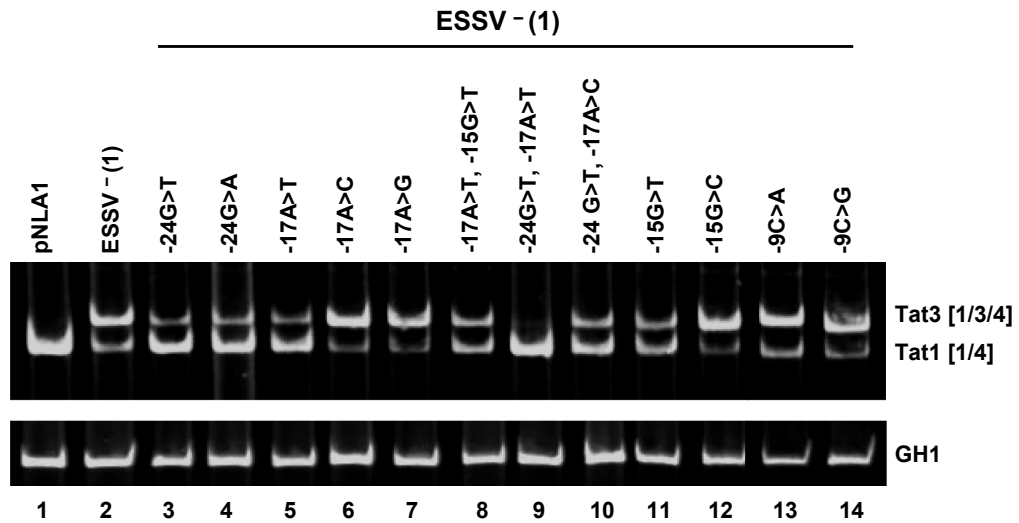


Fig.III-22: Detailed mutational analysis reveals an ESE sequence located within the 3'-terminal region of HIV-1 exon 3.

(A) Schematic drawing of the HIV-1 based LTR ex2 ex3 splicing reporter. The exon 3 and mutant sequences flanked by the intronic AG and GT dinucleotides are given below exon 3. Mutated ESE nucleotide residues are denoted by their position relative to the GT-dinucleotide. Positions of the RT-PCR primers are indicated by arrows. (B) 1µg of each of the constructs were transiently transfected into 2.5×10^5 HeLa-T4⁺ cells together with 0.2µg of SVctat and 1µg of pXGH5. 30h after transfection RNA was harvested from the cells and subjected to RT-PCR analyses with primer pairs #1544/#2588 and #1224/#1225 (GH1). PCR products were separated by 8% non-denaturing polyacrylamide gel electrophoresis and stained with ethidium bromide. HIV-1 RNA species are indicated on the right.

Up to this point, the exact nucleotide positions of an enhancing sequence within exon 3 was not absolutely certain. Increased splicing in absence of the ESSV might exclusively result in a relieved repression on the flanking splice sites, which theoretically opens the possibility that the double mutation (-24G>T; -17A>T) accidentally created a silencer instead of an inactivating enhancer. To discriminate between these two alternatives, the candidate ESE sequence (and mutant derivatives) were tested in the context of another HIV-1 based splicing reporter system in which the expression of spliced mRNAs is strictly dependent on an enhancer in the proximity of the 5'ss (Fig.III-23A, also see III.1.1). Enhancer-positive constructs accumulate spliced reporter mRNAs, which can easily be monitored by RT-PCR analysis and since those mRNAs encode for eGFP also by fluorescence microscopy or immunoblotting. This is different from plasmids, which do not have an enhancing sequence in the proximity of the 5'ss. Because here the U1 snRNP cannot efficiently bind to the 5'ss, the reporter mRNA cannot be stabilized and therefore is committed to degradation inside the nucleus of the transfected cells [Fig.III-1B, (a) and (b)]. However, reporter constructs were transiently transfected into human HeLa cells and analyzed by semi-quantitative RT-PCR to determine whether they can facilitate expression of spliced eGFP mRNAs. Indeed, the wildtype sequence considerably enhanced splicing, whereas both a splicing neutral control sequence (neutral3 (N3), also see Fig.III-3) and the double-mutant (-24G>T; -17A>T) failed to activate expression of spliced eGFP mRNAs (Fig.III-23B, lanes 1, 2 and 5). Both single-mutations were found to splice at a reduced level, as was predicted by the preceding experiment (Fig.III-23B, lanes 3 and 4). Immunoblot analysis revealed that accumulation of eGFP protein was exclusively observed for the wildtype sequence, while each of the single-mutations were already sufficient to abolish eGFP protein expression. Accordingly, splicing within two completely unrelated reporter systems confirmed that the HIV-1 exon 3 contains an ESE immediately downstream of the ESSV defined by the 16nt long core sequence "GAATATCAAGCAGGAC".

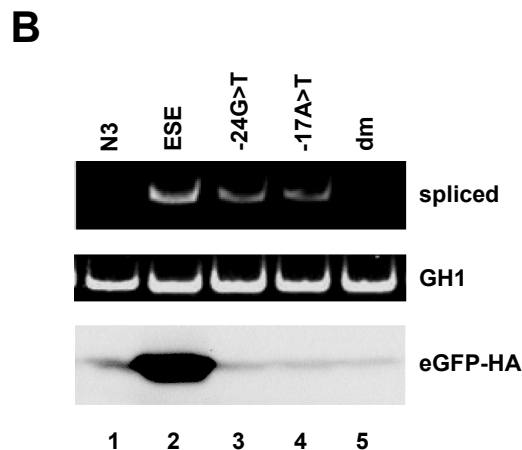
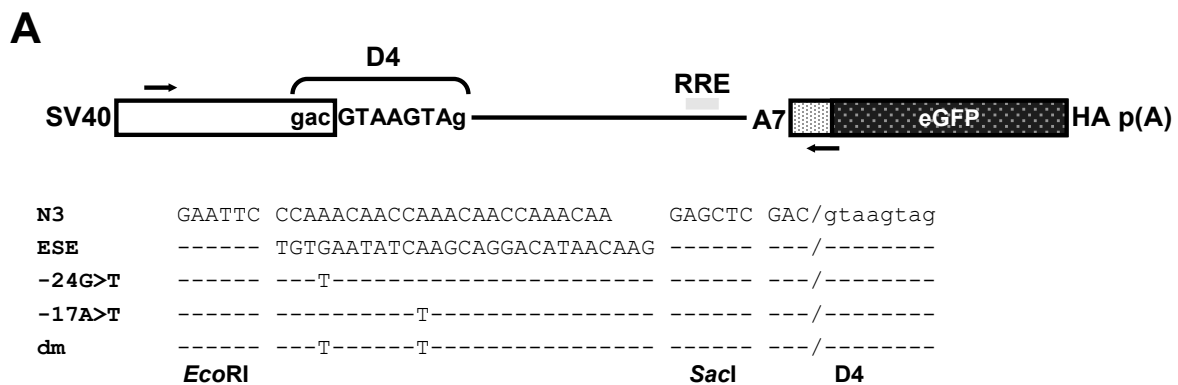


Fig.III-23: The putative ESE sequence supports 5'ss D4 recognition by the spliceosomal U1 snRNP.

(A) Schematic diagram of the SV-env/eGFP-HA expression plasmid carrying the env-eGFP fusion ORF. HIV-1 exon 3 ESE and mutant sequences were inserted upstream of 5'ss D4 using *EcoRI*/*SacI* restriction sites and are enlarged below. As a control, a splicing neutral sequence (N3, see chapter III-1)

was cloned upstream of 5'ss D4. (B) 2.5×10^5 HeLa-T4⁺ cells were transiently transfected with 1 μ g of each of the constructs and 1 μ g of pXGH5 to control for equal transfection efficiencies. 30h post transfection RNA was isolated from the cells and treated with DNase I to remove traces of plasmid DNA. Subsequently, RNA was revers transcribed and used in a PCR reaction with primer pair #3210/#3211 or #1224/#1225, respectively to specifically amplifying spliced eGFP- and GH1-mRNAs. RT-PCR products were resolved in an 8% non-denaturing polyacrylamide gel and stained with ethidium bromide. Cellular lysates were also collected from the cells and separated by 10% SDS-PAGE. After transfer to a nitrocellulose membran, samples were probed with an antibody directed against the HA tag to specifically detect eGFP-HA expressed in the transfected cells. [SV40: Simian Virus 40 early promotor; RRE: Rev Responsive Element; HA: Hemagglutinin epitope; dm: double mutant; GH1: human Growth Hormone 1]

III.2.2 Inactivation of the exonic splicing enhancer restores viral replication in the context of an ESSV-negative provirus

III.2.2.1 Silent mutations chosen not to affect the overlapping open reading frame for the viral Vif protein disrupt the enhancer activity

To determine the importance of this ESE for viral splicing regulation, mutations were tested, predicted (by the hexamer algorithm) to decrease exon 3 inclusion, but which simultaneously do not affect the amino acid sequence of the overlapping HIV-1 *vif* reading frame. Two mutations were chosen of which one should selectively disrupt a TDP-43 “ugug”-binding motif (-25T>C; also see Fig.III-30) and the other was predicted to reduce exon 3 inclusion using the hexamer algorithm described in III.2.1 (-16A>G) (see Appendix Fig.VII-1.5). Both represented transitions in the third-base wobble positions and therefore should maintain the respective Vif amino acids C133 and G136 (Fig.III-24A). Additionally, silent ESE mutants were tested in the context of an ESSV-negative sequence [ESSV⁻(2)], maintaining the overlapping Vif amino acid sequence (Fig.III-24A). HeLa cells were transiently transfected with the parental and mutant minigenes to examine their exon 3 splicing phenotype by semi-quantitative and quantitative RT-PCR analyses (Fig.III-24B-C). In the presence of the ESSV almost the entire reporter transcripts underwent splicing without inclusion of exon 3 (Fig.III-24B, lane 1). Consequently the mutations within the ESE sequence had no or only little effects (Fig.III-24B, lanes 2-4). Interestingly, the single mutation “-25T>C” within the TDP-43 binding site rather increased inclusion of exon 3, which was in agreement with the prediction of the algorithm (see Appendix Fig.VII-1.5), but argued against a positive role of the “ugug” sequence bordering ESSV and ESE for splicing (Fig.III-24B, lane 2). As expected, mutation of the ESSV relieved the repression of the flanking splice sites and caused a strong increase in exon 3 splicing (Fig.III-24B, lane 5). However, point-mutation “-16A>G” and the double-mutant “-25T>C; -16A>G” could efficiently reverse this splicing phenotype and shifted back to those reporter transcript isoforms lacking exon 3 (Fig.III-24B, lanes 7 and 8). The relative amounts of exon 3 including or excluding reporter mRNA species were also quantified using real-time PCR assays with primers specific for Tat3 and Tat1 splice junctions (Fig.III-24C).

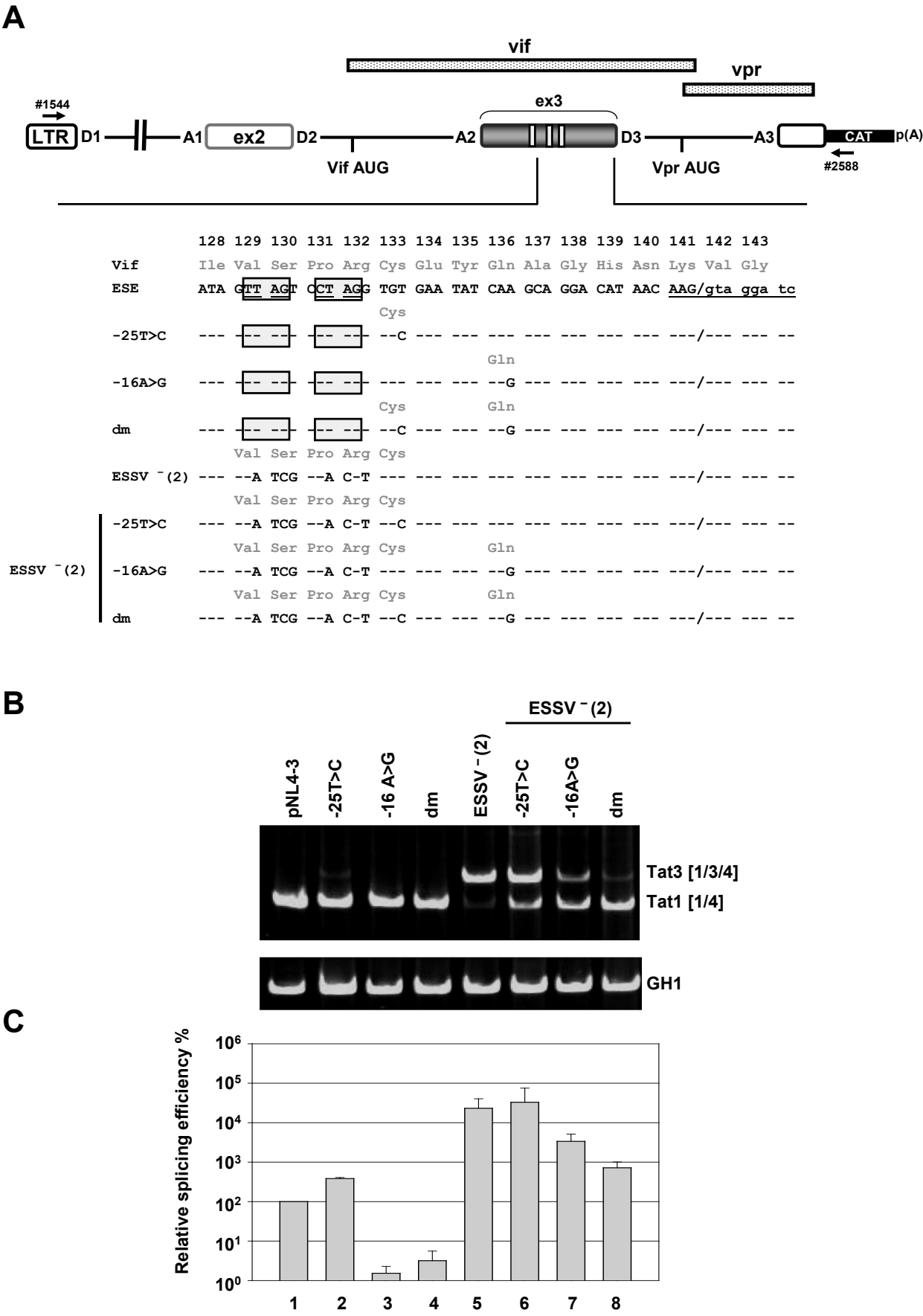


Fig.III-24: A silent double-mutation within the ESE sequence abolishes HIV-1 exon 3 inclusion.
[continued on next page]

For normalization, the total level of minigene transcripts in each sample was measured with a primer pair detecting HIV-1 exon1. Again, the single-mutation “-25T>C” rather seemed to slightly enhance exon 3 inclusion in the presence of the ESSV (~4-fold) (Fig.III-24C, lane 2). Furthermore, it moderately compensated for the reduction of exon 3 splicing by “-16 A>G” when combined in a double-mutant (Fig.III-24C, lanes 3 and 4). However, once the ESSV was mutated, this led to a more than 200 fold higher ratio of Tat3 versus Tat 1 mRNAs (Fig.III-24C, cf. lane 1 and 5), confirming the dominant negative activity on exon 3 inclusion emanating from the ESSV. The single mutation “-16A>G” substantially decreased exon 3 inclusion into the reporter mRNAs in the absence of the ESSV (Fig.III-24C, lane 7). Interestingly, in the context of ESSV-lacking mRNAs, “-25T>C” once more slightly potentiated the decrease in exon 3 splicing exerted by “-16A>G” (Fig.III-24C, lane 8). RT-PCR analyses therefore suggested a relatively minor role for the “ugug” TDP-43 binding motif in exon 3 splicing, while the point mutation “-16A>G” provided further evidence for an ESE activity within the sequence immediately downstream.

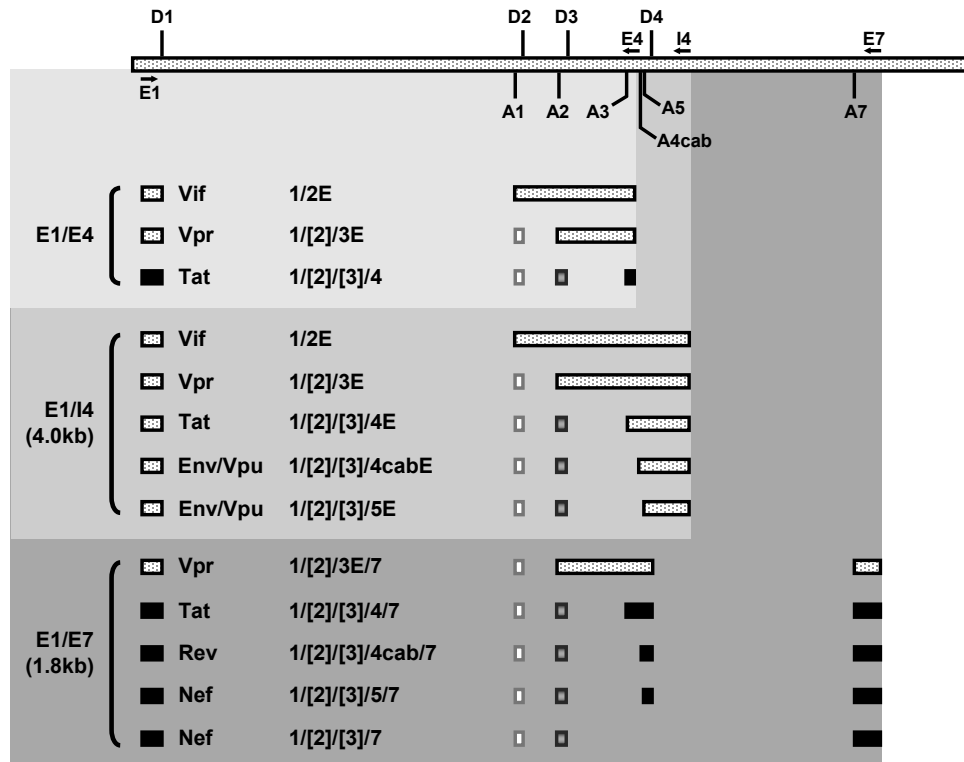
Fig.III-24: [continued]

To extend the analysis to replication-competent virus, it was screened for additional mutations predicted to inactivate the ESE, which did not alter the amino acid sequence of the overlapping *vif* open reading frame and thus, exclude splicing effects due to different levels of Vif protein expression. **(A)** Schematic drawing of the pNL4-3 derived LTR ex2 ex3 splicing reporter. Wildtype and mutated exon 3 sequences are depicted below. The three (Py)UAG motifs of ESSV are highlighted by grey boxes. Single point-mutations are denoted by their relative position to the GT-dinucleotide of 5'ss D3. RT-PCR primers are indicated by arrows. **(B)** RT-PCR analysis of RNA taken from HeLa cells after transfection with 1µg of the HIV-1 exon 3 splicing reporter LTR ex 2 ex3 or mutant derivatives, 0,2µg of SVctat and 1µg of pXGH5 for transfection control. RT-PCR products for the alternatively spliced reporter mRNAs and GH1 as control for transfection efficiency were resolved in an 8% non-denaturing polyacrylamide gel and stained with ethidium bromide. Spliced *tat*-mRNA isoforms are indicated on the right. **(C)** Quantitative SBYR RT-PCR assays performed with primer pairs specifically detecting Tat3 (#3397/#3633) and Tat1 (#3631/#3632) mRNA species. cDNA samples were prepared as described in **(B)**. The graph depicts the mean values of three independent RT-PCR experiments in a logarithmic scale. Tat3/Tat1 ratio of wildtype exon 3 sequence (lane 1) is set as 100%. [dm: double mutation; GH1: human Growth Hormone 1; RT: Reverse Transcriptase].

III.2.2.2 Balanced exon 3 splicing is under combinatorial control of the ESSV and the adjacent ESE_{vpr}

The following experiments aimed to figure out the role of the enhancer (from here on referred to as ESE_{vpr}) in the context of a replication-competent virus. Therefore, the silent mutations shown to decrease the enhancer activity were introduced into the proviral clone pNL4-3 (GenBank Accession No. [M19921](#)), which affords a suitable and robust model system to study HIV-1 replication. HeLa and HEK293T cells transfected with the proviral plasmid facilitate the release of virus particles into the culture medium (2). However, due to the lack of the HIV-1 receptor on their cellular surface these cells could not be re- or superinfected. On the other hand, both cell lines express the whole repertoire of cellular proteins required for regulation of HIV-1 alternative splicing. Exon 3 splicing was analyzed by semi-quantitative RT-PCR using primer pairs specific for Tat-, intron-containing or fully spliced mRNA species, respectively (Fig.III-25A). Total-RNA was isolated from HEK 293T cells transfected with infectious pNL4-3 plasmids and subsequently, the splicing pattern was analyzed by semi-quantitative RT-PCR (Fig.III-25B). In ESSV-positive clones, the single mutation “-16A>G” as well as the double-mutation “-25T>C; -16A>G” (dm) caused a decrease in exon 3 inclusion into the *tat*-, *nef*- and *env*-mRNAs (Tat3, Env8, Nef4, Rev7+8) (Fig.III-25B, lanes 1-4). In addition, *vpr*-mRNAs were rarely detected after insertion of either the “-16 A>G” single mutation or the double-mutation into the provirus (Fig.III-25B, lanes 1-4). As expected, the exon 3-containing mRNAs within both viral mRNA classes accumulated following disruption of the ESSV (Fig.III-25B, cf. lanes 1 and 5). Consistently, the *nef*-, *rev*-, *tat*- and *env*-mRNA species shifted to their respective exon 3-containing isoforms. Furthermore, an increased level of *vpr*-mRNA splicing could be detected for the ESSV-negative provirus (Fig.III-25B, cf. lanes 1 and 5). However, exon 3 splicing was gradually decreased almost all to wildtype levels, starting from “-25T>C”, followed by “-16A>G” and then the double-mutation (Fig.III-25B, lanes 6-8).

A



B

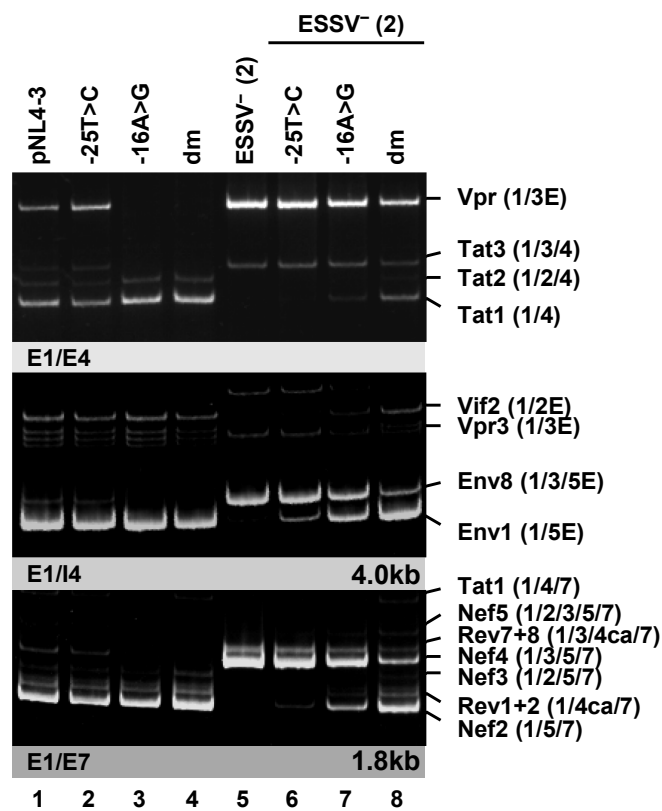


Fig.III-25: Inactivation of the ESE_{vpr} by mutagenesis restores appropriate exon 3 splicing in the context of the ESSV-negative replication-competent virus. [continued on next page]

Fig.III-25: continued

Silent-mutations were introduced into the infectious proviral plasmid pNL4-3. **(A)** Schematic representation of the RT-PCR primer set ups employed to specifically detect the levels of exon 3 inclusion into the main HIV-1 mRNA species. Viral splice sites and primer positions (arrows) are indicated above the viral genomic RNA on top of the figure. Primer pair E1/E4 was used to specifically detect exon 3 splicing into *tat*-mRNA species (see Appendix Fig.VII-1.6). Primer pairs E1/I4 and E1/E7 were utilized to monitor exon 3 inclusion within the 4.0kb and 1.8kb size major mRNA classes, respectively. Exons in each mRNA are designated by rectangles [black: early viral genes; patterned: late viral genes]. Either one or both of the two non-coding leader exons 2 and 3 can be spliced into each of the different viral mRNA species [white: exon 2; shaded grey: exon 3]. Exon arrangements are denoted for each mRNA species. **(B)** 2.5×10^5 HEK 293T cells were transiently transfected with 1 μ g of each of the infectious plasmids. 30h after transfection, total-RNA was isolated from the cells and subjected to RT-PCR analyses with the primer pairs from (A). RT-PCR products were separated by 8% non-denaturing polyacrylamide gel electrophoresis and visualized using ethidium bromide staining. HIV-1 mRNA species are indicated on the right and correspond to the nomenclature published previously (425).

Results obtained from the RT-PCR analysis demonstrated that the ESE contributes to regulated exon 3 inclusion into each of the viral mRNA species. To thoroughly examine the ESE_{vpr} for its impact on regulation of HIV-1 exon 3 splicing, quantitative RT-PCR analyses was performed. Different primer pairs were used to specifically quantitate the relative levels of viral unspliced, spliced, *vpr*-, *vif*- and exon 3 containing mRNAs (see Appendix, Fig. VII-1.7). With regard to differences in the overall amount of viral mRNAs a primer pair was designed directed against terminal HIV-1 exon 7 to normalize the relative levels of each mRNA species (see Appendix, Fig. VII-1.7; #3387 and #3388). Quantitative RT-PCR assays showed that the ESE_{vpr} mutations did not significantly alter the levels of unspliced, spliced and *vif*-mRNA in the context of the ESSV-positive virus [Fig.III-26, (a)-(c), lanes 1-4]. However, the single point mutation “-16A>G” alone was capable to downmodulate the relative amount of *vpr*-mRNA indicating that the ESE_{vpr} is important for activation of 3'ss A2 even in the presence of the dominant-negative ESSV [Fig.III-26, (d), lanes 3 and 4]. Furthermore, the levels of exon 3-containing mRNA species were greatly reduced [Fig.III-26, (e), lanes 3 and 4]. In line with previous work (329), the ESSV disruption resulted in a large reduction (~10-20 fold) of unspliced mRNA [Fig.III-26, (a), cf. lanes 1 and 5]. Moreover, the relative amount of multiply spliced-mRNA was upregulated by approximately 10-fold [Fig.III-26, (b), cf. lanes 1 and 5]. In addition, loss of ESSV function by mutagenesis induced a great decrease in *vif*-mRNA levels up to 20-fold [Fig.III-26, (b), cf. lanes 1 and 5]. By contrast, expression of Vpr- and exon 3-containing viral mRNAs was detected at strongly elevated levels [Fig.III-26, (d) and (e), cf. lanes 1 and 5]. These results were consistent with recent studies, showing that disruption of the ESSV causes a dramatic deregulation of viral splicing. However, second site mutations within the ESE_{vpr} element could compensate for the lack of the ESSV activity [Fig.III-26, (a)-(e), lanes 6-8]. ESE_{vpr} double-mutants restored at least normal levels of unspliced, spliced and *vif*-mRNAs (Fig.III-26, (a)-(c), cf. lane 1 and 8). Surprisingly, the *vpr*-mRNA remained present in a relatively high abundance despite the ESE_{vpr} double mutation (Fig.III-26, (d), cf. lane 1 and 8), suggesting either residual enhancer activity or that activation of 3'ss A2 in the absence of the ESSV is to some extent unrelated to this ESE_{vpr}. The former was supported by the finding that the expression of exon 3-including mRNAs also not entirely returned to normal levels (Fig.III-26, (e), cf. lane 1 and 8).

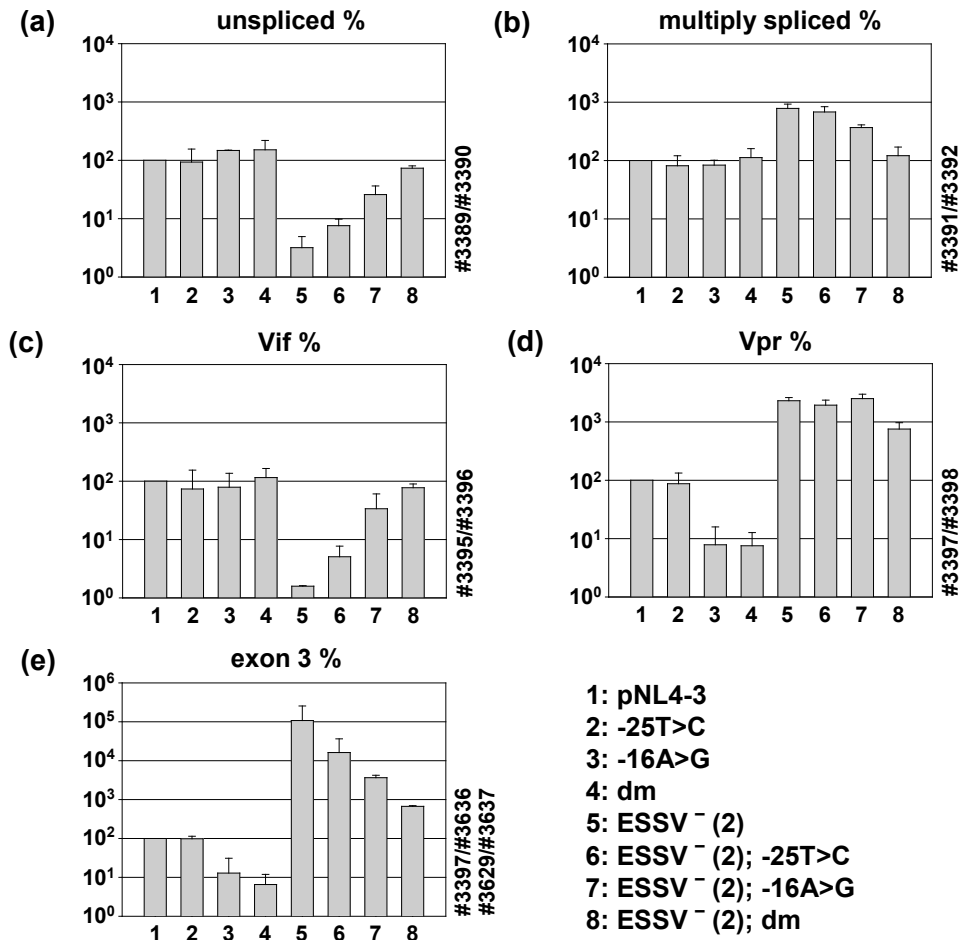


Fig.III-26: Inactivation of the ESE_{vpr} by mutagenesis restored appropriate exon 3 splicing in the context of ESSV negative replication-competent virus.

cDNA samples were prepared as described in Fig.III-25B and used in realtime PCR assays to specifically quantitate the relative abundance of unspliced (a), multiply spliced (b), Vif (c), Vpr (d) or exon 3 including mRNA species (e). For normalization primers P1 and P2 were used, detecting the total content of viral mRNA for each sample. Data represent expression relative to wildtype pNL4-3 (lane 1), which is set equal to 100%. Values and error bars show the average \pm standard deviation of three independent transfection experiments.

However, experiments were complemented by western blot analyses of intracellular viral proteins (Fig.III-27). To evaluate the importance of this ESE_{vpr} for viral replication, cell-free supernatants were harvested from HEK293T cells transfected with infectious HIV-1 plasmids and the viral release for each sample was specifically detected via p24 levels. In correspondence with the data obtained from the real-time PCR assays, western blot results revealed relatively unaltered intracellular levels of Gag and Vif proteins for the ESE_{vpr} mutants in the presence of the ESSV (Fig.III-27, lanes 2-5). Furthermore, viral replication seemed not to be significantly impaired

since comparable levels of viral capsid (CA, p24) protein could be detected within the supernatant (Fig.III-27, lanes 2-5), whereas the mutation of the ESSV expectedly led to low amounts of intracellular Gag-proteins and Vif (Fig.III-27, lane 6). Moreover, a striking defect in Gag processing could be observed, characterized by loss of the gag-precursor p55 cleavage products p41 and p24 (Fig.III-27, lane 6). As expected from the RT-PCR results, the expression of Vpr protein was drastically increased (Fig.III-27, lane 6). In agreement with previous studies, the level of p24 within the supernatant was severely decreased for the ESSV mutant, suggesting a defective replication (Fig.III-27, lane 6). This was supposed to result from insufficient amounts of intracellular Gag needed to drive virus assembly at the cellular plasma membrane. The production and normal processing of intracellular Gag and expression of Vif were restored as soon as the ESE_{vpr} was double-mutated as well (Fig.III-27, lane 9). Additionally, the “-16 A>G” mutation and the double-mutation were able to overcome defects in virion production due to the absence of the ESSV (Fig.III-27, lanes 8 and 9). These results indicated that the novel enhancer positively acts on the flanking splice sites A2 and D3 and that it is involved in regulated HIV-1 exon 3 splicing, which is pivotal for viral replication. Additionally, these observations demonstrated that a functional enhancer is critical for the expression of viral Vpr protein, noteworthy even when the ESSV is active, indicating that in the interplay between ESSV and ESE_{vpr}, the latter has even greater effects on viral splicing than expected.

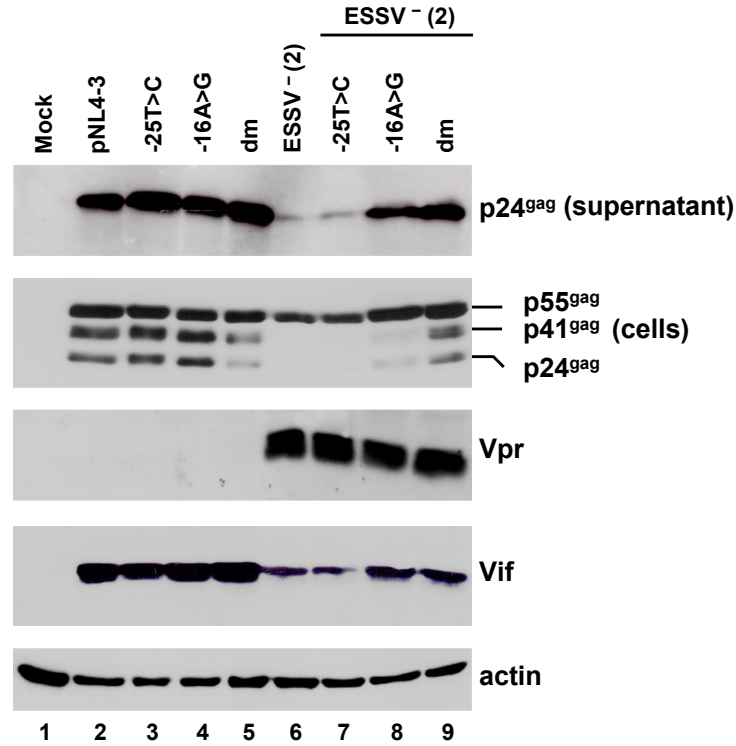


Fig.III-27: Inactivation of ESE_{vpr} by mutagenesis restores appropriate exon 3 splicing in the context of ESSV negative replication-competent virus.

2.5 x 10⁵ HEK 293T cells were transiently transfected with 1µg of each of the infectious plasmids. 48h post transfection viral supernatants were collected, layered onto 20% sucrose solution and centrifuged at 28.000 rpm for 1.30h at 4°C to pellet the released virions. In addition, cells were harvested and resuspended in lysis buffer. Supernatants and cellular lysates were resolved in 12% SDS-PAGE and electroblotted on nitrocellulose membranes. To determine virus particle production and the expression of viral proteins, samples were probed with primary antibodies specifically detecting structural p24^{gag} (CA) and the viral infectivity factor Vif and Vpr. Equal amounts of cell lysates were controlled by detection of β-actin. [E: Extended exon; dm: double mutation; HEK: Human Embryonic Kidney; RT: Reverse Transcriptase; CA: Capsid protein]

III.2.3 Recognition of to the viral 5'ss D3 is critical for viral Vpr expression

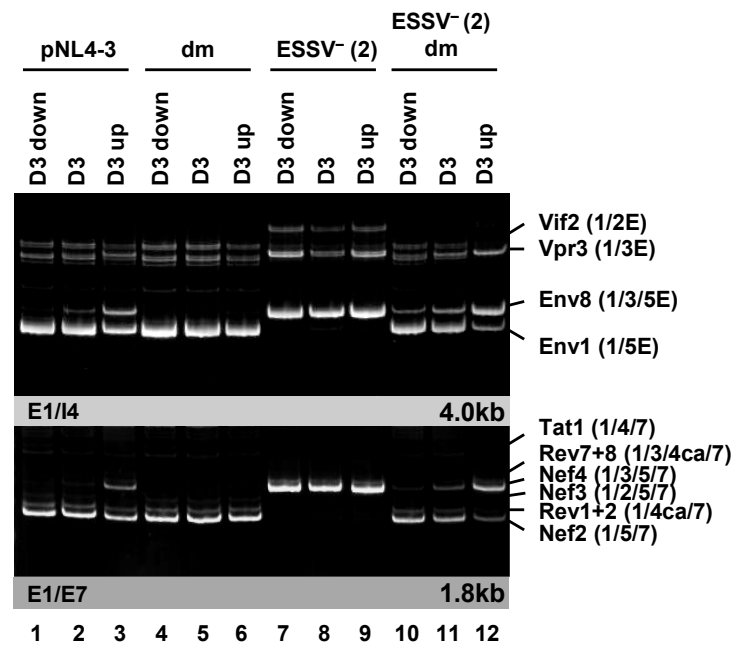
III.2.3.1 *Vpr*-mRNA expression can be modulated by up and down mutations of 5'ss D3

Use of 3'ss A2 results in the formation of *vpr*-mRNA, but only when splicing at the downstream 5'ss D3 is suppressed because this would remove the Vpr translational initiation codon within intron 3 from the mature transcripts. Remarkably, the ESE_{*vpr*} was shown to be critical for *vpr*-mRNA expression, although it is located close to 5'ss D3 and separated from 3'ss A2 by the repressing ESSV. It was found previously that efficient recognition of a 5'ss by the U1 snRNP exerts a positive feedback on the assembly of splicing factors at the upstream 3'ss – most likely via interactions across the exon (11, 215, 337). To analyse the interdependence between ESE_{*vpr*}, 5'ss D3 and *vpr*-mRNA expression, mutations predicted to either decrease (D3 down) or increase (D3 up) the intrinsic strength of the viral 5'ss D3 were tested in the context of a replication-competent provirus (Fig.III-28A). Mutations were chosen so that the overlapping Vif open reading frame was not changed. Following transient transfection of HEK 293T cells with proviral DNA, the exon 3 abundance within the viral mRNA species was determined for each of the mutant proviruses by RT-PCR analysis (Fig.III-28B). As expected, the extent of the complementarity between the U1 snRNA and the 5'ss was in principle directly correlated with the amounts of exon 3 present in the viral transcripts (Fig.III-28B). In the presence of the ESE_{*vpr*} activity weakening 5'ss D3 caused a decrease in the levels of exon 3-containing isoforms within both major viral mRNA classes, whereas an increase in the complementarity of 5'ss D3 partially overcame the general repression of exon 3 splicing by the dominant-negative ESSV (Fig.III-28B, lanes 1-3). This was in line with the hypothesis that the stability of U1 snRNP binding to a 5'ss plays a pivotal role for recognition of the entire exon. However, ESE_{*vpr*} mutants showed no detectable exon 3 inclusion regardless of their intrinsic 5'ss strength (Fig.III-28B, lanes 4-6), indicating a strict requirement for a functional ESE_{*vpr*} to enable stable binding of the U1 snRNP to 5'ss D3.

A

	141	142	143	144		HBs	Maxent
Vif	Lys	Val	Gly	Ser			
D3	AAG/gta	gga	tct	aAGGUAgGauc		14.00	9.45
	Lys	Val	Gly	Ser			
D3 down	AAG/gtt	ggt	tct	aAGGUugGUuc		11.70	8.46
	Lys	Val	Gly	Ser			
D3 up	AAG/gta	ggt	agc	aAGGUAgGUAg		18.00	10.29

B



C

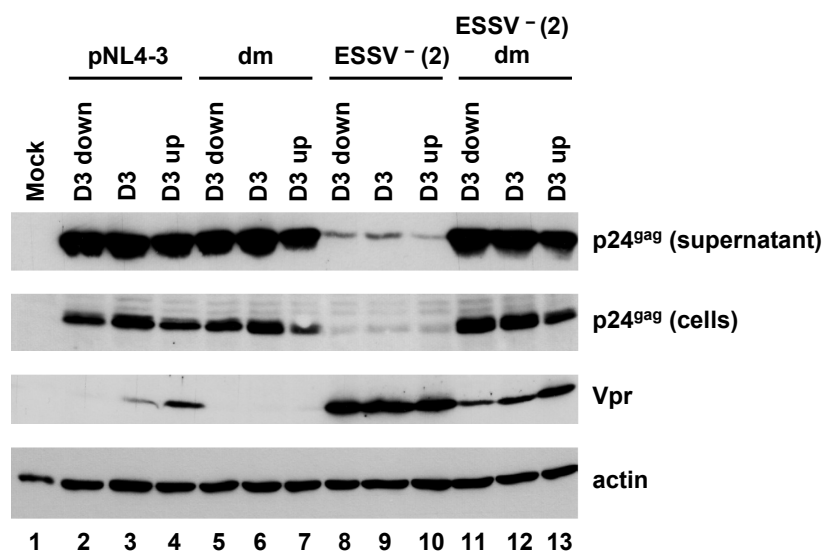


Fig.III-28: 5'ss D3 up and down mutations modulate HIV-1 exon 3 splicing and *vpr*-mRNA formation. [continued on next page]

The ESSV mutant was also associated with a poor response to the distinct 5'ss variants and efficient exon 3 inclusion was detected in each case irrespective of up or down mutations within D3 (Fig.III-28B, lanes 7-9), suggesting that in absence of the ESSV activity recognition of a weaker 5'ss can be equalized by stronger activation of 3'ss A2. Finally, when both ESSV and ESE_{vpr} were mutated the 5'ss strength again up- or downmodulated the frequency of exon 3 inclusion into the viral mRNAs (Fig.III-28B, lanes 10-12), This reinforces the notion that in a less favorable environment with regard to the enhancer strength, efficiency of exon inclusion exhibits a higher dependency on the capability of a splice site to bind the U1 snRNP on its own. Taken as a whole, overall efficiency of HIV-1 exon 3 splicing is adjusted by the individual strength of the preceding exonic splicing regulatory elements and the 5'ss D3. This is reminiscent to the observations made earlier, namely that the overall ability of the spliceosome to recognize a splice site is essentially controlled by the splice site itself and the sequence context in which it is embedded (Fig.III-6). Western blot analyses were consistent with these results and revealed that Vpr expression is under combinatorial control of ESSV, ESE_{vpr} and 5'ss D3 (Fig.III-28C).

Fig.III-28: continued

(A) Silent mutations predicted to decrease or increase the complementarity to the 5'-end of the endogenous U1 snRNA were introduced into the viral 5'ss D3 (*left*). Exonic nucleotides are denoted in upper case letters, intronic nucleotides in lower case letter. Complementarity and predicted intrinsic strength by HBond Score (HBS) and MaxEnt Score algorithms are each depicted at the right. Nucleotides complementary to the U1 snRNA are indicated by capital letters, while mismatches to the U1 snRNA are given in lower case letters. **(B)** 2.5×10^5 HEK 293T cells were transiently transfected with 1 μ g of each of the different infectious clones. RNA was isolated from the cells, DNase I digested and revers transcribed. Resultant cDNA served as DNA template in semi-quantitative PCR reactions using primer pairs E1/I4 and E1/E7 to specifically detect viral 4.0kb and 1.8kb size viral mRNAs, respectively. Proviral mutants are given above the panels. The main HIV-1 mRNA species are indicated at the right of the gels. **(C)** Protein lysates and viral supernatants were collected from HEK 293T cells transfected with 1 μ g of pNL4-3 or mutant derivates. Samples were loaded on 12% SDS polyacrylamide gels and after separation transferred to nitrocellulose membranes. Viral proteins and β -actin as a loading control were determined by probing with specific primary antibodies. For detection appropriate HRP-conjugated antibodies and ECLTM detection reagent were applied. [HBS: HBond Score; MaxEnt: MaxEnt Score; dm: double mutation; E: Extended exon]

Correspondingly, optimization of 5'ss D3 was accompanied by a higher abundance of Vpr protein within the transfected cells, whereas a reduction of the intrinsic strength showed the opposite effect on Vpr expression (Fig.III-28C, lanes 2-4 and 11-13). However, in presence of only one intact SRE, ESE_{vpr} or ESSV, exon 3 splicing efficiency was either too low or too high to allow tuning by alterations of the 5'ss strength (Fig.III-28C, lanes 5-7 and 8-10). These results substantiated the observations that U1 snRNP binding to 5' ss D3 enhances the use of the upstream 3'ss A2 and that splicing of exon 3 can be considered as the integrated outcome of exonic elements (ESE_{vpr} and ESSV) and intrinsic strength of 5'ss D3.

III.2.3.2 Binding of the U1 snRNP to a non-functional 5'ss is sufficient to augment splicing at the upstream 3'ss A2

In principle, these results suggested that U1 snRNP binding to the 5'ss has two primary functions during pre-mRNA splicing, i) it enhances the formation of exon definition complexes and therefore promotes recognition of the upstream 3'ss (11, 215, 337) and ii) it commits the bound 5'ss to splice site pairing with a 3'ss across the downstream intron in the pre-spliceosome (308). It is here hypothesized that *vpr*-mRNA splicing represents a case in which the former is required, whereas the latter has to occur at least with lower efficiency. To gain a broader understanding of how *vpr*-mRNA expression is regulated by U1 snRNP binding, 5'ss D3 was replaced by a sequence, termed "GTV". This highly to U1 snRNA complementary sequence has been shown to be a splicing-incompetent U1 snRNP binding site due to a G-to-C mutation at position +1 (253) (Fig.III-29A). To control that the GT mutation itself was not responsible for the phenotype, 5'ss D3 was inactivated by a G-to-C mutation at position +1 without increasing its complementarity to the U1 snRNA. The resulting set of variants ranging from a functional 5'ss, which supports efficient binding of the U1 snRNP as well as splicing (D3) to sequences either allowing only binding (GTV) or neither binding nor splicing (D3 +1G>C), were tested in the context of ESSV-negative or ESSV/ ESE_{vpr}-double negative proviruses. Splicing efficiency at 3'ss A2 was determined by semi-quantitative RT-PCR analyses (Fig.III-29B).

Results

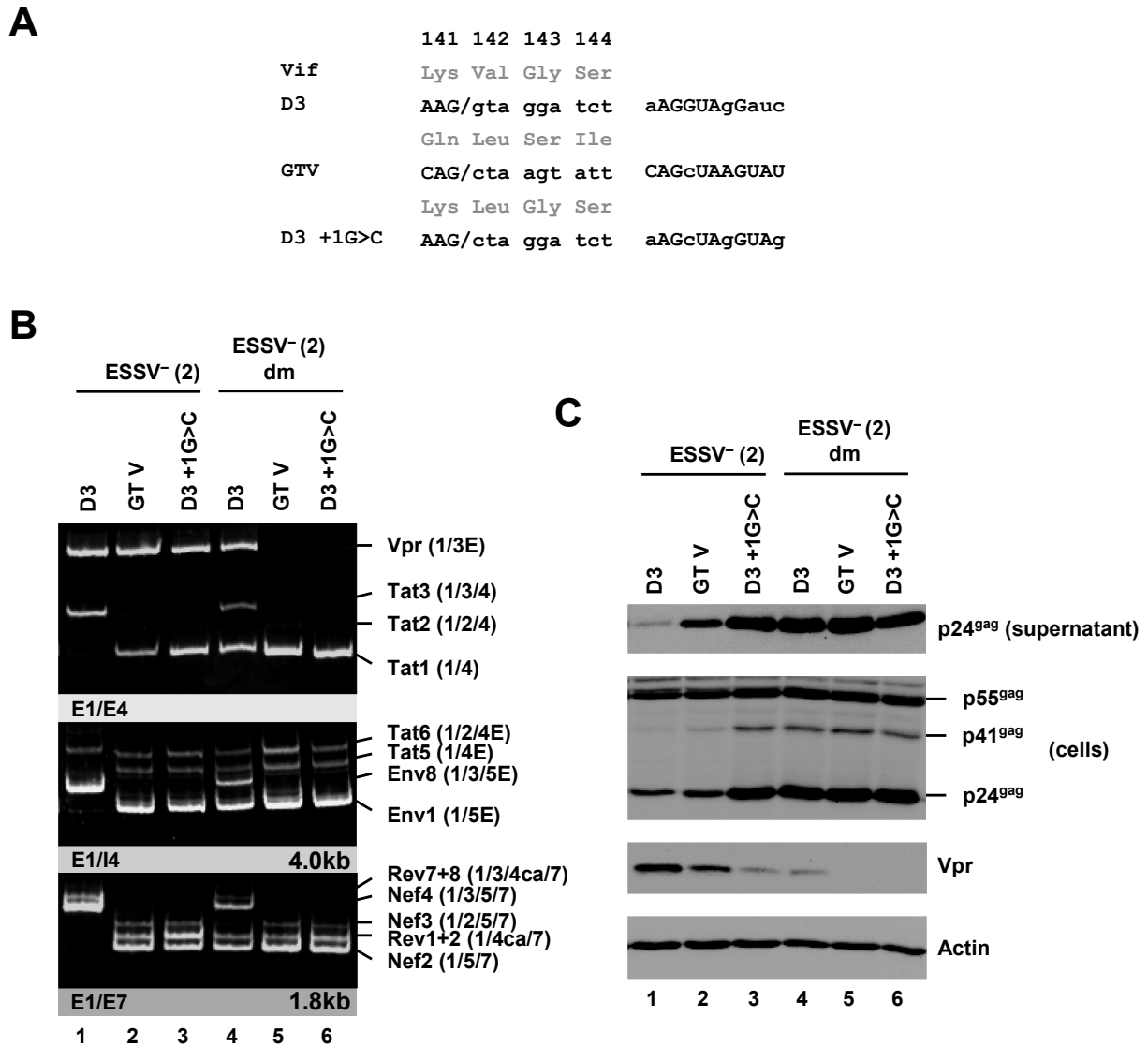


Fig.III-29: U1 snRNP binding to a splicing-incompetent 5'ss enhances *vpr*-mRNA expression.

(A) 5'ss D3 was replaced by a splicing incompetent sequence, which perfectly matches the free 5'-end of the cellular U1 snRNA except for position +1 (GTV). As a control, 5'ss D3 was disabled to splice due to a G-to-C mutation at position +1 without increasing its complementarity to the U1 snRNA (D3 +1G>C). Complementarity patterns are shown at the right. Matches to the U1 snRNA are indicated by upper case letters, residues not complementary are given in lower case letters. **(B)** 2.5×10^5 HEK 293T cells were transiently transfected with 1 μ g of each of the proviral constructs and analyzed by semi-quantitative RT-PCR. RT-PCR products were resolved by PAGE followed by ethidium bromide staining. Mutants are depicted at the top. Main viral mRNAs are indicated on the right and the left of the panels. **(C)** Cellular lysates and viral supernatants were obtained from transfected HEK 293T cells and loaded onto 12% SDS polyacrylamide gels. After transfer to nitrocellulose membranes, viral proteins were determined using specific antibodies for p24^{gag} and Vpr. To ensure loading of equal protein amounts, membrane was also probed with an antibody against directed against cellular β -actin. [dm: double mutant; E: Extended exon; HEK: Human Embryonic Kidney; RT: Reverse Transcriptase].

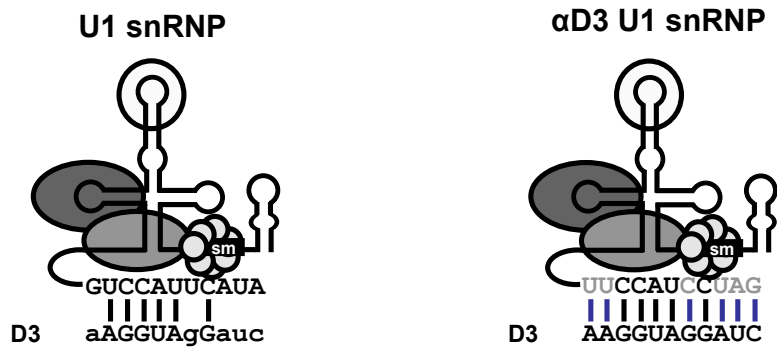
While 5'ss D3 was efficiently used in 293T cells transfected with an ESSV-mutant, neither "GTV" nor "D3 +1G>C" were spliced and thus, allowed accumulation of exon 3 containing mRNAs (Fig.III-29B, lanes 1-3). However, 5'ss D3 and the splicing-incompetent 5'ss-like sequence "GTV" each enhanced expression of *vpr*-mRNA, whereas "D3 +1G>C" as expected from its weaker complementarity to the U1 snRNA led to reduced splicing at the upstream 3'ss A2 (Fig.III-29B, lanes 1-3). This was in agreement with the hypothesis by which U1 snRNP binding alone suffices to augment cross-exon interactions and splicing at 3'ss A2. In ESSV/ ESE_{*vpr*}-double negative clones 5'ss D3 showed decreased levels of exon 3 inclusion and *vpr*-mRNA expression (Fig.III-29B, lane 4). However, substitution of 5'ss D3 with "GTV" or "D3 +1G>C", completely abolished splicing at 3'ss A2 (Fig.III-29B, lanes 5 and 6). This indicated that binding of the U1 snRNP to GTV is less efficient than to 5'ss D3 despite an overall higher complementarity.

Western blot analyses were consistent with the data obtained from the RT-PCR analysis and demonstrated that Vpr expression could be mostly maintained by "GTV", while it was strongly reduced by "D3 +1G>C" (Fig.III-29C, lanes 2 and 3). Moreover, it was observed that inactivation of splicing at the D3 could rescue virus-particle production of ESSV-negative provirus defective in replication (Fig.III-29C, lanes 1-3). These findings recapitulated earlier studies showing that the two functions of U1 snRNP binding to the 5'ss can be dissected; the first one, in formation of exon definition complexes and the second in assembly of a pre-spliceosome across a downstream intron (11, 215, 337).

III.2.3.3 A modified U1 snRNA fully complementary to 5'ss D3 strongly activates exon 3 inclusion and *vpr*-mRNA expression

To extend these analyses, a 5'-end mutated U1 snRNA was generated perfectly matching 5'ss D3 (Fig.III-30A). 293T cells were transiently cotransfected with the vector expressing the modified U1 snRNA and infectious plasmids either containing wildtype or mutant exon 3 sequences. RT-PCR analysis revealed a dramatic shift towards *vpr*- and/or exon 3 spliced mRNAs upon coexpression of the mutated U1 snRNA regardless of the tested ESSV/ESE_{*vpr*} mutations (Fig.III-30B, cf. lanes 1-8 and lanes 9-16), indicating that the coexpressed U1 snRNA seemed to assemble correctly into mature snRNPs. However, in absence of a functional ESE_{*vpr*}, U1 coexpression predominantly activated *vpr*-mRNA splicing, while only a minor influence on exon 3 inclusion could be found (Fig.III-30B, cf. lanes 1-4 and lanes 9-12), indicating that either the ESSV negatively acts on steps of spliceosome assembly later than early 5'ss recognition or the mutated ESE_{*vpr*} may also function later in the splicing reaction. Noteworthy, absence of both, ESSV and ESE_{*vpr*}, allowed efficient exon 3 inclusion into the viral mRNAs upon coexpression of the modified U1 snRNA (Fig.III-30B, lanes 5-8 and lanes 13-16), rather arguing for a bidirectional activity of the ESSV that represses downstream splicing after initial 5'ss recognition that is counteracted by an active ESE_{*vpr*}. Different mechanisms underlying the observed splicing phenotypes will be extensively discussed in chapter IV. Western blot analysis of the p24 levels within the supernatant suggested that the coexpression of the U1 snRNA could induce excessive exon 3 splicing even in the presence of the ESSV, thereby dramatically reducing virus particle production (Fig.III-30B, cf. lanes 1-2 and lanes 9-10). However, the coexpressed U1 snRNA relied on the presence of the ESE_{*vpr*} to elicit excessive activation of the exon 3 splice sites and a severely reduced viral replication (Fig.III-30B, cf. lanes 3-4 and lanes 11-12). When the ESSV was disrupted, U1 coexpression could efficiently inhibit viral replication independent on the ESE_{*vpr*}'s activity (Fig.III-30B, cf. lanes 5-8 and lanes 13-16). These findings emphasized that the complementarity between the 5'ss D3 and the U1 snRNA is not the sole determinant for exon 3 inclusion.

A



B

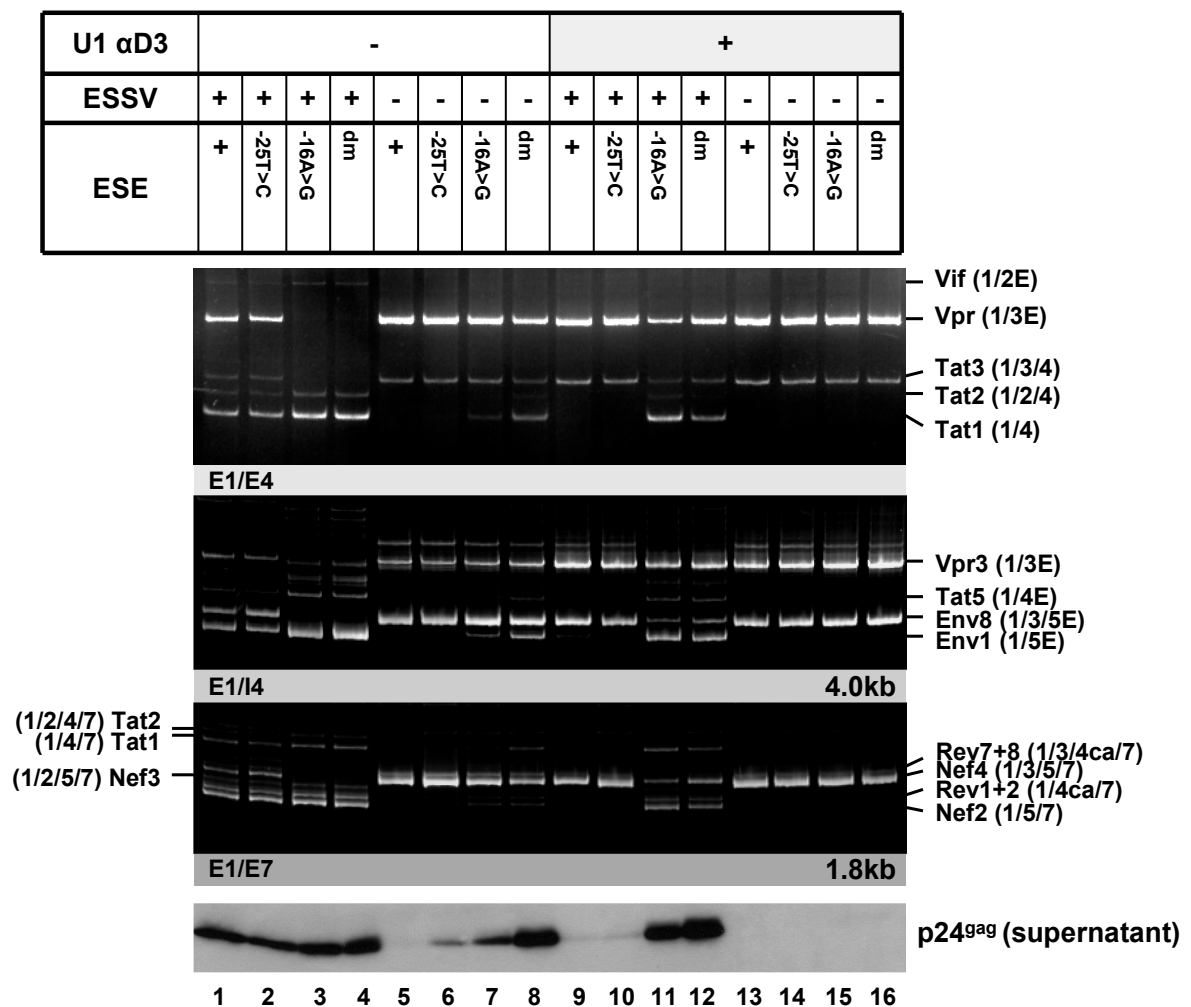


Fig.III-30: Coexpression of a modified U1 snRNP with full complementarity to 5'ss D3 induces HIV-1 exon 3 splicing and *vpr*-mRNA expression. [continued on next page]

By contrast, splicing efficiency is determined by an rewired interplay between 5'ss D3, ESSV and ESE_{vpr}. Moreover, it was demonstrated that exon 3 splicing and *vpr*-mRNA expression can be strongly induced by optimizing the binding efficiency of the U1 snRNP to the 5'ss D3. Detection of Vif and Vpr proteins within pNL4-3-transfected cells revealed that while Vpr expression was induced following coexpression of the mutated U1 snRNA, Vif was less abundant (see Appendix, Fig.VII-1.8, cf. lanes 2 and 3). Taken together, U1 coexpression studies reiterated the importance of U1 snRNP binding to the 5'ss D3 for use of upstream 3'ss A2 and *vpr*-mRNA splicing. In addition, they revealed that strengthening the interactions between U1 and the 5'ss D3 can override the dominant-negative activity of the ESSV in the context of a functional ESE_{vpr} and that the requirement of ESE_{vpr} for *vpr*-mRNA expression can be bypassed by an increased complementarity between 5'ss D3 and the U1 snRNA.

Fig.III-30: continued

(A) Schematic drawing of a 5'-end modified U1 snRNA perfectly matching 5'ss D3 sequence. Mutated nucleotides are indicated by grey capital letters. Additional basepairing interactions between 5'ss D3 and the optimised 5'-end of the U1 snRNA are indicated by vertical red lines.

(B) 2.5×10^5 HEK 293T cells were transiently cotransfected with 1 μ g of each, infectious plasmid and U1 expression plasmid. Total-RNA was isolated and subjected to RT-PCR analyses. PCR products were resolved by PAGE and stained with ethidium bromide. RT-PCR samples are given at the top. Main viral mRNA species are indicated at the left and right of the panels. Viral supernatants were collected as well and analyzed for viral p24^{gag} concentrations by immunoblotting (*lower panel*).

III.2.3.4 Modified U1 snRNAs with higher complementary to the splicing-incompetent 5'ss D3 +1 G>C efficiently rescues *vpr*-mRNA expression

To further substantiate these results, the 5'-end modified U1 snRNA was coexpressed either with D3 or mutant D3 +1G>C containing proviral plasmids and the outcome for splicing was determined by RT-PCR analysis (Fig.III-31A). As expected, *vpr*-mRNA splicing could also be activated when 5'ss D3 was disabled by the +1G>C mutation (Fig.III-31B). In the case of "D3 +1G>C" the extent of *vpr*-mRNA splicing was more dependent on the presence of a functional ESE_{*vpr*}, probably because of the mismatch to the mutated U1 snRNA at position +1 (Fig.III-31B, lanes 5-8 and 13-16). Interestingly, excessive exon 3 splicing seemed to be incompatible with recognition of the upstream exon 2, an observation already made earlier (Fig.III-31B, e.g. lanes 1 and 2).

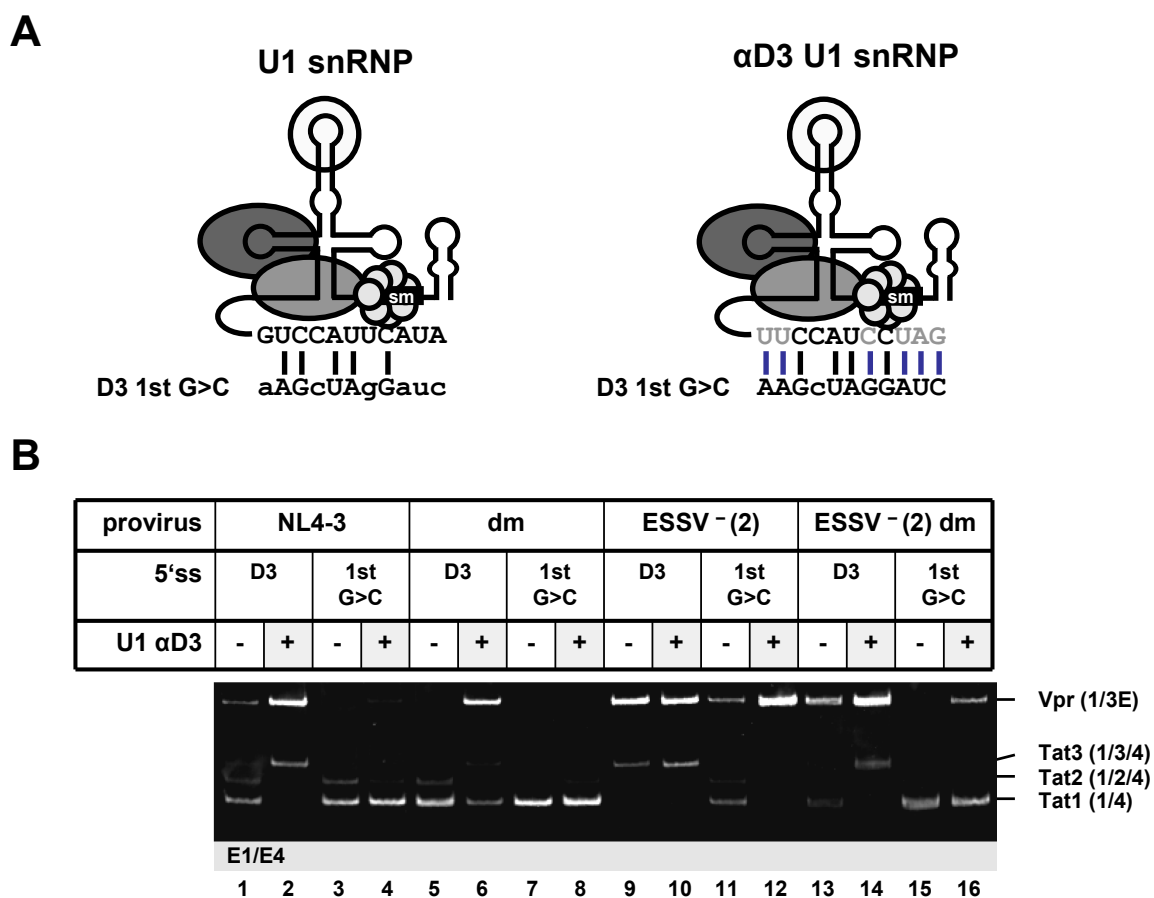


Fig.III-31: Binding of U1 snRNP to 5'ss D3 positively acts on use of the upstream 3'ss A2.

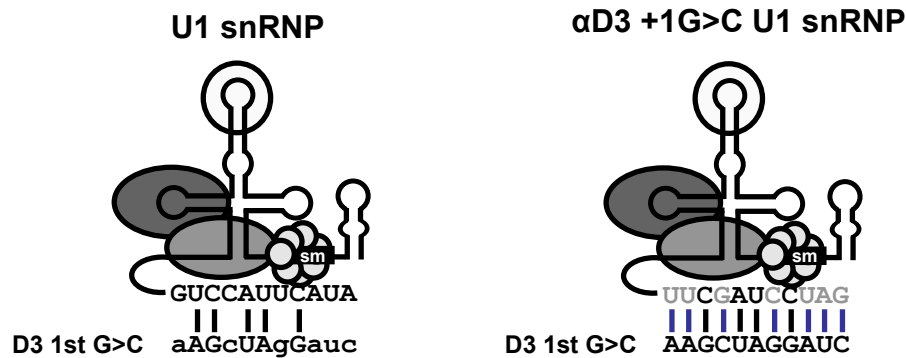
[continued on next page]

Fig.III-31: continued

(A) Sketch of the complementarity between the α D3 U1 snRNA and 5'ss D3 +1 G>C. Red vertical lines represent additional basepairing interactions exclusively formed between 5'ss D3 +1G>C and the mutated U1 snRNA. **(B)** Semi-quantitative RT-PCR analysis of total-RNA isolated from transfected HEK 293T cells. Samples are given at the top of the panel. Viral mRNAs are indicated at the right of the gel.

This suggested a delicate balance between splicing of the two leader exons whose perturbation results in an apparent mutually exclusive use of only one of both. To fortify these results a second U1 snRNA was designed, fully matching the splicing-incompetent 5'ss "D3 +1G>C" target sequence (Fig.III-32A). The U1 snRNA was coexpressed together with infectious plasmids within 293T cells and subsequently, changes in the splicing pattern and viral release were determined by RT-PCR and immunoblotting (Fig.III-32B). Coexpression of the mutated U1 snRNA shifted the splicing pattern for all exon 3 mutants exclusively to *vpr*-mRNA expression (Fig.III-32B, cf. lanes 1-4 and 5-8), obviously to a greater extent than the U1 snRNA containing only a single mismatch at position +1 (see Fig.III-32B). Furthermore, excessive *vpr*-mRNA production was accompanied by a severe defect to produce virus particles (Fig.III-32B, cf. lanes 1-4 and 5-8). These results solidified the observation that regardless of whether the U1 snRNP may or may not be used in the splicing reaction, binding to the 5'ss suffices to establish productive interactions across the upstream exon and augment 3'ss A2 use.

A



B

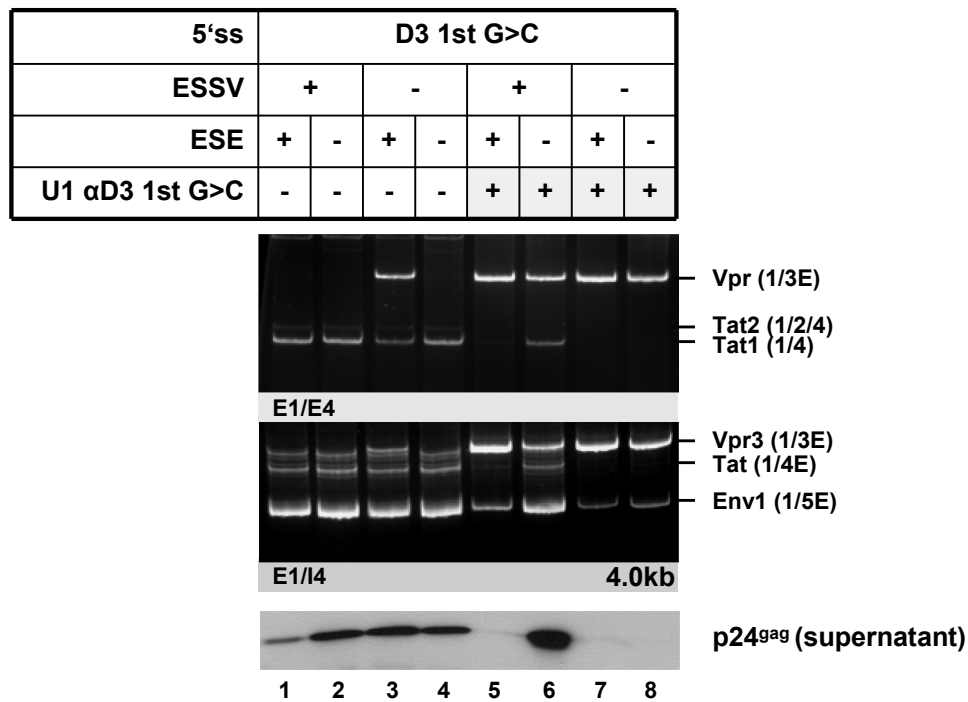


Fig.III-32: Binding of U1 snRNP to the splicing-incompetent 5'ss D3 +1G>C augments use of the upstream 3'ss A2.

(A) Diagram of a U1 snRNA optimized to entirely match 5'ss D3 +1G>C (αD3 +1G>C U1 snRNA).

(B) Semi-quantitative RT-PCR analyses of transfected HEK 293T cells with exon 3 mutants and the U1 expression vector. Samples are indicated at the top. Primer pairs used for amplification are given below the gels. Main HIV-1 mRNA species are indicated at the right. Viral supernatants were also harvested and analyzed by immunoblotting with a primary antibody specifically detecting structural p24^{gag}.

However, to detect the impact of the “αD3 U1 snRNA” on Vpr and Vif protein levels, western blot analyses were performed (Fig.III-33). Immunoblot analyses of the viral protein expression levels showed that Vpr expression could be up- or downmodulated by coexpression of the modified U1 snRNA (Fig.III-33). In absence of the ESSV, relatively high levels of Vpr protein could be detected for 5'ss D3. Coexpression of the 5'-end modified U1 snRNA increased the inherently efficient splicing at 5'ss D3 even more, presumably at the expense of incompletely spliced *vpr*-mRNAs, leading to a reduction in the levels of Vpr protein detected (Fig.III-33, lanes 1 and 2). By contrast, Vpr expression of the “D3 +1G>C” mutant was rarely detectable until the 5'-end modified U1 snRNA was coexpressed (Fig.III-33, lanes 3 and 4). However, despite an entire shift towards *vpr*-mRNA expression in the RT-PCR analysis (Fig.III-32B, lanes 10 and 12), less Vpr protein was detected than for 5'ss D3 (Fig.III-33, lanes 2 and 4). The ESSV-mutant containing 5'ss D3 could efficiently splice exon 3 into the viral-mRNA species and therefore expressed Tat and Rev protein from the respective exon 3-containing isoforms. The dramatic shift from *tat*-mRNAs towards *vpr*-mRNAs observed for non-functional 5'ss “D3 +1G>C” in the RT-PCR analysis was associated with the failure to direct the higher U1 snRNP binding activity into splicing of exon 3 into viral *tat*- and *rev*-mRNAs. This might lead to reduced Tat and Rev protein levels within the cells, impairing transcription and export of intron-containing viral mRNAs. This might also explain why despite excessive *vpr*-mRNA splicing, less Vpr protein was detected within the cells. However, when the ESE_{*vpr*} was additionally mutated in the context of the functional 5'ss, decreased Vpr expression could be rescued by coexpression of the modified U1 snRNA (Fig.III-33, lanes 5 and 6), indicating that targeting a U1 snRNA with maximal complementarity to D3 compensated for the absence of ESE_{*vpr*} activity. In contrast, Vpr expression of “D3 +1G>C” remained below detectable levels irrespective of coexpression of the U1 snRNA expression plasmid and despite an increase of *vpr*-mRNA confirmed by RT-PCR (Fig.III-33, lanes 7 and 8), reiterating the observation made for the ESSV-mutant. Since neither p24 nor Vif, both which are expressed from intron-containing viral mRNAs, could be detected following inactivation of the ESSV, this again indicated a defect in the production of *rev*-mRNAs (Fig.III-33, lanes 5-6 and 7-8). Noteworthy, intracellular levels of Vpr seemed to be inversely correlated with the amount of Vif protein that could be detected by immunoblotting (Fig.III-33, e.g. cf. lanes 1 and 5). Accordingly, excessive splicing at exon 3 and the

Results

corresponding high expression of Vpr excluded not only proper expression of unspliced mRNA necessary to produce viral structural proteins (p24^{gag}) (329), but also of Vif, again indicating a regulated balance between leader exon 2 and 3 splicing. Experimental results reinforced the notion that exonic signals (ESSV, ESE_{vpr}) and the intrinsic strength of 5'ss D3 act in concert to allow balanced HIV-1 exon 3 splicing, efficient enough to provide appropriate expression levels of *vpr*-mRNA, but not too efficient so that unspliced and *vif*-mRNAs can accumulate. Furthermore, it was demonstrated that U1 snRNP binding to 5'ss D3 directs recognition of 3'ss A2 by the splicing machinery, but that downstream intron-definition between 5'ss D3 and 3'ss A3 or 5'ss D3 and 3'ss A5 (or A4cab), respectively, has to occur at least with a lower efficiency than upstream definition of exon 3 in order to allow formation of intron-containing *vpr*-mRNAs.

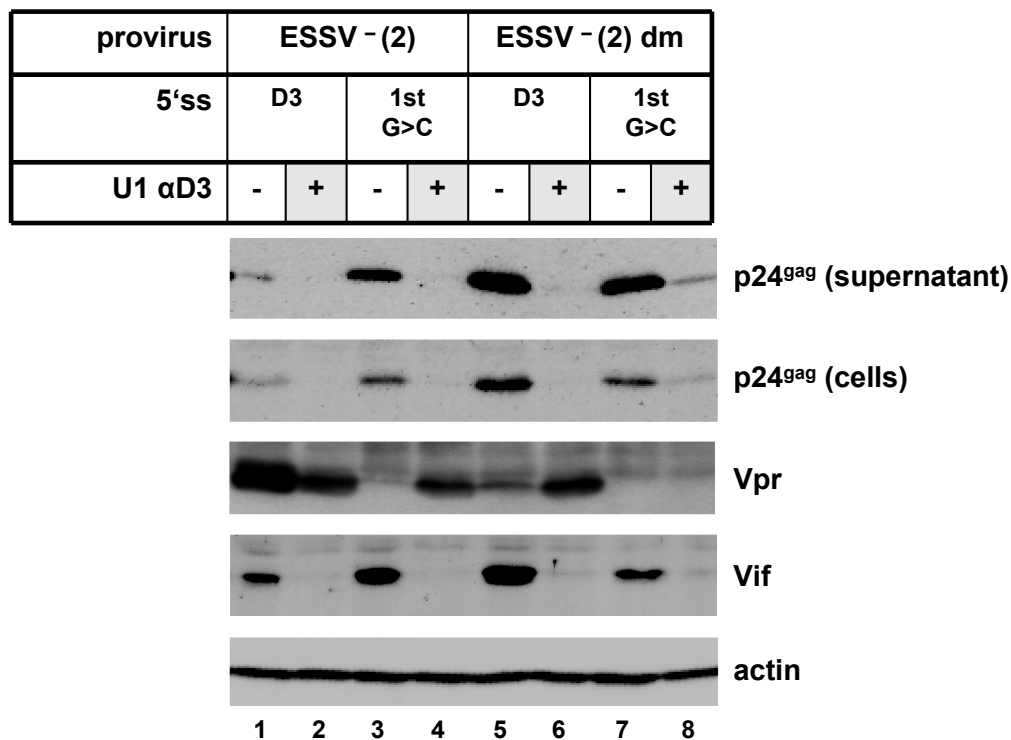


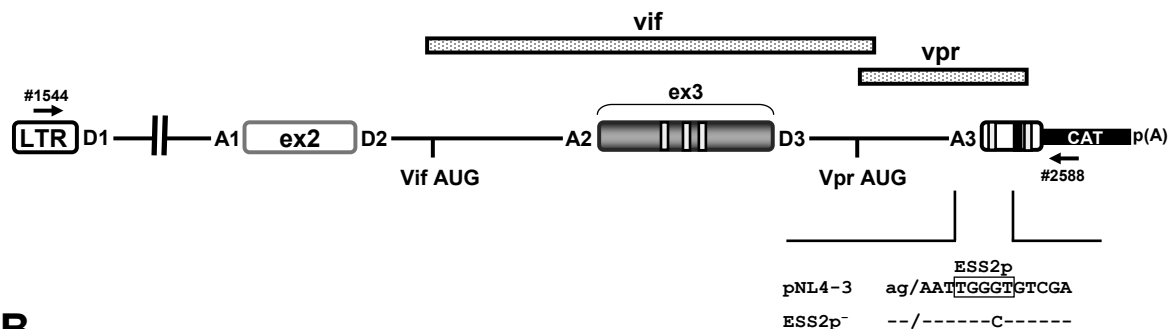
Fig.III-33: Vpr protein expression depends on binding of the U1 snRNP to the 5'ss.

Western blot analyses of viral p24^{gag}, Vpr and Vif after transfection of HEK 293T cells with different proviral exon 3 mutants and the αD3 U1 snRNA expression plasmid. Cellular lysates and viral supernatants were resolved in 12% SDS PAGE, transferred to nitrocellulose membranes and incubated with appropriate antibodies.

III.2.4 Increased activation of 3'ss A3 reduces *vpr*-mRNA expression in the context of the HIV-1 based 4-exon minigene

It is here hypothesized that splicing out the *vpr*-intron is predominantly regulated at the downstream 3'ss and the efficiency by which these are used for splice site pairing with 5'ss D3. Since the recruitment of U1 snRNP, controlled by the novel upstream ESE_{*vpr*} is absolutely pivotal for expression of *vpr*-mRNAs and studies by others could show that this binding is not negatively influenced by the ESSV activity (131), it is likely that downstream intron definition occurs with lower efficiency relative to early 5'ss D3 recognition. Therefore, the intrinsic strength of the *tat*-mRNA specific 3'ss A3 was increased to determine whether its strength has an influence on the production of viral *vpr*-mRNA. Immediately downstream of 3'ss A3 an hnRNPF/H-dependent exonic splicing silencer, termed ESS2p, is localized, which was shown to downregulate splicing (238) (Fig.III-34A). Therefore, ESS2p was inactivated by site-directed mutagenesis to increase the frequency of 3'ss A3 use. This was then examined within the context of different exon 3 mutants in the HIV-1 based 4-exon minigene (see III.2.1). HeLa cells were transiently transfected with each of the reporters and RNA was isolated from the cells for analysis by RT-PCR (Fig.III-34B). Proper expression of intron-containing *vpr*-mRNAs relies on binding of the viral protein Rev to the Rev-responsive element (RRE) embedded within the *env*-intron. Since neither Rev has been coexpressed nor the reporter mRNAs contained a Rev-responsive element (RRE), it was expected to generally detect lower levels of *vpr*-mRNA due to its retention within the nucleus and subsequent withdrawal by degradation. Indeed, *vpr*-mRNAs were underrepresented relative to the Tat-specific mRNAs (Fig.III-34B). Furthermore, at a first glance it seemed to be conflicting with what was found before in transfection experiments with full-length proviral DNAs, since the intrinsic 5'ss D3 strength was negatively correlated with *vpr*-mRNA accumulation and only detectable for the weakest 5'ss D3 down mutation within the context of the ESSV mutants (Fig.III-34B, cf. lanes 3, 7 and 11). However, this could be explained by absence of competitive sequences downstream of 3'ss A3 such as 3'ss A5 that normally might be chosen over 3'ss A3.

A



B

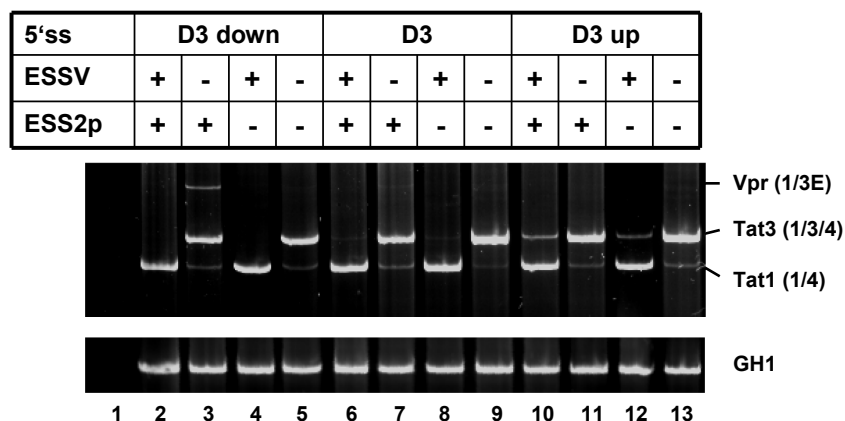


Fig.III-34: *Vpr*-mRNA formation requires inefficient spliceosome formation across the downstream intron.

(A) Schematic drawing of the pNL4-3 derived LTR ex2 ex3 minigene. Wildtype and mutated ESS2p sequences are given below. (B) 2.5×10^5 HeLa-T4⁺ cells were transfected with 1 μ g of each of the constructs, 0,2 μ g of SVctat and 1 μ g of pXGH5 to control transfection efficiency. Total-RNA was isolated 30h post transfection and assayed by RT-PCR using primer pairs specific for alternatively spliced viral mRNAs (#1544/#2588) or constitutively spliced GH1 mRNA (#1224/#1225), respectively. PCR products were loaded onto 8% non-denaturing polyacrylamide gels, electrophoresed and visualized by ethidium bromide staining. Constructs are denoted at the top of the gel. Spliced viral mRNAs are indicated at the right. [RT: Reverse Transcriptase; GH1: human Growth Hormone 1]

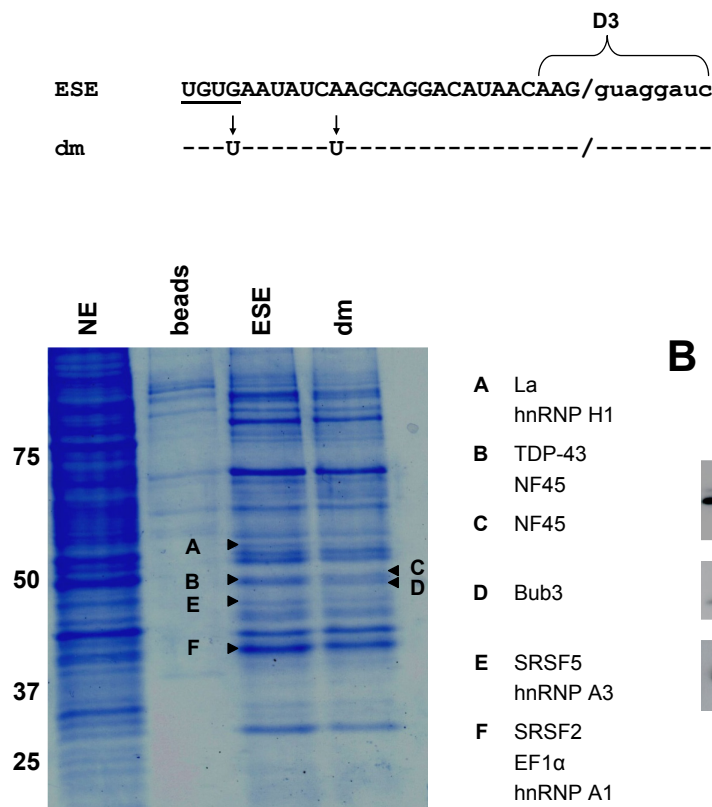
In this case, *vpr*-intron retention would rely on inefficient splice site pairing between 5'ss D3 and 3'ss A3 on the one hand, but also an occasionally inhibited intron definition between 3'ss D3 and one of the downstream 3'ss. Noteworthy, *vpr*-mRNA splicing is decreased in ESS2p mutants (Fig.III-34B, cf. lanes 3 and 5), further indicating that a weak downstream 3'ss A3 recognition is a prerequisite for *vpr*-intron retention. Although it awaits further studies in the context of proviruses, *vpr*-mRNA splicing may necessitate repression of splicing at the downstream 3'ss A3 and 3'ss A4cab.

III.2.5 TAR DNA binding protein-43 (TDP-43) is not involved in positive splicing regulation of HIV-1 exon 3

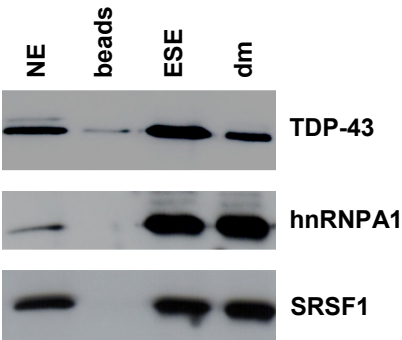
To figure out, which cellular factors bind to the ESE_{vpr} within HIV-1 exon 3, RNA affinity purification experiments were performed, incubating short *in vitro* transcribed RNA substrates either representing the wild-type or double-mutated ESE_{vpr} (-24G>T; -17A>T) sequence in HeLa cell nuclear extract (Fig.III-35A). Using mass spectrometric analysis, TAR-DNA binding protein-43, termed TDP-43, was found to be exclusively associated with the wildtype ESE sequence. This was confirmed by western blotting with a specific TDP-43 antibody (Fig.III-35B). While the levels of hnRNP A1 and SRSF1 were not altered by the mutation, TDP-43 showed a significantly reduced binding to the mutated ESE sequence. TDP-43 is a highly conserved RNA-binding protein implicated to play a role in alternative splicing (56). TDP-43 binds to “*ug*”-repeats (544), which supports specific binding of TDP-43 to exon 3, since “-24G>T” single mutation affects an “*ugug*”-motif within the ESE_{vpr} sequence. To elucidate whether TDP-43 is functionally involved in the inclusion of exon 3, an siRNA mediated knock-down of TDP-43 was performed to determine the effects of TDP-43 downmodulation on viral HIV-1 exon 3 splicing regulation. Despite the efficient reduction of the TDP-43 expression levels within HeLa cells (Fig.III-35C), no changes within the relative abundance of exon 3-containing mRNAs could be detected (Fig.III-35D). Although it cannot formally excluded that residual TDP-43 expression after siRNA treatment might have been supplied sufficient amounts of protein for ESE_{vpr} binding, results argued against a role of TDP-43 as exon 3 inclusion factor. In addition, results were consistent with the comparable minor outcome of the “-25 T>C” mutation within the “*ugug*” TDP-43 binding motif for exon 3 splicing. Furthermore, overexpression experiments using a TDP-43 expressing plasmid supported the conclusion that TDP-43 does not bind to the exonic splicing enhancer within exon 3 to promote splicing (*data not shown*).

Results

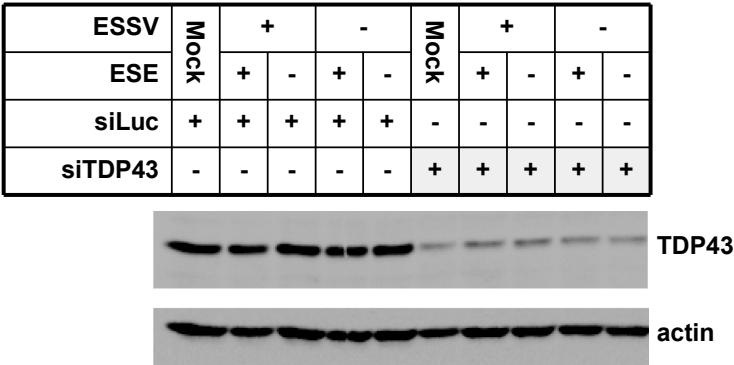
A



B



C



D

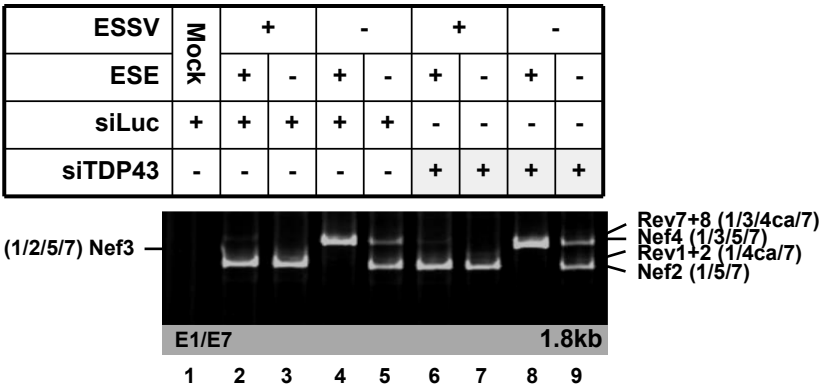


Fig.III-35: TDP-43 is not involved in HIV-1 exon 3 splicing regulation. [continued on next page]

Fig.III-35: continued

(A) RNA pull-down analysis of *in vitro* transcripts of either non-mutated or mutated ESE_{vpr} sequence with HeLa cell nuclear extract were performed to identify host cell proteins, which bind to the ESE_{vpr}. *In vitro* RNA substrates used are shown. NL4-3 sequence is shown in uppercase letters, mutated residues are indicated below. dm: double mutant (-24G>T; -17A>T). The precipitated proteins were resolved by denaturing 10% SDS polyacrylamide gel electrophoresis and stained with Coomassie Brilliant Blue. Bands were eluted and analyzed by mass spectrometry (BMFZ). The following proteins were identified: A) LA, hnRNP H1; B) NF45, TDP-43; C) NF45; D) BUB3; E) SRSF5, hnRNP A3; F) SRSF2, EF1 α , hnRNP A1. Bands are indicated. **(B)** Immunoblot analysis with antibodies specific for TDP-43 or hnRNPA1 and SRSF1 could confirm that TDP-43 levels were significantly reduced in the context of the double mutant. **(C)** To determine the exact role of TDP-43 for exon 3 splicing, HeLa cells were transfected with a small interfering RNA (siRNA) against TDP-43 and a negative siRNA control (siLuc). Cells were treated two times with siRNAs prior to transfection with proviral constructs. For gene silencing of human TDP-43 the following siRNA duplex was used: GCAAAGCCAAGAUGAGCCUdTdT. The siRNA against TDP43 was kindly provided by Dr. Emanuele Buratti. Equal amounts of proteins were loaded onto a 10% denaturing SDS polyacrylamide gel, transferred to a nitrocellulose membrane and incubated with a specific antibody for TDP-43 to confirm protein knockdown. **(D)** RT-PCR analysis of siRNA-treated HeLa cells transfected with pNL4-3 and the various exon 3 mutants. RNA from the transfected cells was reverse transcribed and cDNA amplified with primer pair E1/E7 specific for 1.8kb size viral mRNA class. Resultant PCR products were resolved on 8% polyacrylamide gels and stained by ethidium bromide. Samples are given at the top of the gel. Main RNA species are indicated at the left and the right. [dm: double mutant; NE: HeLa Nuclear Extract; RT: Reverse Transcriptase]

IV. Discussion

IV.1 Position-dependent splicing activation and repression by SR and hnRNP proteins

Splicing regulatory elements (SREs) have been accepted to be major regulators of alternative splicing. The classical view usually defines SR proteins as global activators and hnRNP proteins as global repressors of splicing. However, this is difficult to reconcile with a growing number of examples in which splicing factors from both protein families show the opposite splicing phenotype; in other words SR proteins repress splicing (58, 120, 213, 229, 256, 426, 483), while hnRNP proteins enhance it [e.g. (204, 316, 569)]. An immanent property of splicing regulatory proteins to guide splicing fate into both directions was further substantiated by genome-wide studies, which revealed dual outcomes of the hnRNP and hnRNP-like proteins Nova (550), Fox1/2 (614), TDP-43 (544) and PTB (316) for splicing. In all these cases conversion from enhancer to repressor (or the other way around) was linked to the position of an SRE relative to the regulated exon. However, comparative, systematic analyses of splicing proteins and their position-dependent effects on splice site use were not available and furthermore, underlying mechanistic principles of the dual splicing activities remained a conundrum. Results obtained here could demonstrate that SR proteins and hnRNP or hnRNP-like proteins exhibit a mirror-inverted splicing phenotype regarding their position relative to the test 5'ss. While SR proteins in exonic positions promoted 5'ss use, their relocation into the downstream intron coincided with repression of splicing. The opposite linkage of splicing activity and position was observed for hnRNP and hnRNP-like proteins. Exonic hnRNP or hnRNP-like protein positions were associated with splicing repression. By contrast, 5'ss use was triggered by placement of the respective binding sites into to the downstream intron. This characteristic of position-dependent activation or repression was found for almost all common splicing factors.

RNA pulldown assays revealed that U1 snRNP recruitment to the 5'ss was enhanced by splicing proteins in their activating positions, suggesting a unifying capability of SR and hnRNP or hnRNP-like proteins to support early splice site recognition. Unexpectedly, splicing repression was not conferred by perturbation of initial U1 snRNP binding to the 5'ss. By contrast, SR proteins enhanced U1 snRNP

recruitment regardless of their relative location to 5'ss. However, splicing complex formation analyses indicated an inability of the U1 snRNP recruited by intronic SR proteins to carry out later steps of spliceosome assembly (145). Even though hnRNP or hnRNP-like proteins enhanced U1 snRNP recruitment exclusively from an intronic position, relocation of their binding sites to the upstream exon at least appeared not to block 5'ss recognition. Taken as a whole, repression by SR proteins was not due to interference with early splice site recognition, but associated with recruitment of a splicing-incompetent, "dead-end" complex at the 5'ss. Since there are no splicing complex formation data available for exonic hnRNP protein positions at this time, it remains elusive whether repression by hnRNP and hnRNP-like proteins relies on a comparable mechanism.

IV.1.1 Almost all common splicing regulatory proteins can enhance U1 snRNP recruitment from defined positions

Using an enhancer-dependent splicing reporter, it was demonstrated that activation of the 5'ss essentially depends on the relative position of the splicing regulatory proteins. While SR proteins enhanced splicing from the exonic side of the 5'ss, hnRNP or hnRNP-like proteins needed strict positioning into the downstream intron (Fig.IV-1). This was demonstrated by both use of natural occurring SREs (Fig.III-4) and in MS2 tethering experiments (Fig.III-7). The latter to formally exclude that the sequence itself influences splicing by formation of local RNA secondary structures or is bound by nuclear proteins other than anticipated.

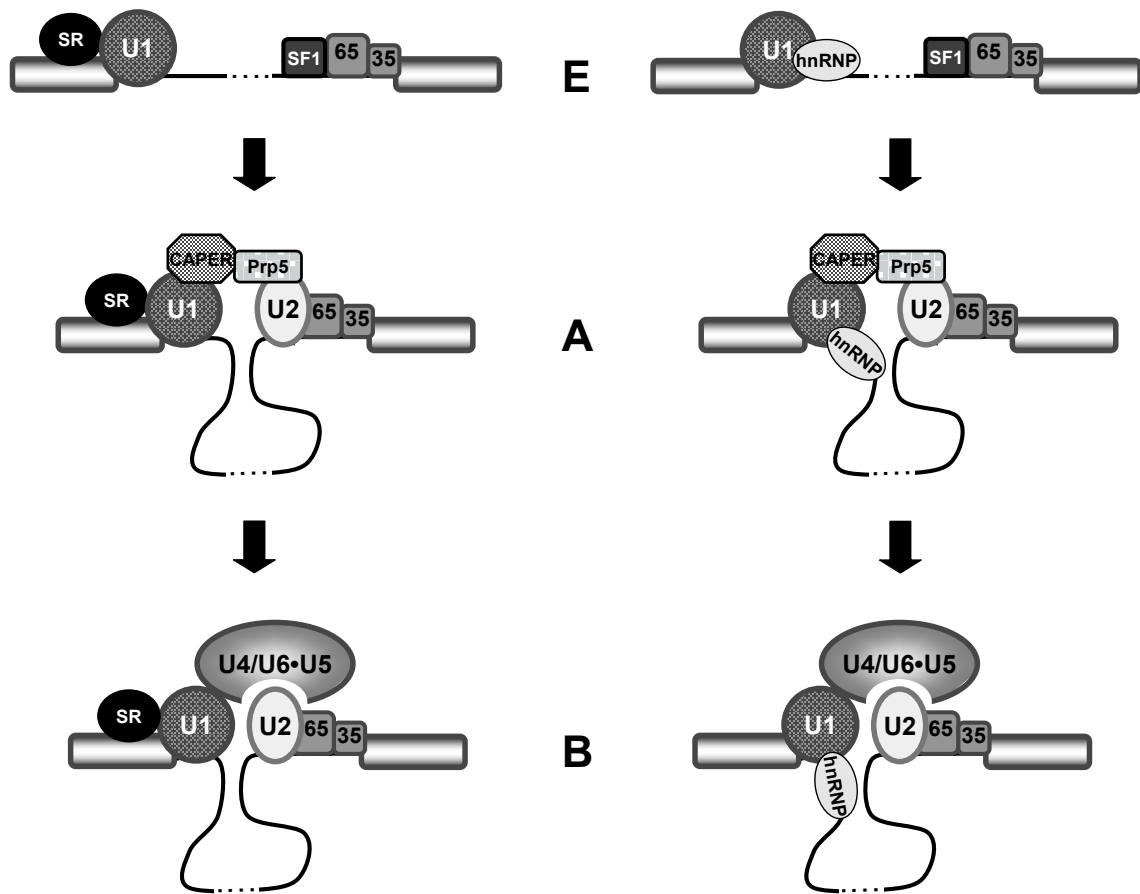


Fig.IV-1: Position-dependent 5'ss activation by SR and hnRNP proteins

Almost all tested splicing factors had the capability to enhance splicing at the nearby 5'ss from an exonic (left) or intronic (right) side. While SR proteins needed strict positioning into the upstream exon, hnRNP or hnRNP-like proteins only could activate splicing from a downstream intronic location relative to the 5'ss. [U1: U1 snRNP; 65: U2AF65; 35: U2AF35; U2: U2 snRNP; E: E complex; A: A complex; B: B complex].

IV.1.1.1 U1 snRNP recruitment by SR proteins

The mirror-inverted splicing phenotype observed could result from distinct strategies by which SR proteins and hnRNP or hnRNP-like proteins recruit the U1 snRNP to the 5'ss. Currently two models are discussed how exonic bound SR proteins interact with the U1 snRNP to direct its binding to the 5'ss. SR proteins have a modular architecture consisting of one or two RNA recognition motifs (RRMs) in the N terminus and a C-terminal arginine-serine (RS) rich domain of variable size.

Phosphorylation of the RS domain was described to be critical for U1 snRNP recruitment to the 5'ss (91, 592). In the commonly accepted pathway SR proteins escort the U1 snRNP to the 5'ss by RS domain-mediated interactions with the U1 snRNP specific protein U1-70K (70, 276, 589). Herein, the RRM confers RNA binding specificity and is only necessary for positioning of the RS domain into the proximity of the 5'ss [Fig.IV-2 (a)]. However, the RS domain dependent mechanism was recently called into question by a study suggesting that U1 snRNP recruitment is primarily achieved by interactions between the RRM domains of SR proteins and U1-70K [Fig.IV-2 (b), (91)]. In line with this hypothesis protein-protein interactions mediated by RRM domains were described precedently [(249, 265, 487); reviewed in (101, 343)]. Other findings support the notion that the RRM of SR proteins may exert functions aside from RNA binding. First, it was documented that an RRM alone could complement *in vitro* splicing of some pre-mRNA substrates in SR protein deficient cytoplasmic HeLa S100 extracts (65, 91, 480, 623, 630). Second, a phosphorylated RS domain on its own failed to efficiently interact with U1-70K in GST pull-down assays (592). However, data are controversial according to the capability of an isolated RRM to contact U1-70K (91, 592). Moreover, SR proteins have also been described to antagonize binding of splicing silencer proteins (194, 397, 606, 624), which does not necessarily require their RS domain (624). This may account for the ability of isolated RRMs to complement *in vitro* splicing in some cases while fail to do in others. However, the results obtained here (Fig. III-7) and by others (72, 186, 187, 483) clearly demonstrate that RS-domains alone can function as potent activators of adjacent splice sites when tethered to an exonic position. This suggests that the RRM domains of SR proteins rather play an ancillary role for binding of the U1 snRNP to the 5'ss although it cannot be ruled out that the RS domains simply stabilizes the basepairing interactions between U1 snRNP and the 5'ss (481, 482) instead of directly interacting with U1-70K. Therefore, it awaits further systematic analyses to delineate the relative contributions of each SR protein domain to 5'ss activation. However, RNA pull-down experiments recapitulated the preceding findings that SR proteins are capable to enhance U1 snRNP recruitment to the 5'ss (Fig.III-12 B and D).

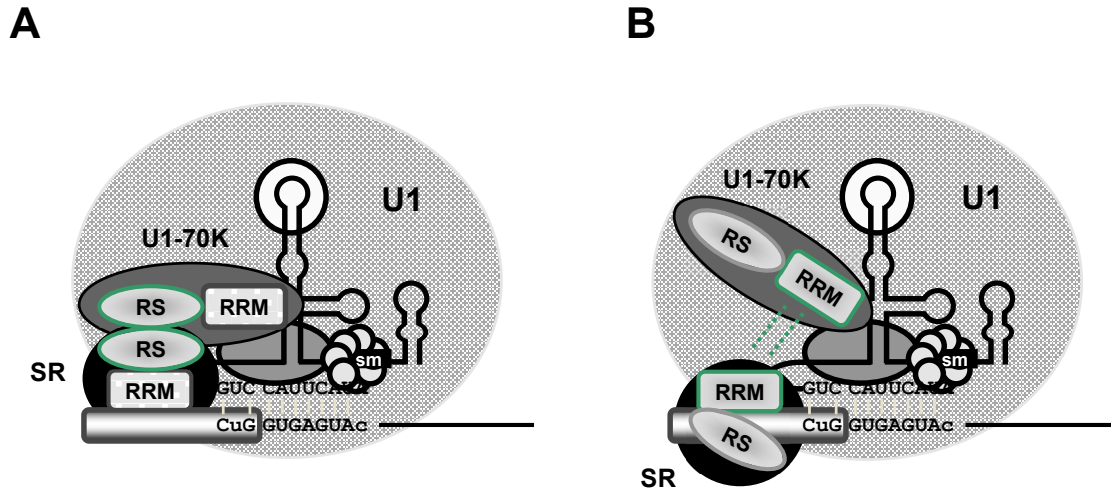


Fig.IV-2: Modes of SR protein-mediated recruitment of the U1 snRNP

In the long-standing model, SR proteins enhance recruitment of the U1 snRNP to the 5'ss by RS-domain mediated interactions with the U1-specific protein U1-70K **(A)** (70, 276, 589). Alternatively, 5'ss recognition is enhanced by RRM-mediated interactions between exonic bound SR proteins and the U1-specific protein U1-70K **(B)**, releasing the RS domains for other protein-protein interactions upon spliceosome assembly (91). Basepairing interactions between the U1 snRNA and the 5'ss and proposed protein:protein contacts are highlighted in green. [RS: arginine-serine-rich domain; RRM: RNA recognition motif; SRSF: SR splicing factor; U1: U1 snRNP].

IV.1.1.2 U1 snRNP recruitment by hnRNP and hnRNP-like proteins

The exact mechanisms by which hnRNP and hnRNP-like proteins recruit the U1 snRNP to an upstream 5'ss are even less understood. However, that almost all hnRNP or hnRNP-like proteins can activate splicing from a downstream intronic position suggests a commonality to directly promote U1 snRNP binding to the 5'ss (Fig.III-4 and III-7). The probably best described case is the recruitment of the U1 snRNP by the hnRNP-like protein TIA-1 bound to the downstream intron (157, 173). TIA-1 is a multidomain protein containing three RRMs of which RRMs 2 and 3 permit RNA binding to uridine (U)-rich sequences (123, 157, 304). However, RRM1 and a C-terminal glutamine (Q)-rich domain are required to direct the U1 snRNP to the upstream 5'ss (157, 173). Herein, the Q-rich domain directly interacts with the U1 snRNP specific protein U1-C [Fig.IV-3 (a)]. In agreement with this finding, U1-C was specifically detected in TIA-1 precipitates (Fig. III-18). However, the RRM1 by itself fails to interact with U1-C, but was shown to facilitate U1 snRNP recruitment even in

absence of the Q-rich domain (157). Therefore, it is conceivable that TIA-1 contacts the U1 snRNP by rewired interactions not exclusively with U1-C.

Although U1-C might be an essential target for TIA-1 (157, 441) and potentially Fox1/2 (394), it remains unclear whether interactions with U1-C are fundamental for 5'ss activation by other hnRNP and hnRNP-like proteins. With exception of TIA-1, hnRNP and hnRNP-like proteins do not contain Q-rich domains. Moreover, equivalent domains, which may functionally substitute for the Q-rich domain have yet not been identified. At least for hnRNP F an inability was shown to interact with U1-C in immunoprecipitation assays (Fig.III-18). However, intronic hnRNP F/H positions markedly enhanced recognition of the 5'ss (Fig.III-13), indicating that there may exist surrogate mechanisms to establish functional interactions with the U1 snRNP.

Almost all hnRNP and hnRNP-like proteins share the salient feature to contain multiple RRM domains. There seems to be a certain degree of functional redundancy among RRM domains with regard to their requirement for pre-mRNA binding (23, 130, 157, 478). This raises the question of whether free RRMs may or may not be available to exert functions in U1 snRNP recruitment to the 5'ss.

The polypyrimidine tract binding (PTB) protein (also hnRNP I) contains four quasi-RRM (qRRM) domains. As the name already implies, PTB preferentially binds to pyrimidine (Y)-rich RNA sequences. Recently it was reported that both qRRM1 and 2 of PTB are each capable to directly interact with the pyrimidine-rich internal stem loop IV of the U1 snRNA [Fig.IV-3 (b), (478)]. The authors suggested this might allow recruitment of the U1 snRNP to the 5'ss by intronic bound PTB [Fig.III-4 and III-7, (478)]. In contrast to stem loops I and II, which are bound by U1-70K and U1-A, respectively, stem loop IV is not occupied by any U1-specific protein within the U1 snRNP particle, leaving it theoretically accessible for interactions with hnRNP and hnRNP-like proteins (421, 519). For splicing activation, PTB might bind to a downstream intron using qRRMs 3 and 4, which preferentially bind to single-stranded RNA targets (100) and then bridge binding of the U1 snRNP to the adjacent 5'ss by qRRMs 1 and 2, which in turn favour binding to the structured stem loop IV of the U1 snRNA (478). Whether the RRMs of other hnRNP proteins directly contacts stem loop IV throughout U1 snRNP recruitment and this is a common mechanism of splicing activation is yet unclear. However, at least the RRM1 of TIA-1 was shown to support U1 recruitment in an U1-C independent manner, which might be due to direct binding of the U1 snRNA (157).

The positioning of high affinity binding sites for hnRNP F/H into the downstream intron of the 5'ss strongly increased the levels of U1 snRNP recruitment (Fig.III-13). HnRNP F/H proteins also contain multiple qRRM domains of which each individual one was described to specifically bind to guanine (G)-run RNA sequences (130). Interestingly, the U1 snRNA displays a putative hnRNP F/H binding sequence within stem loop IV and the qRRMs of hnRNP F/H proteins were also found to efficiently bind to G-run sequences, which are embedded within RNA secondary structures (69, 130, 235). Given these facts, it is plausible that intronic bound hnRNP F/H proteins may act as molecular adaptors between U1 snRNP and the 5'ss by RRM mediated interactions with the pre-mRNA and the U1 snRNA.

It is tempting to speculate that multiple RRM domains may afford hnRNP proteins to build ternary complexes with the pre-mRNA and the U1 snRNP simply by RNA interactions, which would provide an alternative mode of U1 snRNP recruitment in addition to protein interactions with U1-70K or U1-C.

However, Fox1/2 proteins only contain a single RRM and it still remains elusive whether they can interact with U1-C (394). The C-terminal domains of Fox1/2 proteins specifically interact with hnRNP H for splicing activation from positions immediately downstream of a regulated exon (349, 526). It is possible that Fox1/2 proteins interpose hnRNP H for U1 snRNP recruitment to get around their own confinement neither possessing multiple RRMs nor a Q-rich domain for interaction with U1-C [Fig.IV-3 (c)].

Among all common splicing factors tested for their position-dependent activation of a 5'ss, hnRNP A1 was the only protein, which could not enhance splicing regardless of its positioning to the upstream exon or downstream intron, respectively (Fig.III-4 and III-7). In contrast to other hnRNP proteins, there are no precedents, indicating that hnRNP A1 has the functional property to directly interact with spliceosomal components, which does not necessarily imply that they cannot promote splicing in a position-dependent manner. Splicing activation by hnRNP A/B proteins reported so far requires at least two binding sites, one at each terminal end of the intron. Self-interactions between the hnRNP A/B proteins then facilitate looping out of the internal intronic sequences, which alleviates end-to-end bridging (153, 345). However, there is no evidence for active recruitment of spliceosomal components by hnRNP A/B proteins to nearby splice sites as yet. HnRNP A/B proteins contain two copies of a canonical RRM in their N-terminus (138). It was shown that both RRM domains

cooperatively form a single binding unit and each of the individual RRM domains fail by itself to efficiently bind to the pre-mRNA (355). This indicates that as opposed to most of the other hnRNP proteins there is no free RRM available, which may serve to contact the U1 snRNA. Furthermore, stem loop IV does not contain any hnRNP A1 binding motif, which might be related to an inability to promote U1 snRNP recruitment to the 5'ss. However, since hnRNP A1 also interacts with hnRNP H (153), one would expect that even if it does not offer a free RRM, it might exploit hnRNP H as suggested for Fox1/2 proteins to direct the U1 snRNP to the 5'ss. However, multiple RRMs are commonly found in hnRNP and hnRNP-like proteins and this redundancy has a heretofore unknown function.

Given this striking structural similarities one can speculate that direct interactions between RRM domains and the U1 snRNA might provide a molecular basis behind the activation of 5'ss by hnRNP or hnRNP-like proteins. The astounding manner that almost all splicing regulatory proteins have the capability to enhance a 5'ss from defined positions came as a surprise and calls for a re-evaluation of the long-held classification of SR proteins and hnRNP proteins into enhancers and repressors of splicing, respectively. However, while SR proteins need the strict positioning into the upstream exons to activate splicing, hnRNP proteins require a position at the opposite side of the 5'ss. There is little to no understanding of the mechanistic principles establishing these opposite positional requirements for 5'ss activation. However, the 3D structure of the U1 snRNP might provide a possible explanation, since it reveals that stem loop IV and U1-C are placed at the opposite side of U1-70K (421, 519). In light of the proposed modes of U1 snRNP recruitment by SR and hnRNP proteins their position-dependent splicing phenotype may simply reflect the opposite spatial arrangement of their interaction partners within the U1 snRNP particle.

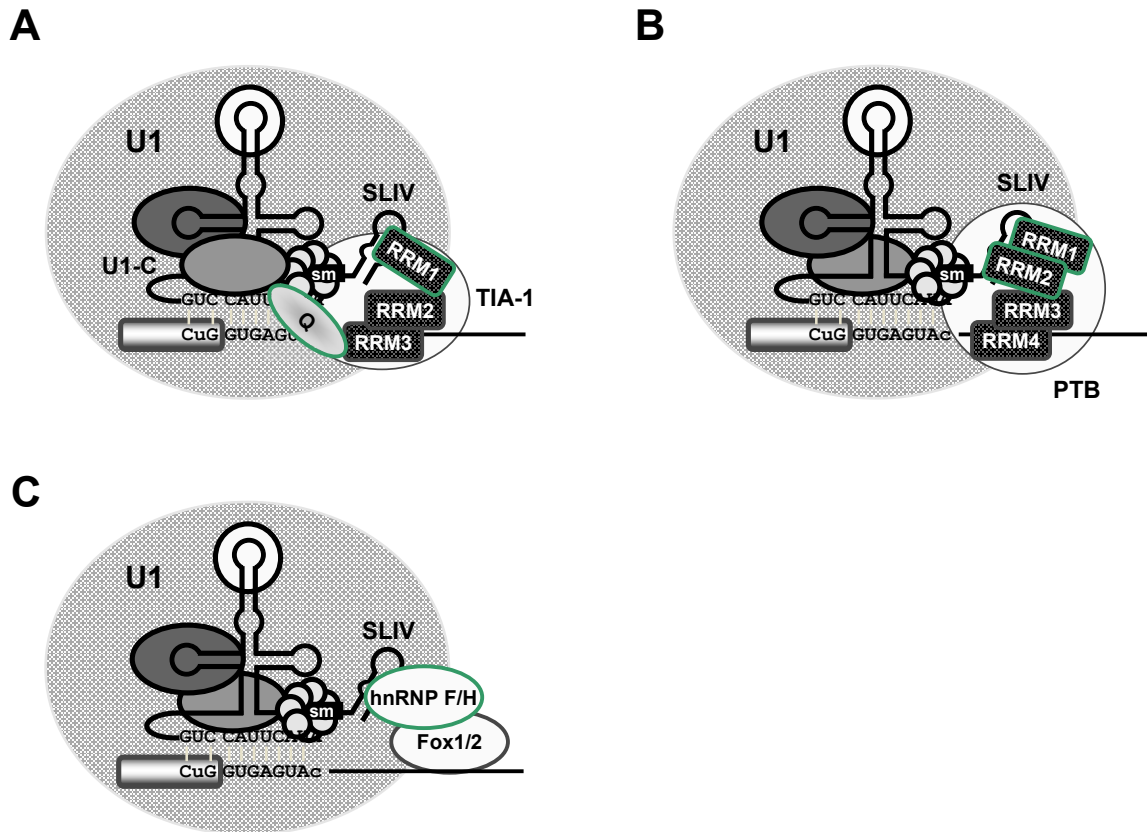


Fig.IV-3: Modes of hnRNP protein-mediated recruitment of the U1 snRNP

The splicing factor TIA-1 enhances binding of the U1 snRNP to the upstream 5'ss by interactions of its glutamine-(Q) rich domain with the U1-specific protein U1-C (**A**) (157). In addition, RRM1 of TIA-1 might directly interact with the U1 snRNA, thereby supporting U1 snRNP recruitment to the 5'ss. The hnRNP protein PTB directly interacts via RRM1 and RRM2 with the stem loop IV (SLIV) of the U1 snRNA (**B**) (478), which might stabilize the U1 snRNP to the upstream 5'ss. Fox1/2 proteins were proposed to interact with hnRNP H for 5'ss activation (526). Herein, hnRNP H proteins might act as bridging factors between the U1 snRNP and intronic bound Fox1/2 proteins by directly contacting SLIV via their RRM domains (**C**). Interactions are highlighted in green. [U1: U1 snRNP; Q: glutamine-rich; RRM: RNA recognition motif; SLIV: stem loop IV].

IV.1.2 The overall ability of the spliceosome to recognize a 5'ss is the net result of U1 snRNA complementarity and enhancer strength

It was found that the number of enhancer binding sites in the vicinity of a 5'ss is positively correlated with its usage in the splicing reaction (Fig.III-5). This observation was in line with previous studies showing an analogous positive impact of multiple enhancers on 3'ss activation (186, 211). Although it can be assumed that in each case only one splicing enhancer protein interacts with the U1 snRNP, multiple sites may increase the likelihood of its binding to the pre-mRNA and consequently the efficiency of 5'ss recognition (211). However, it was investigated whether weak 5'ss in dire straits to recruit the U1 snRNP on their own can be rescued by multisite enhancers in their proximity (Fig.III-6). Even the weakest 5'ss with a dramatically lessened intrinsic strength (HBS 10.8, MaxEnt 7.59) could be activated once it was combined with a strong enhancer. Furthermore, systematically testing 5'ss of variable strength in combination with enhancers of variable strength emphasised that the splicing output is functionally defined by the complementarity of the 5'ss to the U1 snRNA and its neighbouring sequences. Even though a sequence may not look like a functional 5'ss it may be rendered into a bona-fide one due to the activity of strong enhancers in its proximity. Therefore, current approaches to predict the utilization of a 5'ss are often constrained by merely regarding the sequence of a candidate 5'ss and blinding out the environment in which it is embedded. However, as argued by the results shown in this thesis, there is a gap between the intrinsic and the functional strength of a 5'ss, which is considered to be filled by the influence of adjacent splicing regulatory elements.

IV.1.3 Disclosing a “SRcret” in splice site regulation: Repression by SR proteins is coupled to “dead-end” U1 snRNP recruitment

As aforementioned almost all tested splicing factors have the capability in common to enhance splicing from one side of the 5'ss (Fig.III-4 and III-7, Fig.IV-1). However, to our surprise essentially all tested splicing proteins also could repress splicing by simply relocating their binding sites to the opposite side of the regulated 5'ss (Fig.III-9-11). The mechanistic aspects by which the position flips the activity of splicing proteins were investigated by spliceosomal complex formation analyses performed in collaboration with the Hertel Lab (UCI) and RNA pulldown assays. The spliceosome is not a preformed complex that has already acquired splicing competence, but rather assembles for each splicing reaction through a series of intermediates (Fig.I-4). Spliceosome assembly starts with early splice site recognition and binding of the U1 snRNP to the 5'ss – this stage is referred to as E (early) complex. Analysis of spliceosome assembly resulted in the unexpected finding that SR proteins regardless of their activating or repressing position did not interfere with 5'ss recognition by the U1 snRNP and E complex formation (Fig.III-13). However, the recruited U1 snRNP appeared to be a trapped or “dead-end” E complex unable to carry out later steps of spliceosome assembly (145). By contrast, U1 snRNP recruitment in the case of high affinity binding sites for hnRNP or hnRNP-like proteins seemed to be more position-dependent. Here, increased levels of U1 specific proteins could exclusively be detected on RNAs containing TIA-1 or hnRNP F/H binding sites in their activating position downstream of 5'ss D1. However, at least it appeared as if exonic hnRNP or hnRNP-like proteins did not perturb U1 snRNP binding to the 5'ss. Unfortunately, data from splicing complex analyses are not available at this time, so that it is not sufficiently clear whether repressing hnRNP or hnRNP-like proteins in exonic position use a common mechanism like SR proteins and traps the U1 snRNP from further progression in spliceosome assembly. However, non-functional, “frozen” U1 snRNP particles in presence of exonic hnRNP proteins have been proposed by others (131, 219, 375, 605). Similar to what was observed for SR proteins, repression by hnRNP proteins in these cases was achieved after early 5'ss recognition. A mechanism was suggested by which hnRNP proteins render the U1 snRNP inactive only when they are bound to an exon flanked by an inherently strong 5'ss (375). Thereby, hnRNP proteins stabilises cross-exon interactions and prevent efficient splice site pairing

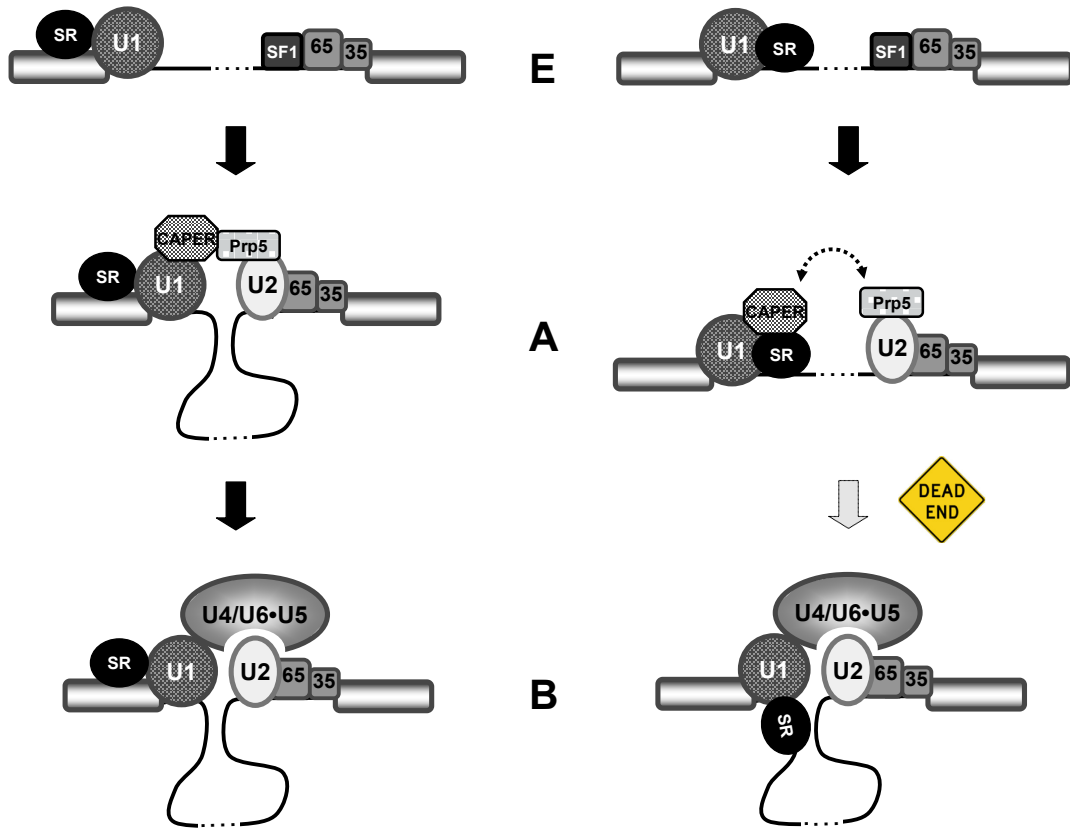
across the flanking introns; in other words they stall the transition from exon definition to intron-definition complexes (375). Although repressing positions of high affinity binding sites for hnRNP or hnRNP-like proteins had little if any effects on the basal levels of U1 snRNP recruitment (Fig.III-13B, cf. lanes 3 and 6; 13D, cf. lanes 3, 6 and 8), they ultimately could not increase the levels of precipitated U1-70K and U1-C proteins as it was observed for inhibitory SR proteins. However, this must not necessarily imply that the remaining U1 particles recruited are not also disarmed. In addition, exonic SR and intronic hnRNP proteins activate splicing by enhancement of U1 snRNP recruitment. This commonality between SR proteins and hnRNP proteins in their mode of activation rather argues for a unifying mechanism by which they repress splicing. Further experiments need to be done to clarify this issue. However, at least for SR proteins it was found that they enhance U1 snRNP recruitment irrespective of their position, but that an exonic location is required to turn the light green for further spliceosome assembly.

The accumulation of splicing-incompetent E complexes stalled to progress into a functional spliceosome was first thought to might result from protein-protein interactions between intronic SR proteins and U1-70K, deviating the U1 away from the 5'ss. To test this hypothesis, the 5'-end of the U1 snRNA was degraded by oligo-directed RNase H cleavage to inhibit basepairing interactions between U1 snRNA and the 5'ss. In all permutations U1 recruitment was greatly abrogated in the presence of 5'-end depleted U1 snRNAs. However, truncation of the free 5'-end of the U1 snRNA might also impair promiscuous basepairing interactions with the pre-mRNA at sites other than the 5'ss, potentially required for the formation of non-functional E complexes (Fig.III-15). However, mutations which inactivate the 5'ss (+1G>C) also generally blocked U1 recruitment (145), providing direct evidence that functional basepairing between U1 and the 5'ss is crucial regardless of whether functional or non-functional E complexes are formed.

IV.1.3.1 Shielding of CAPER by intronic SR proteins impair functional splice site pairing

Splice site pairing is established during A complex formation (281, 308). It was found that inhibitory SR proteins do not interfere with initial U1 recruitment and E complex formation. However, the E complexes appeared to be “dead-end” and could not be chased into A complexes. Productive association of the 5’ss and 3’ss during A complex seems to be established by RS domain mediated cross-intron interactions between the bridging factors CAPER (at the 5’ss) and Prp5 (at the 3’ss) (475). Distortion of this molecular bridge between the terminal ends of the intron results in an inability to form A complexes (475). It seems plausible that intronic SR proteins compete for CAPER or Prp5 binding, since they recapitulated the defect in A complex formation emerging from a Prp5 mutant (475). In this case, RS domain-mediated interactions between CAPER and SR protein bound immediately downstream of the 5’ss could prevent functional bridging between the splice sites and consequently A complex formation (Fig.IV-4). CAPER interacts with the U1-specific protein U1-A (475). Intronic SR proteins could stabilize U1 snRNP binding either by RS domain-mediated interactions with CAPER or RRM mediated interactions with U1-70K, explaining why swapping the position of the SR protein binding sites relative to the 5’ss does not affect the levels of U1 recruitment. However although this model provides a simple explanation for the formation of non-functional E complexes in the presence of intronic SR proteins, it depends on RS domain-mediated interactions and therefore rather suggest that hnRNP proteins repress splicing by a distinct mechanism.

A



B

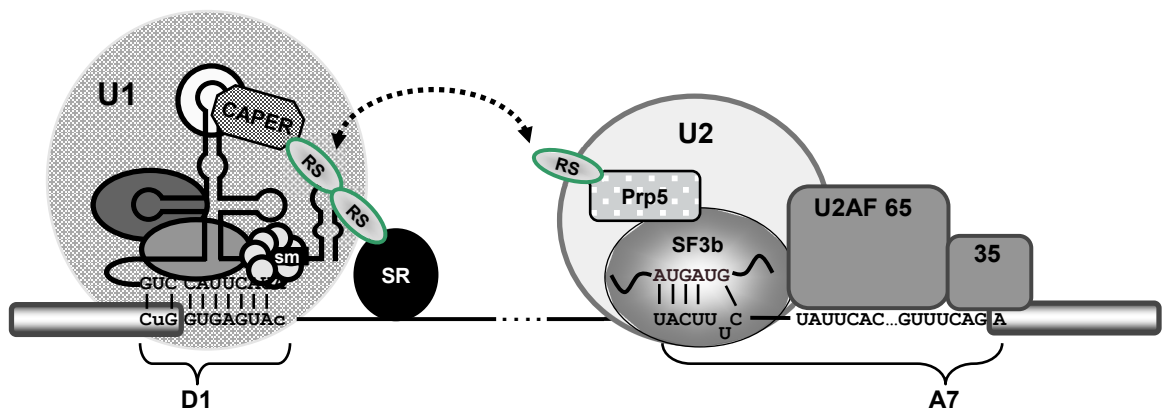


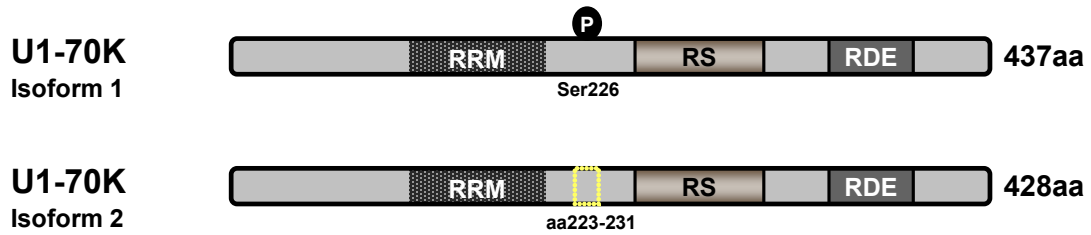
Fig.IV-4: Intronic SR proteins interfere with productive splice site pairing across the intron.

(A) SR proteins bound to the downstream intron of the regulated exon interfere with spliceosome assembly at a later step than 5'ss recognition (E complex). Repressing SR proteins perturb functional splice site pairing for transition into A complex. (B) RS domain-mediated shielding of CAPER prevents the establishing of productive interactions between CAPER and Prp5 across the intron and thereby arrest further progression into a functional spliceosome (475). Interactions are highlighted in green. [U1: U1 snRNP; 65: U2AF65; 35: U2AF35; U2: U2 snRNP].

IV.1.3.2 Recruitment of qualitatively distinct U1 snRNPs to the 5'ss

Some studies indicated that cells could express U1 snRNPs, which are different in their protein composition (209, 290) or which are formed on the basis of a U1 snRNA variant (289). Although this introduces the idea that dependent on their position splicing factors might recruit qualitatively distinct U1 snRNPs to the 5'ss, a splicing competent “good” or a splicing-incompetent “bad” one, it is mechanistically hard to imagine how one and the same splicing factor simply knows by its location which of both it shall recruit. However, the U1 snRNP can alternatively incorporate two distinct U1-70K isoforms each preferring either tight binding to U1-C (isoform 1) or SmB/B' (isoform 2), respectively (209). It was proposed that this might affect the spatial organization of the U1 snRNP. Notably, the proposed region of these interactions was implicated to be play a role for the formation of the U1/U2 network and juxtapositioning of the splice sites (132). Dependent on these considerations, repressing SR proteins might prefer recruitment of those U1 snRNPs containing U1-70K isoform 2, since this is potentially more accessible for RS domain-mediated interactions from a position within the downstream intron. However, the tighter interactions between U1-70K isoform 2 and SmB/B' might also shield the proximal U1 snRNP surface, which is necessary for correct interaction with U2 snRNP and productive splice site pairing (132, 209). Alternatively, SmB/B' might be impaired to function as a binding platform for additional splicing factors, which have to join the U1 snRNP for further spliceosome assembly (26, 27). Interestingly, it was shown that the C-terminal tail of SmB/B' interacting with U1-70K isoform 2 also directly binds to FBP21 (formin-binding protein 21), which is likewise Prp5 and CAPER supposed to serve as a bridging factor bringing both splice sites into proximity (26, 222, 273). As a conclusion, U1-70K isoform 2 favored by intronic SR proteins might compete for SmB/B' binding and prevent appropriate cross-intron bridging, thereby rendering E complexes inactive (Fig.IV-5). Taken together, repression by SR protein might prefer recruitment of U1-70K isoform 2-incorporating U1 snRNP particles incapable to recruit FBP21 and establish functional splice site pairing. Both U1-70K isoforms only differ in 9 amino acids, less to be sufficiently resolved by the polyacrylamide gels used in the RNA pulldown experiments. Accordingly, it awaits further studies to elucidate a potential preference of intronic SR proteins for recruitment of U1 snRNP particle specifically incorporating U1-70K isoform 2.

A



B

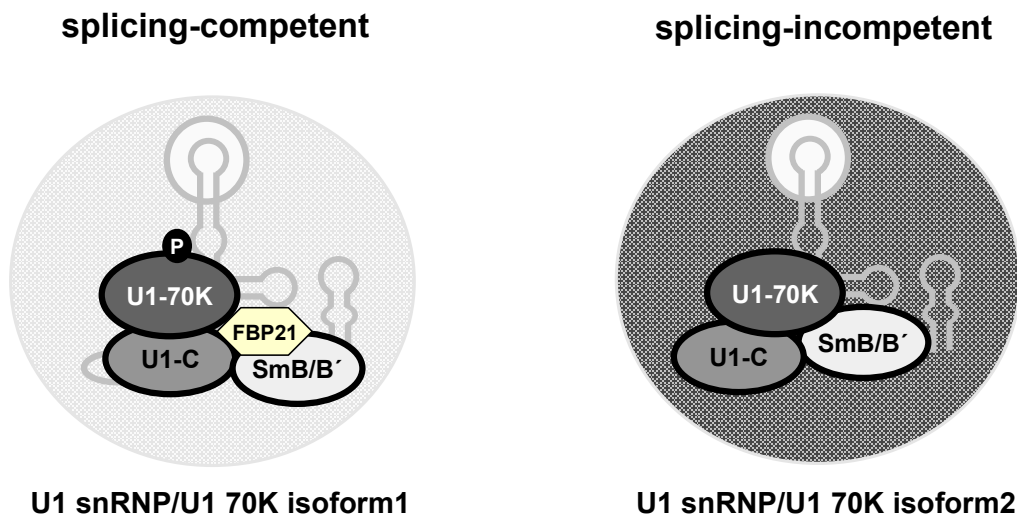


Fig.IV-5: Intronic SR proteins preferentially interact with U1 snRNP particle containing U1-70K isoform 2, incapable to perform the splicing reaction.

(A) Two distinct isoforms of U1-70K can be incorporated into the U1 snRNP particle of which isoform 2 lacks nine amino acids (aa223-231) harbouring a phosphorylated serine at position 228 (209). (B) U1-70K isoform 1 preferentially interact with the C-terminal tail of the U1-specific protein U1-C, leaving the C-terminus of SmB/B' available for interaction with the putative bridging protein FBP21 (26). Intronic SR proteins might prefer recruitment of U1 snRNP particles containing U1-70K isoform 2, since 3D structure indicate a better accessibility of the RS domain. Those U1 snRNP's might be rendered inactive for splicing due to competition of U1-70K isoform 2 with FBP21 for binding to SmB/B'. [Ser: serine; P: phosphorylation; U1: U1 snRNP].

IV.1.4 Inhibition of spliceosome assembly after 5'ss recognition might promote intron retention versus exon skipping

Repression by SR (and potentially hnRNP) proteins occurs at later steps than U1 snRNP recruitment and E complex formation. There are remaining questions, which needs to be addressed; among the first: i) are repressing splicing factor positions commonly found in the vicinity of regulated 5'ss within the genome and ii) why to make an effort enhancing the recruitment of an ultimately splicing-incompetent U1 snRNP.

Computational analyses revealed that while constitutive exons are enriched in splicing factor binding sites within activating positions, alternatively spliced exons display a higher frequency of intronic SR and exonic hnRNP protein binding sites (145). Constitutively spliced exons would be expected to avoid intronic SR or exonic hnRNP protein binding sites interfering with their proper inclusion into the mRNA. By contrast, repressing positions should be overrepresented within alternatively spliced exons to allow the regulated use of their splice sites. Although it is not known whether these putative sites may be engaged in splicing, the genome analysis suggests a general importance of position-dependent splicing by SR and hnRNP proteins.

It is not fully clear for which reason intronic SR proteins should recruit a non-functional U1 to the 5'ss or at least a functional one only to render it splicing-incompetent later. The U1 snRNP is the most abundant snRNP present within the nucleus and it carries out functions aside from its prominent role in pre-mRNA splicing. SR proteins in intronic positions may be crucial to recruit the U1 snRNP to the pre-mRNA and protect the pre-mRNA against degradation (253) or inhibit promiscuous polyadenylation at unspecific sites (1, 158, 250). However, this is not to say that recruitment of a splicing-incompetent U1 snRNP by intronic SR proteins may not also be necessary for specific modes of alternative splicing. In a multi-exon context non-recognition of the splice sites at an alternative exon might lead to exclusion of the exon and its flanking intron sequences from the mRNA, while efficient splice site recognition would be expected to result in exon inclusion. When considering a splicing-incompetent U1 snRNP bound to an internal exon is defect in splice site pairing across the downstream intron but simultaneously capable to pair with the upstream 3'ss, this would favor specific retention of the downstream intron

versus skipping of the entire exon (Fig.IV-6). In accordance with this model, SR proteins in intronic positions might be rather intron retention than exon skipping factors. There are some cases, which support the notion of a “semi-functional” U1 snRNP active in upstream exon definition, while defect to form A complexes across the downstream intron. For instance, it was found that an intronic positioning of the SR protein-dependent GAR ESE (11, 72, 253) immediately downstream of an alternatively spliced exon largely activates the upstream 3'ss, while leading to retention of the intron in which it is embedded (140). Tethering of SR protein-derived MS2 fusion proteins to the same position recapitulated this finding (140, 499). Furthermore, recruitment of U1 snRNPs with optimized complementarities to splicing-defect 5'ss sequences markedly activate the upstream 3'ss and promote a shift from exon-skipping isoforms to those, which specifically retain the downstream intron (Fig.III-29). This reinforces the idea that intronic SR proteins are specific intron retention factors recruiting splicing-incompetent U1 snRNP to the 5'ss, which maintain their capability to promote exon definition and enforce activation of the upstream 3'ss. Furthermore this suggests that the modes of action by SR proteins can be separated into position-dependent and position-independent effects. Accordingly, SR proteins are enhancers of early 5'ss recognition and formation of exon definition complexes regardless of their relative position to the 5'ss. However, their exonic or intronic position determines whether the U1 snRNP may or may not be engaged in the splicing reaction. Moreover, this proposes a model in which intron retention is essentially related to early recognition of the splice sites at an alternative exon.

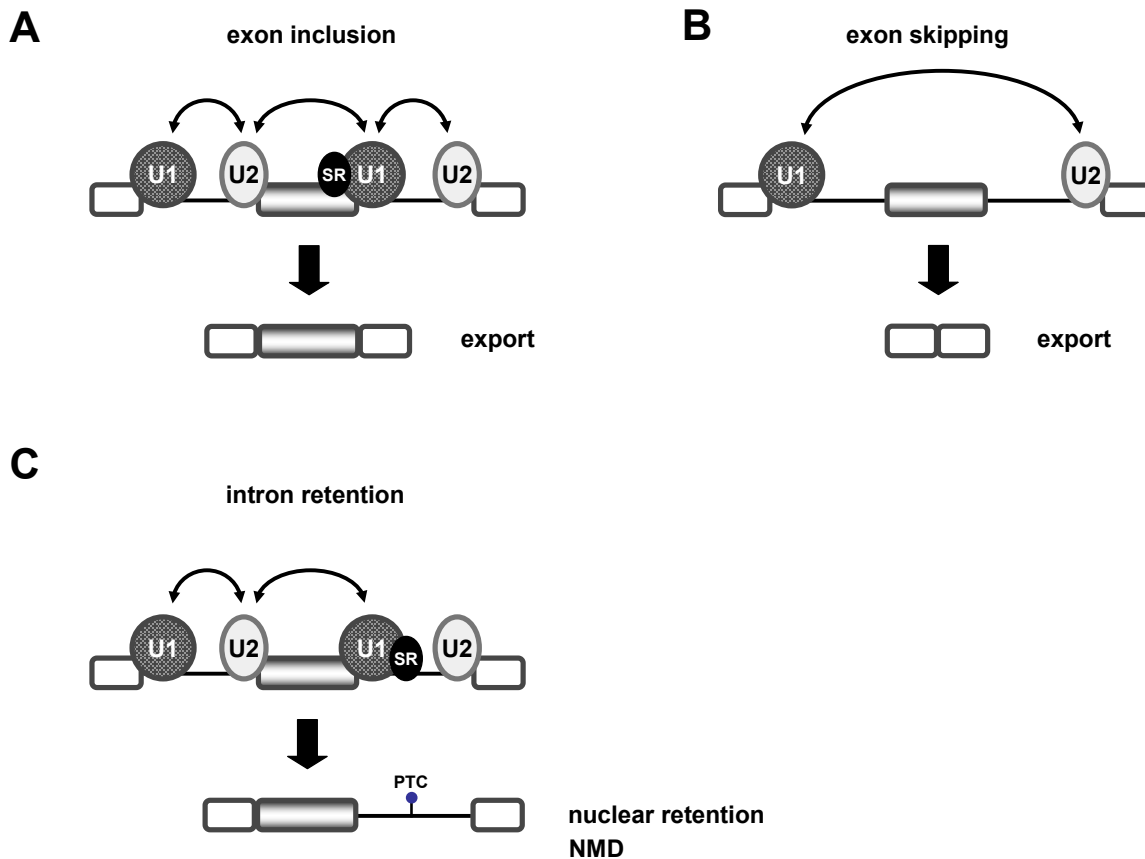


Fig.IV-6: Intronic SR proteins may promote intron retention.

SR proteins bound to the exon enhance recruitment of the U1 snRNP to the 5'ss, thereby leading to the inclusion of a regulated internal exon (**A**). In absence of an enhancer protein, the flanking splice sites are not recognized and the exon is excluded from the mRNA (**B**). SR proteins bound immediately downstream of the regulated exon promote U1 snRNP recruitment to the nearby 5'ss and splice site pairing across the upstream exon. However, the U1 snRNP may not be used in the splicing reaction, leading to selective retention of the downstream intron (**C**). Intron retention sequesters the mRNA within the nucleus and increases the likelihood for the occurrence of premature termination codons (PTCs), eliciting nonsense-mediated decay (NMD). [U1: U1 snRNP; U2: U2 snRNP; PTC: premature termination codon; NMD: nonsense-mediated decay].

Therefore, intronic SR proteins might exert a specific intron retention activity linking alternative splicing to a reduced expression of cellular mRNAs. Intron-containing mRNAs normally fail to enter the cytoplasm and are readily degraded within the nucleus (390, 451, 507). Furthermore retention of an intron may lead to the inclusion of premature stop codons (PTCs) into the mRNA, marking it for destruction by nonsense-mediated decay (NMD) (295). The expression of intron-containing splice isoforms was shown to vary along different tissues (269, 301, 423), indicating that relative concentrations of SR proteins may determine the frequency of intron

retention events in a specific cell. Furthermore, retained introns should follow a trend to be relatively enriched in intronic SR protein binding sites. This indeed appears to be the case, since a recent study could find that putative SR protein binding sites are found in higher frequency within retained introns when compared to non-retained ones (452). Interestingly, SR proteins are thought to negatively autoregulate themselves by coupling their own mRNAs to degradation by NMD (296, 527, 528). Transcripts are rendered sensitive to NMD through recognition of PTCs included by unproductive splicing events into the mRNA (230). Interestingly, ultraconserved elements were found within the introns of SR protein pre-mRNAs that appear to be linked to the emergence of PTC-containing isoforms (296, 386). Furthermore, retention of introns harbouring the ultraconserved elements was found to be a prevalent cause for the inclusion of PTCs into SR protein mRNAs (296). It is tempting to speculate that SR proteins may bind to ultraconserved conserved elements (or other sites) within introns of their own pre-mRNAs, thereby triggering dead-end splicing complex formation, intron retention and ultimately destruction of the resulting mRNAs by NMD to autoregulate their own expression and maybe that of other cellular proteins. Retention of a complete intron versus exon skipping or inclusion of a cryptic exon may be advantageous, since human introns are considerably larger when compared to their flanking exons, which should render them more likely to contain a PTC. SR proteins were proposed to directly enhance NMD of spliced PTC-containing reporter mRNAs (618). Interestingly, overexpression of SR proteins also reproducibly increased intron retention in varying experimental set-ups (data not shown), and it was not fully clear to which extent RNAs were filtered off by sequestration within the nucleus or PTCs within the intron and their subsequent degradation by other mechanisms (618). However, intronic SR protein binding might also be of importance for the formation of specific intron-containing HIV-1 mRNA species (chapter IV-2), which require activation of a 3'ss, but simultaneous retention of the downstream intron.

IV.1.5 Future prospects

It was clearly shown that SR and hnRNP proteins can activate and repress splicing from defined positions relative to the 5'ss. In addition, it was demonstrated that SR proteins enhance recruitment of the U1 snRNP to the 5'ss regardless of their location within the upstream exon or downstream intron, respectively. However, intronic SR proteins render the U1 snRNP splicing-incompetent to carry out later stages of spliceosome assembly. Immunodepletion of CAPER and Prp5 followed by complementation of HeLa nuclear extracts with a recombinant CAPER-Prp5 fusion protein or CAPER and Prp5 each containing the glycine-rich interaction domain of hnRNP A1 instead of the RS domain will show whether repression by SR proteins can be bypassed. Apart from this, higher resolution polyacrylamide gels will clarify whether intronic SR proteins preferentially recruit U1 snRNP containing U1-70K isoform 2. Probing with a specific antibody directed will reveal whether the recruited U1 snRNPs by intronic SR proteins are incapable to interact with FBP21. It requires splicing complex formation investigations to figure out whether exonic hnRNP and hnRNP-like proteins use a similar mode for repression and inhibit splicing after early splice site recognition. Use of the MS2 tethering system will help to characterize the particular contribution of respective protein domains of SR and hnRNP proteins for activation and repression of splicing. Furthermore, it will be of interest to determine whether relative position of SREs also plays a role for activation or repression of regulated 3'ss. To substantiate the proposal that intronic SR proteins recruit a U1 snRNP to the 5'ss, which is still capable to productively pair with the upstream 3'ss, but cannot be used in the splicing reaction, analysis should be extended to multi-exon splicing reporters.

IV.2 An exonic splicing enhancer (ESE_{vpr}) within HIV-1 exon 3 is required for exon 3 splicing and *vpr*-mRNA expression

Splice site recognition is commonly found to be under combinatorial control of multiple splicing regulatory elements in the vicinity, which can either compete or cooperate to regulate splicing activation. It was previously shown that the HIV-1 non-coding leader exon 3 harbours a negative regulatory splicing element – termed ESSV – within its central portion that selectively represses upstream 3'ss A2 (131) and concomitantly inclusion of exon 3 into the different viral mRNA species (329). Disruption of the ESSV results in strong accumulation of *vpr*-mRNA and exon 3-containing isoforms, which is detrimental for viral replication. However, it was found in this thesis that the excessive splicing phenotype is related to the activity of a novel splicing enhancer located immediately downstream of the ESSV. Simultaneous mutation of the enhancer could reset exon 3 over-splicing and inhibition of viral replication in absence of the ESSV. However, also in presence of the ESSV, a functional enhancer was necessary for expression of *vpr*-mRNA and exon 3 splicing. In support of the exon definition hypothesis, ESE_{vpr}-mediated binding of the U1 snRNP to the 5'ss D3 controlled the usability of the upstream 3'ss A2, which is necessary for *vpr*-mRNA expression. However, *vpr*-mRNA expression furthermore relies on the retention of the downstream intron. It is suggested that splice site pairing between 3'ss A2 and 5'ss D3 across HIV-1 exon 3 must occur with higher efficiency than productive splice site pairing between 5'ss D3 and the downstream 3'ss A3, A4cab and A5. Expression of *vpr*-mRNA might rely on the mechanism proposed in chapter IV-1 wherein intron retention is achieved after early splice site recognition by formation of non-functional E complexes.

IV.2.1 An exonic splicing enhancer (ESE_{vpr}) promotes exon 3 inclusion and formation of *vpr*-mRNA

The mutational analysis revealed that a sequence immediately downstream of the ESSV is required for exon 3 splicing and expression of *vpr*-mRNAs. Disruption of the enhancer abolished inclusion of exon 3 into the different viral mRNA species and impaired production of *vpr*-mRNAs (Fig.III-25-26). In absence of the negative

regulatory ESSV, the enhancer was shown to be responsible for emergence of an excessive exon 3 splicing phenotype, dramatically reducing the relative levels of unspliced mRNAs and viral replication. However, simultaneous inactivation of the ESSV and the enhancer restored the wildtype splicing profile and re-enabled accumulation of unspliced mRNA and virus particle production. It was concluded that balanced exon 3 splicing is under combinatorial control of the ESSV and the enhancer (Fig.IV-7).

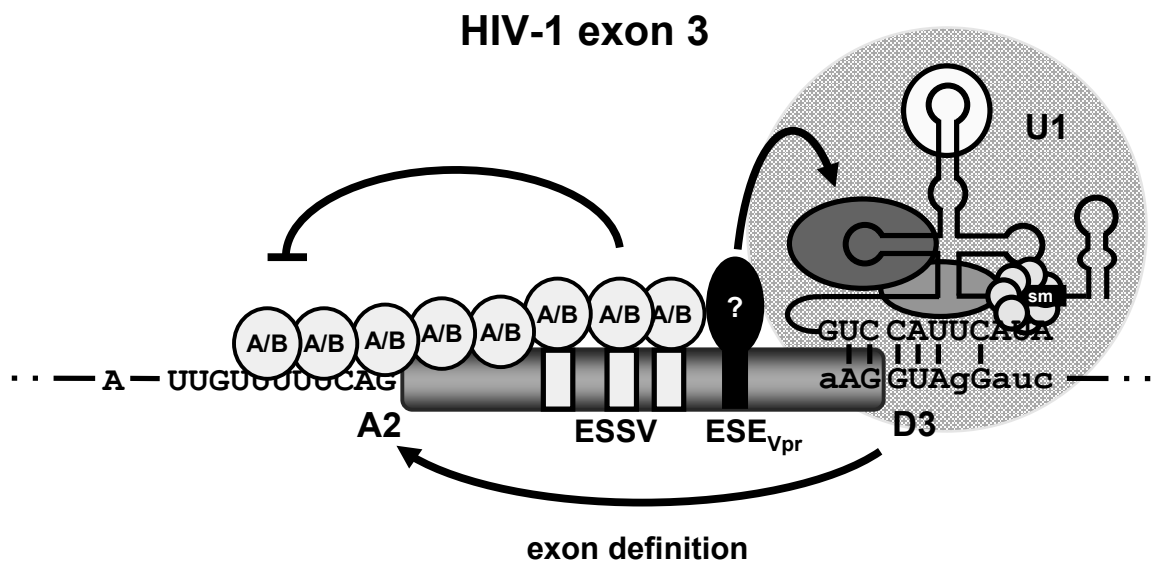


Fig.IV-7: HIV-1 exon 3 splicing is under combinatorial control of ESSV and ESE_{vpr}.

Splice site recognition at HIV-1 exon 3 is regulated by the exonic splicing silencer V (ESSV) and a novel exonic splicing enhancer (ESE_{vpr}) embedded in the region upstream of 5'ss D3. The ESSV is associated with hnRNP A/B proteins, which may multimerize along the 5'-end of exon 3, occluding 3'ss A2. The ESE_{vpr} might enhance binding of the U1 snRNP to 5'ss D3, in turn emanating interactions across the upstream exon and activation of 3'ss A2. [U1: U1 snRNP].

The relative location of the enhancer downstream of the ESSV anticipated that it may act on recognition of the adjacent 5'ss D3. Revisiting the exon definition hypothesis that binding of the U1 snRNP to the 5'ss promote interactions across the exon and thereby activation of the upstream 3'ss (11, 215, 337), 5'ss D3 variations were analyzed either with decreasing or increasing complementarity to the U1 snRNA. The 5'ss mutants were tested in the context of the varying exon 3 mutants. In agreement with the notion that the splicing enhancer promotes U1 snRNP recruitment and thereby exon definition, up and down mutations of 5'ss D3 increased or decreased

vpr-mRNA expression levels and exon 3 splicing into the viral mRNA species, respectively (Fig.III-28). However, in absence of the enhancer, even the strongest 5'ss failed to activate *vpr*-mRNA and exon 3 splicing, underlying the necessity of a functional enhancer upstream of the 5'ss regardless of its intrinsic strength. By contrast, even the weakest 5'ss could not reduce the levels of both in absence of the silencer. However, these results were reminiscent to what was observed earlier (see chapter IV-1.2), namely that the splicing output is not solely determined by the intrinsic strength of a 5'ss but essentially by its embedment into a regulatory network of neighbouring splicing regulatory elements. By the use of a splicing-incompetent 5'ss (GTV), which otherwise perfectly matches to the U1 snRNA, it was demonstrated that binding of the U1 snRNP to the 5'ss suffices to enhance exon definition and activation of 3'ss A2 for *vpr*-mRNA formation (Fig.III-29). This finding was reiterated by inactivation of 5'ss D3 (+1G>C) and coexpression of a fully complementary U1 snRNA, which largely activated 3'ss A2 and *vpr*-mRNA splicing (Fig.III-32). Altogether, these data provided strong evidence that 3'ss A2 and *vpr*-mRNA splicing are coupled to U1 snRNP-elicited exon definition. Unfortunately, so far all attempts to identify host proteins, which bind to the enhancer sequence, failed and thus it awaits further experiments to clarify, which splicing factors confers activation of 5'ss D3. However, a recent study suggested that the high mobility group A protein 1a (HMGA1a) might bind to the candidate enhancer sequence, but in order to repress the use of the downstream 5'ss D3 [also see chapter IV.2.4; (548)].

IV.2.2 Intronic SR proteins might be implicated in specific retention of the *vpr*-intron

However, it is poorly understood how formation of intron-containing HIV-1 mRNAs is regulated. In particular, *vpr*-mRNA expression requires activation of 3'ss A2, while the downstream intron harbouring the *vpr*-AUG must be retained. Therefore one can hypothesize that splice site pairing between 5'ss D3 and any of the 3'ss A3, A4cab or A5 needs to occur at least less efficiently than upstream splice site pairing between 5'ss D3 and 3'ss A2. Consistent with the model proposed in chapter IV-1, intron-retention for proper *vpr*-mRNA expression might involve the inhibition of spliceosome assembly after early splice site recognition. In this case, either the U1 snRNP bound

to 5'ss D3 or the U2 snRNP bound to any 3'ss could be rendered incapable of progressing on with splicing complex formation. This could also be achieved by an SR protein-dependent mechanism. The 3'ss A3 was found to be positively regulated by the SRSF2-dependent enhancer (ESE2) located within the downstream exon (194, 606). It was proposed that SRSF2 promotes activation of 3'ss A3 by antagonizing binding of hnRNP A1 to an overlapping silencer termed ESS2 (4, 73, 493). A second repressing element (ESS2p) was identified, which is placed immediately downstream of 3'ss A3 and binds to hnRNP H (238). However, SRSF2 bound to ESE2 is in an exonic position relative to 3'ss A3, but in an intronic position relative to the more distant 3'ss cluster A4cab and A5. In agreement with the idea that intronic SR proteins compete for CAPER (at the 5'ss) or Prp5 (at the 3'ss), ESE2 binding might interfere with productive splice site pairing across the intron and A complex formation (475).

SRSF2 bound to the ESE2 might inhibit the establishment of productive 5'ss D3 and 3'ss A4cab or A5 interactions by competing for Prp5 (Fig.IV-8). This is supported by the model of Hallay et al. in which SRSF2 does not actively recruit spliceosomal components to the upstream 3'ss A3, but blocks cooperative binding of hnRNP A1 molecules along the 5'-end of the exon (194). This would release the RS domain for inhibitory interactions with Prp5 and would result in the failure of the U2 snRNP to engage the occupied 3'ss. Furthermore 3'ss A3 could be defined as proximal and 3'ss A4cab or A5 as distal relative to 5'ss D3. It was shown previously that SR proteins bound between the proximal and the distal 3'ss enhance proximal splicing and intron retention at the same time (229).

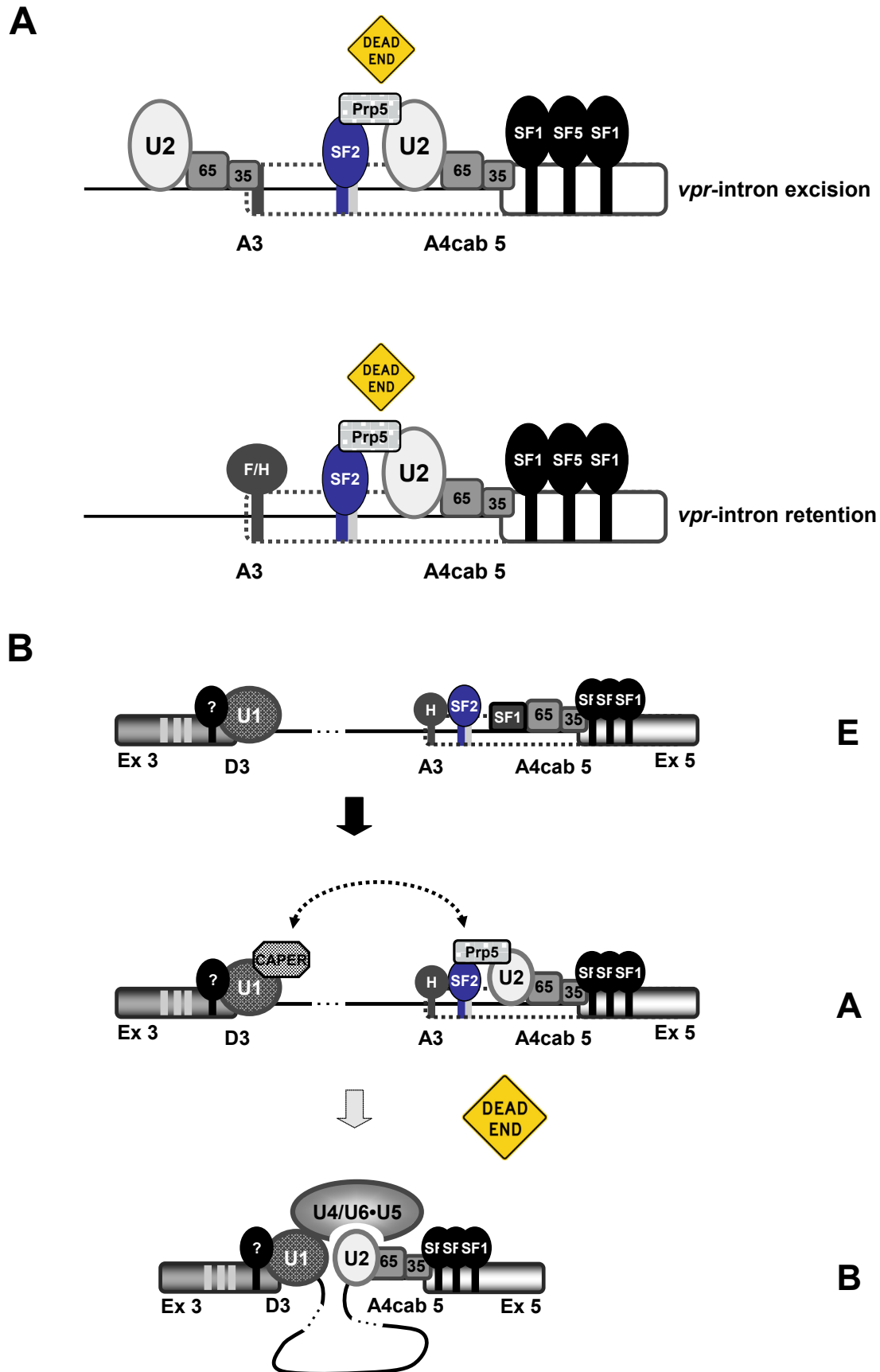


Fig.IV-8: SRSF2 and hnRNP F/H might act together to promote *vpr*-intron retention. [continued on next page]

However, little is known about the changes in the relative expression levels of splicing factors during HIV-1 infection. A recent study indicates that there could exist a time window during infection in which the levels of SRSF2 and hnRNP H are relatively higher than those of hnRNP A/B proteins (135). Unfortunately, *vpr*-mRNA levels have not been determined in this study and thus, the importance of the relative nuclear levels of SRSF2, hnRNP H and hnRNP A/B proteins for *vpr*-intron retention remains unclear.

Dead-end complex formation at 3'ss A4cab or A5 would shift the respective *rev*-, *nef*- and *env*-mRNAs to unspliced and intron-containing *vif*- and *vpr*-mRNAs. Recognition of exon 2 and exon 3 in combination with a dead-end U2 snRNP at 3'ss A4cab and 5 would lead to the formation of *vif*- or *vpr*-mRNAs. However, recognition of neither exon 2 nor exon 3 would promote expression of unspliced-mRNAs. On basis of this mechanism, it would be expected that overexpression of each protein alters the relative levels of intron retention versus intron excision. Splicing of *vpr*-mRNAs would rely on binding of both, hnRNP H to negatively regulate 3'ss A3, and SRSF2 to render U2 snRNP at the downstream 3'ss A4cab or A5 inactive for functional splice site pairing.

Fig.IV-8: continued

(A) SRSF2 binds to the ESE2 downstream of 3'ss A3. SRSF2 renders U2 snRNP bound at the 3'ss A4cab and A5 inactive by competing for the bridging factor Prp5. However, absence of hnRNP F/H proteins at the ESS2p leads to predominante activation of 3'ss A3 and *tat*-mRNA rather than *vpr*-mRNA expression. Binding of hnRNP F/H proteins to the ESS2p interferes with 3'ss A3 recognition. Dead-end U2 snRNP formation by SRSF2 promotes retention of the upstream *vpr* intron.

(B) Lower concentrations of hnRNP A/B proteins ensure activation of the *vpr*-specific 3'ss A2 and access of SRSF2 to the ESE2 to trap the U2 snRNP at the downstream 3'ss. At the same time increased levels of hnRNP H inhibit recognition of 3'ss A3 through binding to the proximal ESS2p. This inhibits that 5'ss D3 splices to 3'ss A3 enhancing formation of *tat*-mRNAs. Altogether, this then leads to activation of 3'ss A2 and retention of the downstream *vpr*-intron. [U1: U1 snRNP; U2: U2 snRNP; 65: U2AF65; 35: U2AF35].

IV.2.3 HMGA1a might trap the U1 snRNP at 5'ss D3 to promote downstream *vpr*-intron retention

An alternative model for *vpr*-intron retention was recently suggested, wherein the high mobility group A protein 1a (HMGA1a) binds upstream of 5'ss D3 and traps the U1 snRNP incapable for splicing (548548). Two candidate binding sequences for HMGA1a were mapped to the 3'-terminal region of HIV-1 exon 3 of which the first one is located within the ESE_{*vpr*} characterized in this thesis. HMGA1a EMSAs revealed that binding to HIV-1 exon 3 was abolished by double-mutation of both putative target sites. However, it was unambiguously shown that ESE_{*vpr*} promotes splicing at the downstream 5'ss and this activity was disrupted by site-specific mutations – notably also within an unrelated enhancer-dependent splicing reporter. Therefore, it seems unlikely that HMGA1a binds to the first putative binding site more distal to 5'ss D3 and hence renders the U1 snRNP splicing-incompetent. Our experiments also indicate that even if the second binding motif, which overlaps with the exonic residues of the 5'ss D3 sequence, is engaged, the upstream enhancer is dominant over the HMGA1a splicing repressor activity. However, HMGA1a was shown to interact with hnRNP A1, which binds to the upstream ESSV exerting a repressing function on 3'ss A2. This proposes a model in which stable binding HMGA1a and the extent of its activity depend on the upstream ESSV and ESE_{*vpr*}, respectively (Fig.IV-9). In presence of the enhancer, U1 snRNP is efficiently recruited to the 5'ss. A certain fraction of the bound U1 snRNP is switched to a splicing-incompetent state by weak HMGA1a binding to the 5'ss proximal target sequence within exon 3. However, ESE_{*vpr*} might function as barrier to upstream bound hnRNP A1 proteins, preventing their spreading along the 3'-end of exon 3, which might stabilize HMGA1a binding. According to this mechanism, ESE_{*vpr*} might be required for efficient recognition of 5'ss D3 but also balances the levels between *vpr*-mRNA and exon 3 splicing by acting as barrier for stable HMGA1a binding to exon 3 [Fig.IV-9, (a)]. Coexpression experiments with a fully complementary U1 snRNA directed against 5'ss D3 support this notion, since the inactivation of ESE_{*vpr*} leads to substantially higher amounts of *vpr*-mRNA splicing, whereas use of 5'ss D3 seems to be avoided (Fig.III-30). The significant upregulation of *vpr*-mRNA expression indicated that the modified U1 snRNP efficiently bound to 5'ss D3 and in agreement with the exon definition hypothesis activated upstream 3'ss A2. However, splicing at

5'ss D3 was rather inefficient in absence of ESE_{vpr} . Consistent with the suggested model, this observation might be due to an hnRNP A1-imparted stabilization of HMGA1a binding that only takes place in absence of ESE_{vpr} , thereby resulting in higher levels of inactive U1 snRNPs and *vpr*-mRNA splicing [Fig.IV-9, (b)]. Accordingly, absence of both, ESSV and ESE_{vpr} then again shows strong exon 3 inclusion into the viral mRNA species, suggesting that hnRNP A1 cannot bind to multimerize along the downstream 3'-end of exon 3 to promote HMGA1a binding [Fig.IV-9, (d)]. Whether HMGA1a really plays a role for *vpr*-mRNA splicing regulation by trapping of the U1 snRNP at the 5'ss D3 awaits further studies.

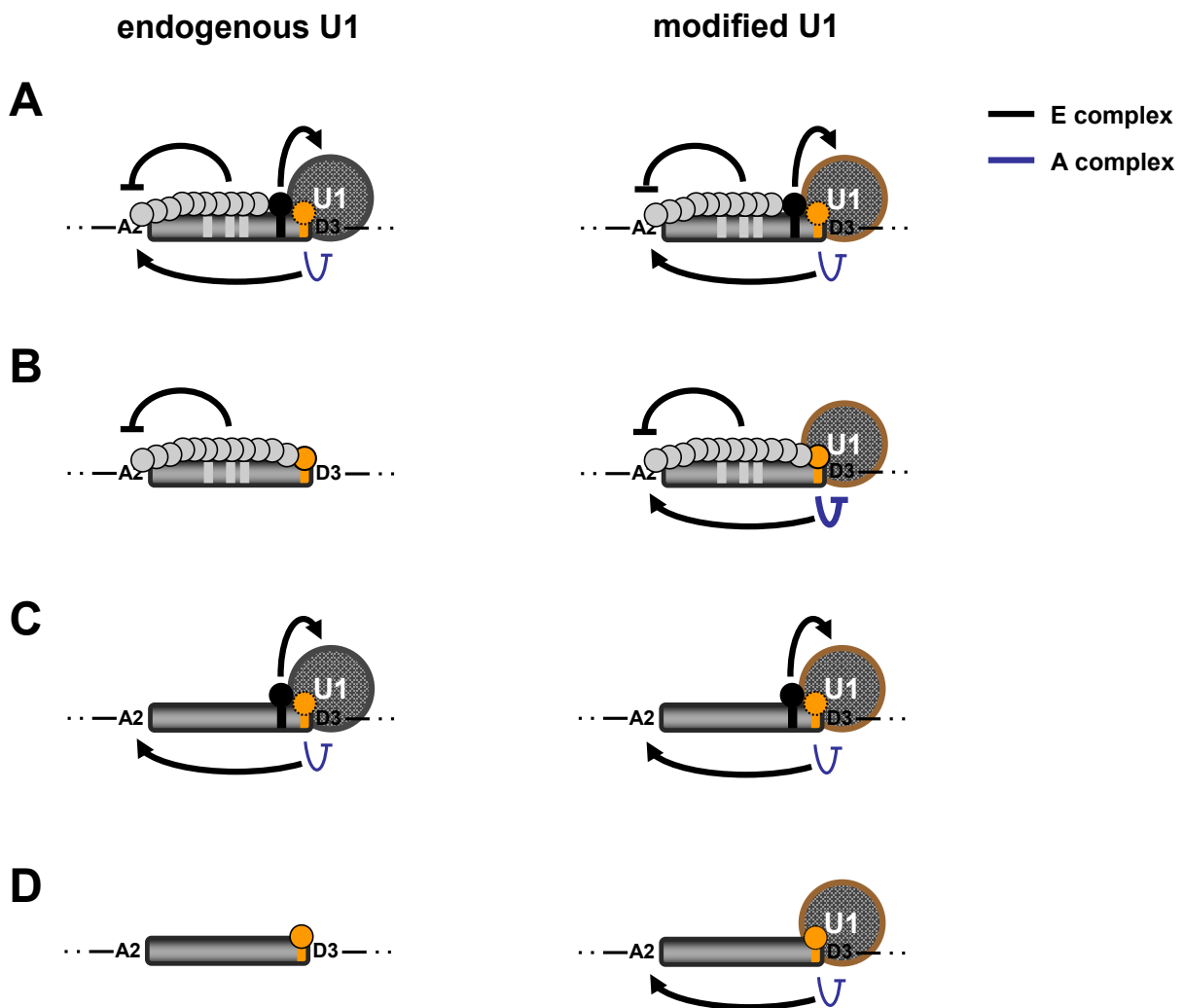


Fig.IV-9: ESSV and ESE_{vpr} control binding of HMGA1a. [continued on next page]

IV.2.4 Future prospects

Future experiments will extend the RNA affinity chromatography analyses to identify host proteins, which bind to the the ESE_{vpr} sequence. Overexpression or siRNA-mediated knockdown will be performed to elucidate the functional importance of the identified protein(s) for ESE_{vpr} activity in the context of the LTR ex2 ex3 splicing reporter and replication-competent provirus. Mutational analysis of ESE2 will reveal whether SRSF2 binding upstream of 3'ss A4cab and A5 is required for dead-end U2 snRNP formation and *vpr*-intron retention. Inactivation of the putative HMGA1a binding site within HIV-1 exon 3 will provide insight into the importance of HMGA1a binding for *vpr*-intron retention.

Fig.IV-9: continued

(A) In presence of both, ESSV and ESE_{vpr}, stable binding of HMGA1a (orange) is prevented by proteins bound to the enhancer sequence (black), acting as barrier for hnRNP A/B protein (grey) spreading along the 3'-end of exon 3 (on the left). As a result, only a minor fraction of U1 snRNPs bound to the 5'ss D3 is rendered inactive. However, hnRNP A/B proteins multimerize across the 5'-end of exon 3, thereby occluding 3'ss A2 and inhibiting efficient exon 3 inclusion and *vpr*-mRNA expression. Coexpression of a 5'-end modified U1 snRNA (on the right) activates exon 3 splicing and *vpr*-mRNA expression. **(B)** In absence of the enhancer, binding of HMGA1a is stabilized through interaction with hnRNPA/B proteins. Recruitment of the U1 snRNP to the 5'ss D3 is significantly impaired, resulting in abrogation of exon 3 splicing and *vpr*-mRNA formation. However, coexpression of a 5'-end modified U1 snRNA substantially increases *vpr*-mRNA expression by activation of 3'ss A2 via interactions across the exon. Stable binding of HMGA1a at the same time renders the U1 snRNP more frequently inactive for the splicing of the downstream intron, favouring *vpr*-mRNA formation against inclusion of exon 3. **(C)** In absence of the ESSV splicing at exon 3 and *vpr*-mRNA expression are increased due to relieved repression on 3'ss A2. However, the enhancer interferes with stable HMGA1a binding, allowing considerable splicing at 5'ss D3. Coexpression of the fully complementary U1 snRNP leading to both, increase of exon 3 inclusion and *vpr*-mRNA expression, reminiscent to sup-optimal HGMA1a binding to exon 3 **(D)** In absence of ESSV and ESE_{vpr} exon 3 splicing and *vpr*-mRNA expression are reduced. Coexpression of the 5'-end mutated U1 snRNA then predominantly shifts towards exon 3 splicing versus *vpr*-mRNA expression due to the absence of stabilizing interactions between hnRNP A/B proteins and HMGA1a. [U1: U1 snRNP].

V. Conclusions

It was found that almost all known splicing regulatory proteins can activate and repress splicing dependent on their relative position to a regulated 5'ss. While SR proteins needed a strict positioning into the upstream exon to enhance 5'ss use, hnRNP or hnRNP-like proteins exclusively facilitated splicing from an intronic location. However, SR and hnRNP or hnRNP-like proteins could also repress splicing by simply relocating them to the opposite side of the 5'ss. Conclusively, the classical definition of SR proteins as enhancers and hnRNP proteins as repressors calls for a reassessment, since almost all splicing factors appear to act as both in a position-dependent manner. In addition, it was demonstrated that the tested splicing factors all enhanced recruitment of the U1 snRNP from their activating positions. Furthermore, it was shown that intronic SR proteins do not perturb recognition of the 5'ss by the U1 snRNP, but rather stall functional spliceosome assembly at a later step than E complex formation. It remains unclear whether hnRNP or hnRNP-like proteins uses a common mode for repression.

In the second part of this thesis, it was found that an exonic splicing enhancer, termed ESE_{vpr}, counteracts the negative regulatory ESSV in the central part of HIV-1 exon 3. It was proposed that the ESE_{vpr} supports 3'ss D3 recognition by the U1 snRNP, which promotes functional splice site pairing across the exon and activation of 3'ss A2 for *vpr*-mRNA expression. Inactivation of the ESE_{vpr} in the context of an ESSV-negative virus could restore exon 3 splicing and viral replication, indicating that ESE_{vpr} activity primarily accounts for the excessive splicing phenotype observed in ESSV mutants. Accordingly, balanced HIV-1 exon 3 splicing is established by an interplay of functional opposing elements, adjusting activation at the flanking splice sites to levels compatible with both (i) viral replication and (ii) *vpr*-mRNA expression.

VI. References

1. **Abad, X., M. Vera, S. P. Jung, E. Oswald, I. Romero, V. Amin, P. Fortes, and S. I. Gunderson.** 2008. Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic acids research* **36**:2338-2352.
2. **Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin.** 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *Journal of virology* **59**:284-291.
3. **Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger.** 1996. CC CKR5: a RANTES, MIP-1alpha, MIP-1beta receptor as a fusion cofactor for macrophage-tropic HIV-1. *Science* **272**:1955-1958.
4. **Amendt, B. A., D. Hesslein, L. J. Chang, and C. M. Stoltzfus.** 1994. Presence of negative and positive cis-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1. *Molecular and cellular biology* **14**:3960-3970.
5. **Amendt, B. A., Z. H. Si, and C. M. Stoltzfus.** 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. *Molecular and cellular biology* **15**:6480.
6. **Amendt, B. A., Z. H. Si, and C. M. Stoltzfus.** 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. *Molecular and cellular biology* **15**:4606-4615.
7. **Amir-Ahmady, B., P. L. Boutz, V. Markovtsov, M. L. Phillips, and D. L. Black.** 2005. Exon repression by polypyrimidine tract binding protein. *RNA* **11**:699-716.
8. **Amrein, H., M. Gorman, and R. Nothiger.** 1988. The sex-determining gene tra-2 of *Drosophila* encodes a putative RNA binding protein. *Cell* **55**:1025-1035.
9. **Anko, M. L., and K. M. Neugebauer.** 2010. Long noncoding RNAs add another layer to pre-mRNA splicing regulation. *Molecular cell* **39**:833-834.
10. **Arhel, N., and F. Kirchhoff.** 2010. Host proteins involved in HIV infection: new therapeutic targets. *Biochimica et biophysica acta* **1802**:313-321.

11. **Asang, C., I. Hauber, and H. Schaal.** 2008. Insights into the selective activation of alternatively used splice acceptors by the human immunodeficiency virus type-1 bidirectional splicing enhancer. *Nucleic acids research* **36**:1450-1463.
12. **Ashe, M. P., A. Furger, and N. J. Proudfoot.** 2000. Stem-loop 1 of the U1 snRNP plays a critical role in the suppression of HIV-1 polyadenylation. *RNA* **6**:170-177.
13. **Ashe, M. P., L. H. Pearson, and N. J. Proudfoot.** 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *The EMBO journal* **16**:5752-5763.
14. **Ausubel, F. M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., Struhl,K. and, and eds.** 1991. *Current protocols in molecular biology*. Greene Publishing Associates and Wiley-Interscience, New York.
15. **Auweter, S. D., and F. H. Allain.** 2008. Structure-function relationships of the polypyrimidine tract binding protein. *Cellular and molecular life sciences : CMLS* **65**:516-527.
16. **Auweter, S. D., R. Fasan, L. Reymond, J. G. Underwood, D. L. Black, S. Pitsch, and F. H. Allain.** 2006. Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *The EMBO journal* **25**:163-173.
17. **Azubel, M., N. Habib, R. Sperling, and J. Sperling.** 2006. Native spliceosomes assemble with pre-mRNA to form supraspliceosomes. *Journal of molecular biology* **356**:955-966.
18. **Bajenova, O., E. Stolper, S. Gapon, N. Sundina, R. Zimmer, and P. Thomas.** 2003. Surface expression of heterogeneous nuclear RNA binding protein M4 on Kupffer cell relates to its function as a carcinoembryonic antigen receptor. *Experimental cell research* **291**:228-241.
19. **Bannai, H., K. Fukatsu, A. Mizutani, T. Natsume, S. Iemura, T. Ikegami, T. Inoue, and K. Mikoshiba.** 2004. An RNA-interacting protein, SYNCRIP (heterogeneous nuclear ribonuclear protein Q1/NSAP1) is a component of mRNA granule transported with inositol 1,4,5-trisphosphate receptor type 1 mRNA in neuronal dendrites. *The Journal of biological chemistry* **279**:53427-53434.

-
20. **Baraniak, A. P., J. R. Chen, and M. A. Garcia-Blanco.** 2006. Fox-2 mediates epithelial cell-specific fibroblast growth factor receptor 2 exon choice. *Molecular and cellular biology* **26**:1209-1222.
 21. **Barash, Y., J. A. Calarco, W. Gao, Q. Pan, X. Wang, O. Shai, B. J. Blencowe, and B. J. Frey.** 2010. Deciphering the splicing code. *Nature* **465**:53-59.
 22. **Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**:868-871.
 23. **Bauer, W. J., J. Heath, J. L. Jenkins, and C. L. Kielkopf.** 2012. Three RNA recognition motifs participate in RNA recognition and structural organization by the pro-apoptotic factor TIA-1. *Journal of molecular biology* **415**:727-740.
 24. **Beck, A. R., Q. G. Medley, S. O'Brien, P. Anderson, and M. Streuli.** 1996. Structure, tissue distribution and genomic organization of the murine RRM-type RNA binding proteins TIA-1 and TIAR. *Nucleic acids research* **24**:3829-3835.
 25. **Bedard, K. M., S. Daijogo, and B. L. Semler.** 2007. A nucleo-cytoplasmic SR protein functions in viral IRES-mediated translation initiation. *The EMBO journal* **26**:459-467.
 26. **Bedford, M. T., R. Reed, and P. Leder.** 1998. WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proceedings of the National Academy of Sciences of the United States of America* **95**:10602-10607.
 27. **Bedford, M. T., D. Sarbassova, J. Xu, P. Leder, and M. B. Yaffe.** 2000. A novel pro-Arg motif recognized by WW domains. *The Journal of biological chemistry* **275**:10359-10369.
 28. **Berget, S. M.** 1995. Exon recognition in vertebrate splicing. *The Journal of biological chemistry* **270**:2411-2414.
 29. **Berglund, J. A., N. Abovich, and M. Rosbash.** 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes & development* **12**:858-867.

30. **Berglund, J. A., K. Chua, N. Abovich, R. Reed, and M. Rosbash.** 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89**:781-787.
31. **Bernard, D., K. V. Prasanth, V. Tripathi, S. Colasse, T. Nakamura, Z. Xuan, M. Q. Zhang, F. Sedel, L. Jourdain, F. Culpier, A. Triller, D. L. Spector, and A. Bessis.** 2010. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *The EMBO journal* **29**:3082-3093.
32. **Berro, R., K. Kehn, C. de la Fuente, A. Pumfery, R. Adair, J. Wade, A. M. Colberg-Poley, J. Hiscott, and F. Kashanchi.** 2006. Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. *Journal of virology* **80**:3189-3204.
33. **Berro, R., C. Pedati, K. Kehn-Hall, W. Wu, Z. Klase, Y. Even, A. M. Geneviere, T. Ammosova, S. Nekhai, and F. Kashanchi.** 2008. CDK13, a new potential human immunodeficiency virus type 1 inhibitory factor regulating viral mRNA splicing. *Journal of virology* **82**:7155-7166.
34. **Beyer, A. L., M. E. Christensen, B. W. Walker, and W. M. LeSturgeon.** 1977. Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell* **11**:127-138.
35. **Bilodeau, P. S., J. K. Domsic, A. Mayeda, A. R. Krainer, and C. M. Stoltzfus.** 2001. 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. *Journal of virology* **75**:8487-8497.
36. **Birney, E., S. Kumar, and A. R. Krainer.** 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic acids research* **21**:5803-5816.
37. **Black, D. L.** 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annual review of biochemistry* **72**:291-336.
38. **Blanchette, M., and B. Chabot.** 1999. Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *The EMBO journal* **18**:1939-1952.
39. **Blaustein, M., F. Pelisch, O. A. Coso, M. J. Bissell, A. R. Kornblihtt, and A. Srebrow.** 2004. Mammary epithelial-mesenchymal interaction regulates

- fibronectin alternative splicing via phosphatidylinositol 3-kinase. The Journal of biological chemistry **279**:21029-21037.
40. **Blencowe, B. J.** 2006. Alternative splicing: new insights from global analyses. Cell **126**:37-47.
41. **Blencowe, B. J.** 2000. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends in biochemical sciences **25**:106-110.
42. **Blencowe, B. J., G. Bauren, A. G. Eldridge, R. Issner, J. A. Nickerson, E. Rosonina, and P. A. Sharp.** 2000. The SRm160/300 splicing coactivator subunits. RNA **6**:111-120.
43. **Blencowe, B. J., J. A. Bowman, S. McCracken, and E. Rosonina.** 1999. SR-related proteins and the processing of messenger RNA precursors. Biochemistry and cell biology = Biochimie et biologie cellulaire **77**:277-291.
44. **Boggs, R. T., P. Gregor, S. Idriss, J. M. Belote, and M. McKeown.** 1987. Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene. Cell **50**:739-747.
45. **Bonnal, S., C. Martinez, P. Forch, A. Bachi, M. Wilm, and J. Valcarcel.** 2008. RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. Molecular cell **32**:81-95.
46. **Bono, F., and N. H. Gehring.** 2011. Assembly, disassembly and recycling: the dynamics of exon junction complexes. RNA biology **8**:24-29.
47. **Boris-Lawrie, K., T. M. Roberts, and S. Hull.** 2001. Retroviral RNA elements integrate components of post-transcriptional gene expression. Life sciences **69**:2697-2709.
48. **Boucher, L., C. A. Ouzounis, A. J. Enright, and B. J. Blencowe.** 2001. A genome-wide survey of RS domain proteins. RNA **7**:1693-1701.
49. **Boukis, L. A., N. Liu, S. Furuyama, and J. P. Bruzik.** 2004. Ser/Arg-rich protein-mediated communication between U1 and U2 small nuclear ribonucleoprotein particles. The Journal of biological chemistry **279**:29647-29653.
50. **Bringmann, P., and R. Luhrmann.** 1986. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. The EMBO journal **5**:3509-3516.

51. **Brow, D. A.** 2002. Allosteric cascade of spliceosome activation. Annual review of genetics **36**:333-360.
52. **Buckanovich, R. J., and R. B. Darnell.** 1997. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. Molecular and cellular biology **17**:3194-3201.
53. **Buckanovich, R. J., J. B. Posner, and R. B. Darnell.** 1993. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. Neuron **11**:657-672.
54. **Buckanovich, R. J., Y. Y. Yang, and R. B. Darnell.** 1996. The onconeural antigen Nova-1 is a neuron-specific RNA-binding protein, the activity of which is inhibited by paraneoplastic antibodies. The Journal of neuroscience : the official journal of the Society for Neuroscience **16**:1114-1122.
55. **Bukrinsky, M.** 2006. SNFing HIV transcription. Retrovirology **3**:49.
56. **Buratti, E., and F. E. Baralle.** 2008. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. Frontiers in bioscience : a journal and virtual library **13**:867-878.
57. **Buratti, E., M. Baralle, L. De Conti, D. Baralle, M. Romano, Y. M. Ayala, and F. E. Baralle.** 2004. hnRNP H binding at the 5' splice site correlates with the pathological effect of two intronic mutations in the NF-1 and TSHbeta genes. Nucleic acids research **32**:4224-4236.
58. **Buratti, E., C. Stuani, G. De Prato, and F. E. Baralle.** 2007. SR protein-mediated inhibition of CFTR exon 9 inclusion: molecular characterization of the intronic splicing silencer. Nucleic acids research **35**:4359-4368.
59. **Burge, C. B., Tuschl, T. and Sharp, P. A.** 1999. Splicing of Precursors to mRNAs by the Spliceosomes. In Gesteland, R. F., Cech, T. R. and Atkins, J. F. (eds.), Cold Spring Harbor Laboratory Press, New York:525-560.
60. **Burgess, S., J. R. Couto, and C. Guthrie.** 1990. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. Cell **60**:705-717.
61. **Burgess, S. M., and C. Guthrie.** 1993. Beat the clock: paradigms for NTPases in the maintenance of biological fidelity. Trends in biochemical sciences **18**:381-384.

-
62. **Burnette, J. M., E. Miyamoto-Sato, M. A. Schaub, J. Conklin, and A. J. Lopez.** 2005. Subdivision of large introns in *Drosophila* by recursive splicing at nonexonic elements. *Genetics* **170**:661-674.
 63. **Buzon, M. J., M. Massanella, J. M. Llibre, A. Esteve, V. Dahl, M. C. Puertas, J. M. Gatell, P. Domingo, R. Paredes, M. Sharkey, S. Palmer, M. Stevenson, B. Clotet, J. Blanco, and J. Martinez-Picado.** 2010. HIV-1 replication and immune dynamics are affected by raltegravir intensification of HAART-suppressed subjects. *Nature medicine* **16**:460-465.
 64. **Caceres, J. F., and A. R. Kornblihtt.** 2002. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends in genetics : TIG* **18**:186-193.
 65. **Caceres, J. F., and A. R. Krainer.** 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *The EMBO journal* **12**:4715-4726.
 66. **Caceres, J. F., T. Misteli, G. R. Sreaton, D. L. Spector, and A. R. Krainer.** 1997. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *The Journal of cell biology* **138**:225-238.
 67. **Caceres, J. F., G. R. Sreaton, and A. R. Krainer.** 1998. A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes & development* **12**:55-66.
 68. **Caceres, J. F., S. Stamm, D. M. Helfman, and A. R. Krainer.** 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* **265**:1706-1709.
 69. **Camats, M., S. Guil, M. Kokolo, and M. Bach-Elias.** 2008. P68 RNA helicase (DDX5) alters activity of cis- and trans-acting factors of the alternative splicing of H-Ras. *PloS one* **3**:e2926.
 70. **Cao, W., and M. A. Garcia-Blanco.** 1998. A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. *The Journal of biological chemistry* **273**:20629-20635.
 71. **Cao, W., S. F. Jamison, and M. A. Garcia-Blanco.** 1997. Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA* **3**:1456-1467.
 72. **Caputi, M., M. Freund, S. Kammler, C. Asang, and H. Schaal.** 2004. A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates

- human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *Journal of virology* **78**:6517-6526.
73. **Caputi, M., A. Mayeda, A. R. Krainer, and A. M. Zahler.** 1999. hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *The EMBO journal* **18**:4060-4067.
74. **Caputi, M., and A. M. Zahler.** 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *The Journal of biological chemistry* **276**:43850-43859.
75. **Cartegni, L., S. L. Chew, and A. R. Krainer.** 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature reviews. Genetics* **3**:285-298.
76. **Cartegni, L., J. Wang, Z. Zhu, M. Q. Zhang, and A. R. Krainer.** 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic acids research* **31**:3568-3571.
77. **Castle, J. C., C. Zhang, J. K. Shah, A. V. Kulkarni, A. Kalsotra, T. A. Cooper, and J. M. Johnson.** 2008. Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. *Nature genetics* **40**:1416-1425.
78. **Cattaruzza, M., K. Schafer, and M. Hecker.** 2002. Cytokine-induced down-regulation of zfm1/splicing factor-1 promotes smooth muscle cell proliferation. *The Journal of biological chemistry* **277**:6582-6589.
79. **Cavaloc, Y., C. F. Bourgeois, L. Kister, and J. Stevenin.** 1999. The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. *RNA* **5**:468-483.
80. **Chabot, B., M. Blanchette, I. Lapierre, and H. La Branche.** 1997. An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Molecular and cellular biology* **17**:1776-1786.
81. **Chabot, B., C. LeBel, S. Hutchison, F. H. Nasim, and M. J. Simard.** 2003. Heterogeneous nuclear ribonucleoprotein particle A/B proteins and the control of alternative splicing of the mammalian heterogeneous nuclear ribonucleoprotein particle A1 pre-mRNA. *Progress in molecular and subcellular biology* **31**:59-88.
82. **Chan, D. C., D. Fass, J. M. Berger, and P. S. Kim.** 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**:263-273.

83. **Chan, R. C., and D. L. Black.** 1997. Conserved intron elements repress splicing of a neuron-specific c-src exon in vitro. *Molecular and cellular biology* **17**:2970.
84. **Chan, R. C., and D. L. Black.** 1995. Conserved intron elements repress splicing of a neuron-specific c-src exon in vitro. *Molecular and cellular biology* **15**:6377-6385.
85. **Chang, Y. F., J. S. Imam, and M. F. Wilkinson.** 2007. The nonsense-mediated decay RNA surveillance pathway. *Annual review of biochemistry* **76**:51-74.
86. **Chasin, L. A.** 2007. Searching for splicing motifs. *Advances in experimental medicine and biology* **623**:85-106.
87. **Chaudhury, A., P. Chander, and P. H. Howe.** 2010. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. *RNA* **16**:1449-1462.
88. **Chen, J. Y., L. Stands, J. P. Staley, R. R. Jackups, Jr., L. J. Latus, and T. H. Chang.** 2001. Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. *Molecular cell* **7**:227-232.
89. **Chen, Y., X. Zhou, N. Liu, C. Wang, L. Zhang, W. Mo, and G. Hu.** 2008. Arginine methylation of hnRNP K enhances p53 transcriptional activity. *FEBS letters* **582**:1761-1765.
90. **Cho, S., A. Hoang, S. Chakrabarti, N. Huynh, D. B. Huang, and G. Ghosh.** 2011. The SRSF1 linker induces semi-conservative ESE binding by cooperating with the RRM. *Nucleic acids research* **39**:9413-9421.
91. **Cho, S., A. Hoang, R. Sinha, X. Y. Zhong, X. D. Fu, A. R. Krainer, and G. Ghosh.** 2011. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proceedings of the National Academy of Sciences of the United States of America* **108**:8233-8238.
92. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical biochemistry* **162**:156-159.

-
93. **Chomczynski, P., and N. Sacchi.** 2006. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature protocols* **1**:581-585.
 94. **Chou, M. Y., N. Rooke, C. W. Turck, and D. L. Black.** 1999. hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Molecular and cellular biology* **19**:69-77.
 95. **Chou, T. B., Z. Zachar, and P. M. Bingham.** 1987. Developmental expression of a regulatory gene is programmed at the level of splicing. *The EMBO journal* **6**:4095-4104.
 96. **Chun, T. W., D. Engel, M. M. Berrey, T. Shea, L. Corey, and A. S. Fauci.** 1998. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America* **95**:8869-8873.
 97. **Chun, T. W., D. Engel, S. B. Mizell, L. A. Ehler, and A. S. Fauci.** 1998. Induction of HIV-1 replication in latently infected CD4+ T cells using a combination of cytokines. *The Journal of experimental medicine* **188**:83-91.
 98. **Chun, T. W., D. Finzi, J. Margolick, K. Chadwick, D. Schwartz, and R. F. Siliciano.** 1995. In vivo fate of HIV-1-infected T cells: quantitative analysis of the transition to stable latency. *Nature medicine* **1**:1284-1290.
 99. **Clavel, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, and et al.** 1986. Isolation of a new human retrovirus from West African patients with AIDS. *Science* **233**:343-346.
 100. **Clerte, C., and K. B. Hall.** 2009. The domains of polypyrimidine tract binding protein have distinct RNA structural preferences. *Biochemistry* **48**:2063-2074.
 101. **Clery, A., M. Blatter, and F. H. Allain.** 2008. RNA recognition motifs: boring? Not quite. *Current opinion in structural biology* **18**:290-298.
 102. **Cochrane, A. W., C. H. Chen, and C. A. Rosen.** 1990. Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **87**:1198-1202.
 103. **Collins, C. A., and C. Guthrie.** 2001. Genetic interactions between the 5' and 3' splice site consensus sequences and U6 snRNA during the second catalytic step of pre-mRNA splicing. *RNA* **7**:1845-1854.

-
104. **Colwill, K., L. L. Feng, J. M. Yeakley, G. D. Gish, J. F. Caceres, T. Pawson, and X. D. Fu.** 1996. SRPK1 and Clk/Sty protein kinases show distinct substrate specificities for serine/arginine-rich splicing factors. *The Journal of biological chemistry* **271**:24569-24575.
 105. **Conte, M. R., T. Grune, J. Ghuman, G. Kelly, A. Ladas, S. Matthews, and S. Curry.** 2000. Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *The EMBO journal* **19**:3132-3141.
 106. **Coolidge, C. J., R. J. Seely, and J. G. Patton.** 1997. Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic acids research* **25**:888-896.
 107. **Corvelo, A., M. Hallegger, C. W. Smith, and E. Eyras.** 2010. Genome-wide association between branch point properties and alternative splicing. *PLoS computational biology* **6**:e1001016.
 108. **Coulter, L. R., M. A. Landree, and T. A. Cooper.** 1997. Identification of a new class of exonic splicing enhancers by in vivo selection. *Molecular and cellular biology* **17**:2143-2150.
 109. **Cullen, B. R.** 2003. Nuclear mRNA export: insights from virology. *Trends in biochemical sciences* **28**:419-424.
 110. **Cullen, B. R.** 2003. Nuclear RNA export. *Journal of cell science* **116**:587-597.
 111. **Cullen, B. R.** 1991. Regulation of human immunodeficiency virus replication. *Annual review of microbiology* **45**:219-250.
 112. **Cummins, N. W., and A. D. Badley.** 2010. Mechanisms of HIV-associated lymphocyte apoptosis: 2010. *Cell death & disease* **1**:e99.
 113. **Daeﬂer, S., M. E. Klotman, and F. Wong-Staal.** 1990. Trans-activating rev protein of the human immunodeficiency virus 1 interacts directly and specifically with its target RNA. *Proceedings of the National Academy of Sciences of the United States of America* **87**:4571-4575.
 114. **Dalgleish, A. G., P. C. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, and R. A. Weiss.** 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**:763-767.
 115. **Daly, T. J., K. S. Cook, G. S. Gray, T. E. Maione, and J. R. Rusche.** 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element in vitro. *Nature* **342**:816-819.

-
116. **Damgaard, C. K., T. O. Tange, and J. Kjems.** 2002. 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**:1401-1415.
 117. **Damianov, A., and D. L. Black.** 2010. Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA* **16**:405-416.
 118. **Damier, L., L. Domenjoud, and C. Branlant.** 1997. 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. *Biochemical and biophysical research communications* **237**:182-187.
 119. **Das, S. R., and S. Jameel.** 2005. Biology of the HIV Nef protein. *The Indian journal of medical research* **121**:315-332.
 120. **Dauksaite, V., and G. Akusjarvi.** 2002. Human splicing factor ASF/SF2 encodes for a repressor domain required for its inhibitory activity on pre-mRNA splicing. *The Journal of biological chemistry* **277**:12579-12586.
 121. **Del Gatto-Konczak, F., C. F. Bourgeois, C. Le Guiner, L. Kister, M. C. Gesnel, J. Stevenin, and R. Breathnach.** 2000. The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5' splice site. *Molecular and cellular biology* **20**:6287-6299.
 122. **Del Gatto-Konczak, F., M. Olive, M. C. Gesnel, and R. Breathnach.** 1999. hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. *Molecular and cellular biology* **19**:251-260.
 123. **Dember, L. M., N. D. Kim, K. Q. Liu, and P. Anderson.** 1996. Individual RNA recognition motifs of TIA-1 and TIAR have different RNA binding specificities. *The Journal of biological chemistry* **271**:2783-2788.
 124. **Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. Di Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau.** 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**:661-666.
 125. **Ding, J. H., X. Xu, D. Yang, P. H. Chu, N. D. Dalton, Z. Ye, J. M. Yeakley, H. Cheng, R. P. Xiao, J. Ross, J. Chen, and X. D. Fu.** 2004. Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. *The EMBO journal* **23**:885-896.

-
126. **Ding, J. H., X. Y. Zhong, J. C. Hagopian, M. M. Cruz, G. Ghosh, J. Feramisco, J. A. Adams, and X. D. Fu.** 2006. Regulated cellular partitioning of SR protein-specific kinases in mammalian cells. *Molecular biology of the cell* **17**:876-885.
 127. **Doitsh, G., M. Cavrois, K. G. Lassen, O. Zepeda, Z. Yang, M. L. Santiago, A. M. Hebbeler, and W. C. Greene.** 2010. Abortive HIV infection mediates CD4 T cell depletion and inflammation in human lymphoid tissue. *Cell* **143**:789-801.
 128. **Doktor, T. K., L. D. Schroeder, A. Vested, J. Palmfeldt, H. S. Andersen, N. Gregersen, and B. S. Andresen.** 2011. SMN2 exon 7 splicing is inhibited by binding of hnRNP A1 to a common ESS motif that spans the 3' splice site. *Human mutation* **32**:220-230.
 129. **Dominguez, C., and F. H. Allain.** 2006. NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic acids research* **34**:3634-3645.
 130. **Dominguez, C., J. F. Fisette, B. Chabot, and F. H. Allain.** 2010. Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nature structural & molecular biology* **17**:853-861.
 131. **Domsic, J. K., Y. Wang, A. Mayeda, A. R. Krainer, and C. M. Stoltzfus.** 2003. Human immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV antagonizes binding of U2AF65 to viral polypyrimidine tracts. *Molecular and cellular biology* **23**:8762-8772.
 132. **Donmez, G., K. Hartmuth, B. Kastner, C. L. Will, and R. Luhrmann.** 2007. The 5' end of U2 snRNA is in close proximity to U1 and functional sites of the pre-mRNA in early spliceosomal complexes. *Molecular cell* **25**:399-411.
 133. **Douek, D. C., L. J. Picker, and R. A. Koup.** 2003. T cell dynamics in HIV-1 infection. *Annual review of immunology* **21**:265-304.
 134. **Douglas, A. G., and M. J. Wood.** 2011. RNA splicing: disease and therapy. *Briefings in functional genomics* **10**:151-164.
 135. **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.** 2008. HIV-1 infection induces changes in expression of cellular splicing factors that regulate

- alternative viral splicing and virus production in macrophages. *Retrovirology* **5**:18.
136. **Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton.** 1996. HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**:667-673.
 137. **Dreyfuss, G., V. N. Kim, and N. Kataoka.** 2002. Messenger-RNA-binding proteins and the messages they carry. *Nature reviews. Molecular cell biology* **3**:195-205.
 138. **Dreyfuss, G., M. J. Matunis, S. Pinol-Roma, and C. G. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. *Annual review of biochemistry* **62**:289-321.
 139. **Dreyfuss, G., M. S. Swanson, and S. Pinol-Roma.** 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends in biochemical sciences* **13**:86-91.
 140. **Dürr, S.** Heinrich-Heine-Universität Düsseldorf, 2007, diploma thesis. Sequenzvoraussetzungen für die Erkennung von viralen und humanen Exons mit schwachen 5' Spleißstellen.
 141. **Dyhr-Mikkelsen, H., and J. Kjems.** 1995. Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. *The Journal of biological chemistry* **270**:24060-24066.
 142. **Emerman, M., R. Vazeux, and K. Peden.** 1989. The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. *Cell* **57**:1155-1165.
 143. **Eperon, I. C., D. C. Ireland, R. A. Smith, A. Mayeda, and A. R. Krainer.** 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *The EMBO journal* **12**:3607-3617.
 144. **Eperon, I. C., O. V. Makarova, A. Mayeda, S. H. Munroe, J. F. Cáceres, D. G. Hayward, and A. R. Krainer.** 2000. Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Molecular and cellular biology* **20**:8303-8318.
 145. **Erkelenz, S., W. F. Mueller, M. S. Evans, A. Busch, K. Schöneweis, K. J. Hertel, and H. Schaal.** 2012. Position-dependent splicing activation and

- p repression by SR and hnRNP proteins rely on common mechanisms. Cell Reports
- submitted**
- .
-
- 146.
- Exline, C. M., Z. Feng, and C. M. Stoltzfus.**
2008. Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. Journal of virology
- 82**
- :3921-3931.
-
- 147.
- Fairbrother, W. G., R. F. Yeh, P. A. Sharp, and C. B. Burge.**
2002. Predictive identification of exonic splicing enhancers in human genes. Science
- 297**
- :1007-1013.
-
- 148.
- Fairbrother, W. G., R. F. Yeh, P. A. Sharp, and C. B. Burge.**
2002. Predictive identification of exonic splicing enhancers in human genes. Science
- 297**
- :1007-1013.
-
- 149.
- Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis.**
1989. rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. Proceedings of the National Academy of Sciences of the United States of America
- 86**
- :1495-1499.
-
- 150.
- Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger.**
1996. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. Science
- 272**
- :872-877.
-
- 151.
- Fic, W., F. Juge, J. Soret, and J. Tazi.**
2007. Eye development under the control of SRp55/B52-mediated alternative splicing of eyeless. PloS one
- 2**
- :e253.
-
- 152.
- Fischer, D. C., K. Noack, I. B. Runnebaum, D. O. Watermann, D. G. Kieback, S. Stamm, and E. Stickeler.**
2004. Expression of splicing factors in human ovarian cancer. Oncology reports
- 11**
- :1085-1090.
-
- 153.
- Fisette, J. F., J. Toutant, S. Dugre-Brisson, L. Desgroseillers, and B. Chabot.**
2010. hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. RNA
- 16**
- :228-238.
-
- 154.
- Fleckner, J., M. Zhang, J. Valcarcel, and M. R. Green.**
1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes & development
- 11**
- :1864-1872.
-
- 155.
- Fogel, B. L., and M. T. McNally.**
2000. A cellular protein, hnRNP H, binds to the negative regulator of splicing element from Rous sarcoma virus. The Journal of biological chemistry
- 275**
- :32371-32378.

-
156. **Forch, P., O. Puig, N. Kedersha, C. Martinez, S. Granneman, B. Seraphin, P. Anderson, and J. Valcarcel.** 2000. The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing. *Molecular cell* **6**:1089-1098.
 157. **Forch, P., O. Puig, C. Martinez, B. Seraphin, and J. Valcarcel.** 2002. The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites. *The EMBO journal* **21**:6882-6892.
 158. **Fortes, P., Y. Cuevas, F. Guan, P. Liu, S. Pentlicky, S. P. Jung, M. L. Martinez-Chantar, J. Prieto, D. Rowe, and S. I. Gunderson.** 2003. Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small nuclear RNAs targeted to terminal exons of pre-mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **100**:8264-8269.
 159. **Fox-Walsh, K. L., Y. Dou, B. J. Lam, S. P. Hung, P. F. Baldi, and K. J. Hertel.** 2005. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America* **102**:16176-16181.
 160. **Fox-Walsh, K. L., and K. J. Hertel.** 2009. Splice-site pairing is an intrinsically high fidelity process. *Proceedings of the National Academy of Sciences of the United States of America* **106**:1766-1771.
 161. **Freund, M.** Heinrich-Heine-Universität Düsseldorf, 2004, PhD thesis. Die Funktion des U1 snRNPs in der HIV-1 env-Expression.
 162. **Freund, M., C. Asang, S. Kammler, C. Konermann, J. Krummheuer, M. Hipp, I. Meyer, W. Gierling, S. Theiss, T. Preuss, D. Schindler, J. Kjems, and H. Schaal.** 2003. A novel approach to describe a U1 snRNA binding site. *Nucleic acids research* **31**:6963-6975.
 163. **Fu, X. D.** 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**:663-680.
 164. **Fu, X. D., and T. Maniatis.** 1992. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proceedings of the National Academy of Sciences of the United States of America* **89**:1725-1729.
 165. **Fu, Y., A. Masuda, M. Ito, J. Shinmi, and K. Ohno.** 2011. AG-dependent 3'-splice sites are predisposed to aberrant splicing due to a mutation at the first nucleotide of an exon. *Nucleic acids research* **39**:4396-4404.

-
166. **Fukumura, K., A. Kato, Y. Jin, T. Ideue, T. Hirose, N. Kataoka, T. Fujiwara, H. Sakamoto, and K. Inoue.** 2007. Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1gamma gene. *Nucleic acids research* **35**:5303-5311.
167. **Furtado, M. R., R. Balachandran, P. Gupta, and S. M. Wolinsky.** 1991. Analysis of alternatively spliced human immunodeficiency virus type-1 mRNA species, one of which encodes a novel tat-env fusion protein. *Virology* **185**:258-270.
168. **Gallo, R. C., S. Z. Salahuddin, M. Popovic, G. M. Shearer, M. Kaplan, B. F. Haynes, T. J. Parker, R. Redfield, J. Oleske, B. Safai, and et al.** 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**:500-503.
169. **Gao, K., A. Masuda, T. Matsuura, and K. Ohno.** 2008. Human branch point consensus sequence is yUnAy. *Nucleic acids research* **36**:2257-2267.
170. **Garneau, D., T. Revil, J. F. Fisette, and B. Chabot.** 2005. Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x. *The Journal of biological chemistry* **280**:22641-22650.
171. **Ge, H., and J. L. Manley.** 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**:25-34.
172. **Geijtenbeek, T. B., D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. van Duijnhoven, J. Middel, I. L. Cornelissen, H. S. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, and Y. van Kooyk.** 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **100**:587-597.
173. **Gesnel, M. C., S. Theoleyre, F. Del Gatto-Konczak, and R. Breathnach.** 2007. Cooperative binding of TIA-1 and U1 snRNP in K-SAM exon splicing activation. *Biochemical and biophysical research communications* **358**:1065-1070.
174. **Ghetti, A., S. Pinol-Roma, W. M. Michael, C. Morandi, and G. Dreyfuss.** 1992. hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic acids research* **20**:3671-3678.

175. **Gil, A., P. A. Sharp, S. F. Jamison, and M. A. Garcia-Blanco.** 1991. Characterization of cDNAs encoding the polypyrimidine tract-binding protein. *Genes & development* **5**:1224-1236.
176. **Goina, E., N. Skoko, and F. Pagani.** 2008. Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. *Molecular and cellular biology* **28**:3850-3860.
177. **Golling, G., A. Amsterdam, Z. Sun, M. Antonelli, E. Maldonado, W. Chen, S. Burgess, M. Haldi, K. Artzt, S. Farrington, S. Y. Lin, R. M. Nissen, and N. Hopkins.** 2002. Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nature genetics* **31**:135-140.
178. **Gooding, C., F. Clark, M. C. Wollerton, S. N. Grellscheid, H. Groom, and C. W. Smith.** 2006. A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones. *Genome biology* **7**:R1.
179. **Gooding, C., G. C. Roberts, and C. W. Smith.** 1998. Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon. *RNA* **4**:85-100.
180. **Gorlach, M., C. G. Burd, D. S. Portman, and G. Dreyfuss.** 1993. The hnRNP proteins. *Molecular biology reports* **18**:73-78.
181. **Gottlieb, M. S., R. Schroff, H. M. Schanker, J. D. Weisman, P. T. Fan, R. A. Wolf, and A. Saxon.** 1981. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *The New England journal of medicine* **305**:1425-1431.
182. **Gugeon, M. L., and M. Piacentini.** 2009. New insights on the role of apoptosis and autophagy in HIV pathogenesis. *Apoptosis : an international journal on programmed cell death* **14**:501-508.
183. **Gozani, O., R. Feld, and R. Reed.** 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes & development* **10**:233-243.

184. **Gozani, O., J. Potashkin, and R. Reed.** 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Molecular and cellular biology* **18**:4752-4760.
185. **Graveley, B. R.** 2000. Sorting out the complexity of SR protein functions. *RNA* **6**:1197-1211.
186. **Graveley, B. R., K. J. Hertel, and T. Maniatis.** 1998. A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *The EMBO journal* **17**:6747-6756.
187. **Graveley, B. R., and T. Maniatis.** 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Molecular cell* **1**:765-771.
188. **Gross, T., K. Richert, C. Mierke, M. Lutzberger, and N. F. Kaufer.** 1998. Identification and characterization of *srp1*, a gene of fission yeast encoding a RNA binding domain and a RS domain typical of SR splicing factors. *Nucleic acids research* **26**:505-511.
189. **Grossman, J. S., M. I. Meyer, Y. C. Wang, G. J. Mulligan, R. Kobayashi, and D. M. Helfman.** 1998. The use of antibodies to the polypyrimidine tract binding protein (PTB) to analyze the protein components that assemble on alternatively spliced pre-mRNAs that use distant branch points. *RNA* **4**:613-625.
190. **Guatelli, J. C., T. R. Gingeras, and D. D. Richman.** 1990. Alternative splice acceptor utilization during human immunodeficiency virus type 1 infection of cultured cells. *Journal of virology* **64**:4093-4098.
191. **Gui, J. F., H. Tronchere, S. D. Chandler, and X. D. Fu.** 1994. Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. *Proceedings of the National Academy of Sciences of the United States of America* **91**:10824-10828.
192. **Guth, S., C. Martinez, R. K. Gaur, and J. Valcarcel.** 1999. Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). *Molecular and cellular biology* **19**:8263-8271.
193. **Guth, S., T. O. Tange, E. Kellenberger, and J. Valcarcel.** 2001. Dual function for U2AF(35) in AG-dependent pre-mRNA splicing. *Molecular and cellular biology* **21**:7673-7681.

-
194. **Hallay, H., N. Locker, L. Ayadi, D. Ropers, E. Guittet, and C. Branlant.** 2006. Biochemical and NMR study on the competition between proteins SC35, SRp40, and heterogeneous nuclear ribonucleoprotein A1 at the HIV-1 Tat exon 2 splicing site. *The Journal of biological chemistry* **281**:37159-37174.
195. **Hallegger, M., A. Sobala, and C. W. Smith.** 2010. Four exons of the serotonin receptor 4 gene are associated with multiple distant branch points. *RNA* **16**:839-851.
196. **Han, K., G. Yeo, P. An, C. B. Burge, and P. J. Grabowski.** 2005. A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS biology* **3**:e158.
197. **Han, S. P., Y. H. Tang, and R. Smith.** 2010. Functional diversity of the hnRNPs: past, present and perspectives. *The Biochemical journal* **430**:379-392.
198. **Hanamura, A., J. F. Cáceres, A. Mayeda, B. R. Franza, Jr., and A. R. Krainer.** 1998. Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *RNA* **4**:430-444.
199. **Hargous, Y., G. M. Hautbergue, A. M. Tintaru, L. Skrisovska, A. P. Golovanov, J. Stevenin, L. Y. Lian, S. A. Wilson, and F. H. Allain.** 2006. Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *The EMBO journal* **25**:5126-5137.
200. **Harris, M. E., and T. J. Hope.** 2000. RNA export: insights from viral models. *Essays in biochemistry* **36**:115-127.
201. **Hartmann, L.** Heinrich-Heine-Universität Düsseldorf, 2005, diploma thesis. Charakterisierung funktionaler Sequenzelemente der HIV-1 Spleißakzeptoren.
202. **Hartmann, L., S. Theiss, D. Niederacher, and H. Schaal.** 2008. Diagnostics of pathogenic splicing mutations: does bioinformatics cover all bases? *Frontiers in bioscience : a journal and virtual library* **13**:3252-3272.
203. **Hastings, M. L., E. Allemand, D. M. Duelli, M. P. Myers, and A. R. Krainer.** 2007. Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65). *PloS one* **2**:e538.
204. **Hastings, M. L., C. M. Wilson, and S. H. Munroe.** 2001. A purine-rich intronic element enhances alternative splicing of thyroid hormone receptor mRNA. *RNA* **7**:859-874.

-
205. **Hautbergue, G. M., M. L. Hung, A. P. Golovanov, L. Y. Lian, and S. A. Wilson.** 2008. Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proceedings of the National Academy of Sciences of the United States of America* **105**:5154-5159.
 206. **Heaphy, S., C. Dingwall, I. Ernberg, M. J. Gait, S. M. Green, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. *Cell* **60**:685-693.
 207. **Hedley, M. L., H. Amrein, and T. Maniatis.** 1995. An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proceedings of the National Academy of Sciences of the United States of America* **92**:11524-11528.
 208. **Heinrichs, V., M. Bach, and R. Luhrmann.** 1990. U1-specific protein C is required for efficient complex formation of U1 snRNP with a 5' splice site. *Molecular biology reports* **14**:165.
 209. **Hernandez, H., O. V. Makarova, E. M. Makarov, N. Morgner, Y. Muto, D. P. Krummel, and C. V. Robinson.** 2009. Isoforms of U1-70k control subunit dynamics in the human spliceosomal U1 snRNP. *PloS one* **4**:e7202.
 210. **Herrmann, F., M. Bossert, A. Schwander, E. Akgun, and F. O. Fackelmayer.** 2004. Arginine methylation of scaffold attachment factor A by heterogeneous nuclear ribonucleoprotein particle-associated PRMT1. *The Journal of biological chemistry* **279**:48774-48779.
 211. **Hertel, K. J., and T. Maniatis.** 1998. The function of multisite splicing enhancers. *Molecular cell* **1**:449-455.
 212. **Hertel, K. J., and T. Maniatis.** 1999. Serine-arginine (SR)-rich splicing factors have an exon-independent function in pre-mRNA splicing. *Proceedings of the National Academy of Sciences of the United States of America* **96**:2651-2655.
 213. **Hicks, M. J., W. F. Mueller, P. J. Shepard, and K. J. Hertel.** 2010. Competing upstream 5' splice sites enhance the rate of proximal splicing. *Molecular and cellular biology* **30**:1878-1886.
 214. **Hilliker, A. K., M. A. Mefford, and J. P. Staley.** 2007. U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. *Genes & development* **21**:821-834.

-
215. **Hoffman, B. E., and P. J. Grabowski.** 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes & development* **6**:2554-2568.
216. **Hoffman, B. E., and J. T. Lis.** 2000. Pre-mRNA splicing by the essential *Drosophila* protein B52: tissue and target specificity. *Molecular and cellular biology* **20**:181-186.
217. **Hoffman, D. W., C. C. Query, B. L. Golden, S. W. White, and J. D. Keene.** 1991. RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proceedings of the National Academy of Sciences of the United States of America* **88**:2495-2499.
218. **Hope, T. J.** 1999. The ins and outs of HIV Rev. *Archives of biochemistry and biophysics* **365**:186-191.
219. **House, A. E., and K. W. Lynch.** 2006. An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition. *Nature structural & molecular biology* **13**:937-944.
220. **Huang, M., J. E. Rech, S. J. Northington, P. F. Flicker, A. Mayeda, A. R. Krainer, and W. M. LeSturgeon.** 1994. The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Molecular and cellular biology* **14**:518-533.
221. **Huang, T., J. Vilardell, and C. C. Query.** 2002. Pre-spliceosome formation in *S.pombe* requires a stable complex of SF1-U2AF(59)-U2AF(23). *The EMBO journal* **21**:5516-5526.
222. **Huang, X., M. Beullens, J. Zhang, Y. Zhou, E. Nicolaescu, B. Lesage, Q. Hu, J. Wu, M. Bollen, and Y. Shi.** 2009. Structure and function of the two tandem WW domains of the pre-mRNA splicing factor FBP21 (formin-binding protein 21). *The Journal of biological chemistry* **284**:25375-25387.
223. **Huang, Y., R. Gattoni, J. Stevenin, and J. A. Steitz.** 2003. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Molecular cell* **11**:837-843.
224. **Huang, Y., and J. A. Steitz.** 2001. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular cell* **7**:899-905.

225. **Huang, Y., and J. A. Steitz.** 2005. SRprises along a messenger's journey. *Molecular cell* **17**:613-615.
226. **Huang, Y., T. A. Yario, and J. A. Steitz.** 2004. A molecular link between SR protein dephosphorylation and mRNA export. *Proceedings of the National Academy of Sciences of the United States of America* **101**:9666-9670.
227. **Huh, G. S., and R. O. Hynes.** 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes & development* **8**:1561-1574.
228. **Huranova, M., I. Ivani, A. Benda, I. Poser, Y. Brody, M. Hof, Y. Shav-Tal, K. M. Neugebauer, and D. Stanek.** 2010. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *The Journal of cell biology* **191**:75-86.
229. **Ibrahim, E. C., T. D. Schaal, K. J. Hertel, R. Reed, and T. Maniatis.** 2005. Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. *Proceedings of the National Academy of Sciences of the United States of America* **102**:5002-5007.
230. **Isken, O., and L. E. Maquat.** 2007. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes & development* **21**:1833-1856.
231. **Ito, T., Y. Muto, M. R. Green, and S. Yokoyama.** 1999. Solution structures of the first and second RNA-binding domains of human U2 small nuclear ribonucleoprotein particle auxiliary factor (U2AF(65)). *The EMBO journal* **18**:4523-4534.
232. **Izquierdo, J. M., N. Majos, S. Bonnal, C. Martinez, R. Castelo, R. Guigo, D. Bilbao, and J. Valcarcel.** 2005. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Molecular cell* **19**:475-484.
233. **Izquierdo, J. M., and J. Valcarcel.** 2007. Two isoforms of the T-cell intracellular antigen 1 (TIA-1) splicing factor display distinct splicing regulation activities. Control of TIA-1 isoform ratio by TIA-1-related protein. *The Journal of biological chemistry* **282**:19410-19417.
234. **Jablonski, J. A., A. L. Amelio, M. Giacca, and M. Caputi.** 2010. The transcriptional transactivator Tat selectively regulates viral splicing. *Nucleic acids research* **38**:1249-1260.

-
235. **Jablonski, J. A., E. Buratti, C. Stuani, and M. Caputi.** 2008. The secondary structure of the human immunodeficiency virus type 1 transcript modulates viral splicing and infectivity. *Journal of virology* **82**:8038-8050.
236. **Jablonski, J. A., and M. Caputi.** 2009. Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity. *Journal of virology* **83**:981-992.
237. **Jacquenot, S., D. Decimo, D. Muriaux, and J. L. Darlix.** 2005. Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. *Retrovirology* **2**:33.
238. **Jacquenot, S., A. Mereau, P. S. Bilodeau, L. Damier, C. M. Stoltzfus, and C. Branlant.** 2001. A second exon splicing silencer within human immunodeficiency virus type 1 tat exon 2 represses splicing of Tat mRNA and binds protein hnRNP H. *The Journal of biological chemistry* **276**:40464-40475.
239. **Jacquenot, S., D. Ropers, P. S. Bilodeau, L. Damier, A. Mougin, C. M. Stoltzfus, and C. Branlant.** 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 research* **29**:464-478.
240. **Jenkins, J. L., H. Shen, M. R. Green, and C. L. Kielkopf.** 2008. Solution conformation and thermodynamic characteristics of RNA binding by the splicing factor U2AF65. *The Journal of biological chemistry* **283**:33641-33649.
241. **Jensen, K. B., K. Musunuru, H. A. Lewis, S. K. Burley, and R. B. Darnell.** 2000. The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proceedings of the National Academy of Sciences of the United States of America* **97**:5740-5745.
242. **Jia, R., C. Li, J. P. McCoy, C. X. Deng, and Z. M. Zheng.** 2010. SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. *International journal of biological sciences* **6**:806-826.
243. **Jimenez-Garcia, L. F., and D. L. Spector.** 1993. In vivo evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell* **73**:47-59.
244. **Jin, Y., H. Suzuki, S. Maegawa, H. Endo, S. Sugano, K. Hashimoto, K. Yasuda, and K. Inoue.** 2003. A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *The EMBO journal* **22**:905-912.

245. **Johnson, V. A., F. Brun-Vezinet, B. Clotet, B. Conway, R. T. D'Aquila, L. M. Demeter, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, A. Telenti, and D. D. Richman.** 2004. Update of the drug resistance mutations in HIV-1: 2004. Topics in HIV medicine : a publication of the International AIDS Society, USA **12**:119-124.
246. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. The EMBO journal **20**:1726-1738.
247. **Jumaa, H., G. Wei, and P. J. Nielsen.** 1999. Blastocyst formation is blocked in mouse embryos lacking the splicing factor SRp20. Current biology : CB **9**:899-902.
248. **Jurica, M. S., and M. J. Moore.** 2003. Pre-mRNA splicing: awash in a sea of proteins. Molecular cell **12**:5-14.
249. **Kadlec, J., E. Izaurralde, and S. Cusack.** 2004. The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. Nature structural & molecular biology **11**:330-337.
250. **Kaida, D., M. G. Berg, I. Younis, M. Kasim, L. N. Singh, L. Wan, and G. Dreyfuss.** 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature **468**:664-668.
251. **Kambach, C., S. Walke, R. Young, J. M. Avis, E. de la Fortelle, V. A. Raker, R. Luhrmann, J. Li, and K. Nagai.** 1999. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell **96**:375-387.
252. **Kamma, H., D. S. Portman, and G. Dreyfuss.** 1995. Cell type-specific expression of hnRNP proteins. Experimental cell research **221**:187-196.
253. **Kammler, S., C. Leurs, M. Freund, J. Krummheuer, K. Seidel, T. O. Tange, M. K. Lund, J. Kjems, A. Scheid, and H. Schaal.** 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.
254. **Kammler, S., M. Otte, I. Hauber, J. Kjems, J. Hauber, and H. Schaal.** 2006. The strength of the HIV-1 3' splice sites affects Rev function. Retrovirology **3**:89.

-
255. **Kanaar, R., S. E. Roche, E. L. Beall, M. R. Green, and D. C. Rio.** 1993. The conserved pre-mRNA splicing factor U2AF from *Drosophila*: requirement for viability. *Science* **262**:569-573.
256. **Kanopka, A., O. Muhlemann, and G. Akusjarvi.** 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* **381**:535-538.
257. **Karni, R., E. de Stanchina, S. W. Lowe, R. Sinha, D. Mu, and A. R. Krainer.** 2007. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nature structural & molecular biology* **14**:185-193.
258. **Kashima, T., and J. L. Manley.** 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nature genetics* **34**:460-463.
259. **Kashima, T., N. Rao, and J. L. Manley.** 2007. An intronic element contributes to splicing repression in spinal muscular atrophy. *Proceedings of the National Academy of Sciences of the United States of America* **104**:3426-3431.
260. **Kataoka, N., J. L. Bachorik, and G. Dreyfuss.** 1999. Transportin-SR, a nuclear import receptor for SR proteins. *The Journal of cell biology* **145**:1145-1152.
261. **Kawakami, A., Q. Tian, X. Duan, M. Streuli, S. F. Schlossman, and P. Anderson.** 1992. Identification and functional characterization of a TIA-1-related nucleolysin. *Proceedings of the National Academy of Sciences of the United States of America* **89**:8681-8685.
262. **Kawakami, A., Q. Tian, M. Streuli, M. Poe, S. Edelhoff, C. M. Disteché, and P. Anderson.** 1994. Intron-exon organization and chromosomal localization of the human TIA-1 gene. *J Immunol* **152**:4937-4945.
263. **Kawano, T., M. Fujita, and H. Sakamoto.** 2000. Unique and redundant functions of SR proteins, a conserved family of splicing factors, in *Caenorhabditis elegans* development. *Mechanisms of development* **95**:67-76.
264. **Kent, O. A., D. B. Ritchie, and A. M. Macmillan.** 2005. Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. *Molecular and cellular biology* **25**:233-240.
265. **Kielkopf, C. L., S. Lucke, and M. R. Green.** 2004. U2AF homology motifs: protein recognition in the RRM world. *Genes & development* **18**:1513-1526.
266. **Kim, J. H., B. Hahm, Y. K. Kim, M. Choi, and S. K. Jang.** 2000. Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *Journal of molecular biology* **298**:395-405.

267. **Kim, S. Y., R. Byrn, J. Groopman, and D. Baltimore.** 1989. Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. *Journal of virology* **63**:3708-3713.
268. **Kinomoto, M., M. Yokoyama, H. Sato, A. Kojima, T. Kurata, K. Ikuta, T. Sata, and K. Tokunaga.** 2005. Amino acid 36 in the human immunodeficiency virus type 1 gp41 ectodomain controls fusogenic activity: implications for the molecular mechanism of viral escape from a fusion inhibitor. *Journal of virology* **79**:5996-6004.
269. **Kirschbaum-Slager, N., G. M. Lopes, P. A. Galante, G. J. Riggins, and S. J. de Souza.** 2004. Splicing factors are differentially expressed in tumors. *Genetics and molecular research : GMR* **3**:512-520.
270. **Kjems, J., and P. Askjaer.** 2000. Rev protein and its cellular partners. *Adv Pharmacol* **48**:251-298.
271. **Kjems, J., M. Brown, D. D. Chang, and P. A. Sharp.** 1991. Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proceedings of the National Academy of Sciences of the United States of America* **88**:683-687.
272. **Klatzmann, D., E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J. C. Gluckman, and L. Montagnier.** 1984. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**:767-768.
273. **Klippel, S., M. Wieczorek, M. Schumann, E. Krause, B. Marg, T. Seidel, T. Meyer, E. W. Knapp, and C. Freund.** 2011. Multivalent binding of formin-binding protein 21 (FBP21)-tandem-WW domains fosters protein recognition in the pre-spliceosome. *The Journal of biological chemistry* **286**:38478-38487.
274. **Klotman, M. E., S. Kim, A. Buchbinder, A. DeRossi, D. Baltimore, and F. Wong-Staal.** 1991. Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes. *Proceedings of the National Academy of Sciences of the United States of America* **88**:5011-5015.
275. **Kogan, M., and J. Rappaport.** 2011. HIV-1 accessory protein Vpr: relevance in the pathogenesis of HIV and potential for therapeutic intervention. *Retrovirology* **8**:25.

276. **Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. A. Garcia-Blanco, and J. L. Manley.** 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* **368**:119-124.
277. **Kol, G., G. Lev-Maor, and G. Ast.** 2005. Human-mouse comparative analysis reveals that branch-site plasticity contributes to splicing regulation. *Human molecular genetics* **14**:1559-1568.
278. **Konarska, M. M., J. Vilardell, and C. C. Query.** 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Molecular cell* **21**:543-553.
279. **Konermann, C.** Heinrich-Heine-Universität Düsseldorf, 2004, diploma thesis. Regulation des alternativen Spleißens der HIV-1 Exons mit der 5' Spleißstelle #4.
280. **Konig, J., K. Zarnack, G. Rot, T. Curk, M. Kayikci, B. Zupan, D. J. Turner, N. M. Luscombe, and J. Ule.** 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology* **17**:909-915.
281. **Kotlajich, M. V., T. L. Crabb, and K. J. Hertel.** 2009. Spliceosome assembly pathways for different types of alternative splicing converge during commitment to splice site pairing in the A complex. *Molecular and cellular biology* **29**:1072-1082.
282. **Kozak, M.** 2002. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**:1-34.
283. **Krainer, A. R., G. C. Conway, and D. Kozak.** 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & development* **4**:1158-1171.
284. **Krecic, A. M., and M. S. Swanson.** 1999. hnRNP complexes: composition, structure, and function. *Current opinion in cell biology* **11**:363-371.
285. **Kress, T. L., N. J. Krogan, and C. Guthrie.** 2008. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. *Molecular cell* **32**:727-734.
286. **Krummheuer, J., C. Lenz, S. Kammler, A. Scheid, and H. Schaal.** 2001. Influence of the small leader exons 2 and 3 on human immunodeficiency virus type 1 gene expression. *Virology* **286**:276-289.

-
287. **Kumar, A. O., M. C. Swenson, M. M. Benning, and C. L. Kielkopf.** 2008. Structure of the central RNA recognition motif of human TIA-1 at 1.95Å resolution. *Biochemical and biophysical research communications* **367**:813-819.
288. **Kuroyanagi, H.** 2009. Fox-1 family of RNA-binding proteins. *Cellular and molecular life sciences* : CMLS **66**:3895-3907.
289. **Kyriakopoulou, C., P. Larsson, L. Liu, J. Schuster, F. Soderbom, L. A. Kirsebom, and A. Virtanen.** 2006. U1-like snRNAs lacking complementarity to canonical 5' splice sites. *RNA* **12**:1603-1611.
290. **Labourier, E., and D. C. Rio.** 2001. Purification of Drosophila snRNPs and characterization of two populations of functional U1 particles. *RNA* **7**:457-470.
291. **Lai, M. C., R. I. Lin, S. Y. Huang, C. W. Tsai, and W. Y. Tarn.** 2000. A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *The Journal of biological chemistry* **275**:7950-7957.
292. **Lamichhane, R., G. M. Daubner, J. Thomas-Crusells, S. D. Auweter, C. Manatschal, K. S. Austin, O. Valniuk, F. H. Allain, and D. Rueda.** 2010. RNA looping by PTB: Evidence using FRET and NMR spectroscopy for a role in splicing repression. *Proceedings of the National Academy of Sciences of the United States of America* **107**:4105-4110.
293. **Lamond, A. I., and D. L. Spector.** 2003. Nuclear speckles: a model for nuclear organelles. *Nature reviews. Molecular cell biology* **4**:605-612.
294. **Landi, A., V. Iannucci, A. V. Nuffel, P. Meuwissen, and B. Verhasselt.** 2011. One protein to rule them all: modulation of cell surface receptors and molecules by HIV Nef. *Current HIV research* **9**:496-504.
295. **Lareau, L. F., R. E. Green, R. S. Bhatnagar, and S. E. Brenner.** 2004. The evolving roles of alternative splicing. *Current opinion in structural biology* **14**:273-282.
296. **Lareau, L. F., M. Inada, R. E. Green, J. C. Wengrod, and S. E. Brenner.** 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**:926-929.
297. **Le Guiner, C., F. Lejeune, D. Galiana, L. Kister, R. Breathnach, J. Stevenin, and F. Del Gatto-Konczak.** 2001. TIA-1 and TIAR activate splicing

- of alternative exons with weak 5' splice sites followed by a U-rich stretch on their own pre-mRNAs. *The Journal of biological chemistry* **276**:40638-40646.
298. **Le Hir, H., and G. R. Andersen.** 2008. Structural insights into the exon junction complex. *Current opinion in structural biology* **18**:112-119.
 299. **Le Hir, H., E. Izaurralde, L. E. Maquat, and M. J. Moore.** 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO journal* **19**:6860-6869.
 300. **Le Hir, H., and B. Seraphin.** 2008. EJCs at the heart of translational control. *Cell* **133**:213-216.
 301. **Ledee, D. R., J. Chen, L. H. Tonelli, H. Takase, I. Gery, and P. S. Zelenka.** 2004. Differential expression of splice variants of chemokine CCL27 mRNA in lens, cornea, and retina of the normal mouse eye. *Molecular vision* **10**:663-667.
 302. **Lewis, H. A., K. Musunuru, K. B. Jensen, C. Edo, H. Chen, R. B. Darnell, and S. K. Burley.** 2000. Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* **100**:323-332.
 303. **Li, T., E. Evdokimov, R. F. Shen, C. C. Chao, E. Tekle, T. Wang, E. R. Stadtman, D. C. Yang, and P. B. Chock.** 2004. Sumoylation of heterogeneous nuclear ribonucleoproteins, zinc finger proteins, and nuclear pore complex proteins: a proteomic analysis. *Proceedings of the National Academy of Sciences of the United States of America* **101**:8551-8556.
 304. **Li, W., Y. Li, N. Kedersha, P. Anderson, M. Emara, K. M. Swiderek, G. T. Moreno, and M. A. Brinton.** 2002. Cell proteins TIA-1 and TIAR interact with the 3' stem-loop of the West Nile virus complementary minus-strand RNA and facilitate virus replication. *Journal of virology* **76**:11989-12000.
 305. **Licatalosi, D. D., A. Mele, J. J. Fak, J. Ule, M. Kayikci, S. W. Chi, T. A. Clark, A. C. Schweitzer, J. E. Blume, X. Wang, J. C. Darnell, and R. B. Darnell.** 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**:464-469.
 306. **Lim, F., and D. S. Peabody.** 2002. RNA recognition site of PP7 coat protein. *Nucleic acids research* **30**:4138-4144.

307. **Lim, L. P., and P. A. Sharp.** 1998. Alternative splicing of the fibronectin EIIIB exon depends on specific TGCATG repeats. *Molecular and cellular biology* **18**:3900-3906.
308. **Lim, S. R., and K. J. Hertel.** 2004. Commitment to splice site pairing coincides with A complex formation. *Molecular cell* **15**:477-483.
309. **Lin, C. H., and J. G. Patton.** 1995. Regulation of alternative 3' splice site selection by constitutive splicing factors. *RNA* **1**:234-245.
310. **Lin, S., and X. D. Fu.** 2007. SR proteins and related factors in alternative splicing. *Advances in experimental medicine and biology* **623**:107-122.
311. **Little, S. J., S. Holte, J. P. Routy, E. S. Daar, M. Markowitz, A. C. Collier, R. A. Koup, J. W. Mellors, E. Connick, B. Conway, M. Kilby, L. Wang, J. M. Whitcomb, N. S. Hellmann, and D. D. Richman.** 2002. Antiretroviral-drug resistance among patients recently infected with HIV. *The New England journal of medicine* **347**:385-394.
312. **Liu, H. X., S. L. Chew, L. Cartegni, M. Q. Zhang, and A. R. Krainer.** 2000. Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Molecular and cellular biology* **20**:1063-1071.
313. **Liu, H. X., M. Zhang, and A. R. Krainer.** 1998. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & development* **12**:1998-2012.
314. **Liu, Z., I. Luyten, M. J. Bottomley, A. C. Messias, S. HOUNGNINOU-MOLANGO, R. Sprangers, K. Zanier, A. Kramer, and M. Sattler.** 2001. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* **294**:1098-1102.
315. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**:402-408.
316. **Llorian, M., S. Schwartz, T. A. Clark, D. Hollander, L. Y. Tan, R. Spellman, A. Gordon, A. C. Schweitzer, P. de la Grange, G. Ast, and C. W. Smith.** 2010. Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nature structural & molecular biology* **17**:1114-1123.

317. **Long, J. C., and J. F. Cáceres.** 2009. The SR protein family of splicing factors: master regulators of gene expression. *The Biochemical journal* **417**:15-27.
318. **Longman, D., I. L. Johnstone, and J. F. Cáceres.** 2000. Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *The EMBO journal* **19**:1625-1637.
319. **Longman, D., T. McGarvey, S. McCracken, I. L. Johnstone, B. J. Blencowe, and J. F. Cáceres.** 2001. Multiple interactions between SRm160 and SR family proteins in enhancer-dependent splicing and development of *C. elegans*. *Current biology : CB* **11**:1923-1933.
320. **Lunde, B. M., C. Moore, and G. Varani.** 2007. RNA-binding proteins: modular design for efficient function. *Nature reviews. Molecular cell biology* **8**:479-490.
321. **Luo, Y., H. Yu, and B. M. Peterlin.** 1994. Cellular protein modulates effects of human immunodeficiency virus type 1 Rev. *Journal of virology* **68**:3850-3856.
322. **Lutzemberger, M., T. Gross, and N. F. Käufer.** 1999. Srp2, an SR protein family member of fission yeast: in vivo characterization of its modular domains. *Nucleic acids research* **27**:2618-2626.
323. **Luukkonen, B. G., and B. Seraphin.** 1997. The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in *Saccharomyces cerevisiae*. *The EMBO journal* **16**:779-792.
324. **Lynch, K. W.** 2007. Regulation of alternative splicing by signal transduction pathways. *Advances in experimental medicine and biology* **623**:161-174.
325. **Ma, C. T., G. Ghosh, X. D. Fu, and J. A. Adams.** 2010. Mechanism of dephosphorylation of the SR protein ASF/SF2 by protein phosphatase 1. *Journal of molecular biology* **403**:386-404.
326. **Mackereth, C. D., T. Madl, S. Bonnal, B. Simon, K. Zanier, A. Gasch, V. Rybin, J. Valcarcel, and M. Sattler.** 2011. Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **475**:408-411.
327. **MacMillan, A. M., P. S. McCaw, J. D. Crispino, and P. A. Sharp.** 1997. SC35-mediated reconstitution of splicing in U2AF-depleted nuclear extract. *Proceedings of the National Academy of Sciences of the United States of America* **94**:133-136.

-
328. **Maddon, P. J., A. G. Dalgleish, J. S. McDougal, P. R. Clapham, R. A. Weiss, and R. Axel.** 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47**:333-348.
329. **Madsen, J. M., and C. M. Stoltzfus.** 2005. An exonic splicing silencer downstream of the 3' splice site A2 is required for efficient human immunodeficiency virus type 1 replication. *Journal of virology* **79**:10478-10486.
330. **Mahalingam, B., P. Boross, Y. F. Wang, J. M. Louis, C. C. Fischer, J. Tozser, R. W. Harrison, and I. T. Weber.** 2002. Combining mutations in HIV-1 protease to understand mechanisms of resistance. *Proteins* **48**:107-116.
331. **Majewski, J., and J. Ott.** 2002. Distribution and characterization of regulatory elements in the human genome. *Genome research* **12**:1827-1836.
332. **Makarova, O. V., E. M. Makarov, and R. Luhrmann.** 2001. The 65 and 110 kDa SR-related proteins of the U4/U6.U5 tri-snRNP are essential for the assembly of mature spliceosomes. *The EMBO journal* **20**:2553-2563.
333. **Maldarelli, F., C. Xiang, G. Chamoun, and S. L. Zeichner.** 1998. The expression of the essential nuclear splicing factor SC35 is altered by human immunodeficiency virus infection. *Virus research* **53**:39-51.
334. **Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338**:254-257.
335. **Malim, M. H., L. S. Tiley, D. F. McCarn, J. R. Rusche, J. Hauber, and B. R. Cullen.** 1990. HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell* **60**:675-683.
336. **Manceau, V., M. Swenson, J. P. Le Caer, A. Sobel, C. L. Kielkopf, and A. Maucuer.** 2006. Major phosphorylation of SF1 on adjacent Ser-Pro motifs enhances interaction with U2AF65. *The FEBS journal* **273**:577-587.
337. **Mandal, D., C. M. Exline, Z. Feng, and C. M. Stoltzfus.** 2009. Regulation of Vif mRNA splicing by human immunodeficiency virus type 1 requires 5' splice site D2 and an exonic splicing enhancer to counteract cellular restriction factor APOBEC3G. *Journal of virology* **83**:6067-6078.
338. **Mandal, D., Z. Feng, and C. M. Stoltzfus.** 2010. Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs. *Journal of virology* **84**:12790-12800.

-
339. **Maniatis, T., and B. Tasic.** 2002. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **418**:236-243.
340. **Manley, J. L., and A. R. Krainer.** 2010. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes & development* **24**:1073-1074.
341. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. *Genes & development* **10**:1569-1579.
342. **Marchand, V., A. Mereau, S. Jacquenet, D. Thomas, A. Mougin, R. Gattoni, J. Stevenin, and C. Branlant.** 2002. A Janus splicing regulatory element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *Journal of molecular biology* **323**:629-652.
343. **Maris, C., C. Dominguez, and F. H. Allain.** 2005. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS journal* **272**:2118-2131.
344. **Martinez-Contreras, R., P. Cloutier, L. Shkreta, J. F. Fisette, T. Revil, and B. Chabot.** 2007. hnRNP proteins and splicing control. *Advances in experimental medicine and biology* **623**:123-147.
345. **Martinez-Contreras, R., J. F. Fisette, F. U. Nasim, R. Madden, M. Cordeau, and B. Chabot.** 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS biology* **4**:e21.
346. **Masur, H., M. A. Michelis, J. B. Greene, I. Onorato, R. A. Stouwe, R. S. Holzman, G. Wormser, L. Brettman, M. Lange, H. W. Murray, and S. Cunningham-Rundles.** 1981. An outbreak of community-acquired *Pneumocystis carinii* pneumonia: initial manifestation of cellular immune dysfunction. *The New England journal of medicine* **305**:1431-1438.
347. **Matlin, A. J., F. Clark, and C. W. Smith.** 2005. Understanding alternative splicing: towards a cellular code. *Nature reviews. Molecular cell biology* **6**:386-398.
348. **Matlin, A. J., J. Southby, C. Gooding, and C. W. Smith.** 2007. Repression of alpha-actinin SM exon splicing by assisted binding of PTB to the polypyrimidine tract. *RNA* **13**:1214-1223.
349. **Mauger, D. M., C. Lin, and M. A. Garcia-Blanco.** 2008. hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc. *Molecular and cellular biology* **28**:5403-5419.

350. **Mayas, R. M., H. Maita, D. R. Semlow, and J. P. Staley.** 2010. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *Proceedings of the National Academy of Sciences of the United States of America* **107**:10020-10025.
351. **Mayas, R. M., H. Maita, and J. P. Staley.** 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nature structural & molecular biology* **13**:482-490.
352. **Mayeda, A., D. M. Helfman, and A. R. Krainer.** 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Molecular and cellular biology* **13**:2993-3001.
353. **Mayeda, A., and A. R. Krainer.** 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68**:365-375.
354. **Mayeda, A., S. H. Munroe, J. F. Caceres, and A. R. Krainer.** 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *The EMBO journal* **13**:5483-5495.
355. **Mayeda, A., S. H. Munroe, R. M. Xu, and A. R. Krainer.** 1998. Distinct functions of the closely related tandem RNA-recognition motifs of hnRNP A1. *RNA* **4**:1111-1123.
356. **Maynard, C. M., and K. B. Hall.** 2010. Interactions between PTB RRMs induce slow motions and increase RNA binding affinity. *Journal of molecular biology* **397**:260-277.
357. **McAlinden, A., L. Liang, Y. Mukudai, T. Imamura, and L. J. Sandell.** 2007. Nuclear protein TIA-1 regulates COL2A1 alternative splicing and interacts with precursor mRNA and genomic DNA. *The Journal of biological chemistry* **282**:24444-24454.
358. **McDougal, J. S., M. S. Kennedy, J. M. Sligh, S. P. Cort, A. Mawle, and J. K. Nicholson.** 1986. Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* **231**:382-385.
359. **McDougal, J. S., P. J. Maddon, A. G. Dalgleish, P. R. Clapham, D. R. Littman, M. Godfrey, D. E. Maddon, L. Chess, R. A. Weiss, and R. Axel.** 1986. The T4 glycoprotein is a cell-surface receptor for the AIDS virus. *Cold Spring Harbor symposia on quantitative biology* **51 Pt 2**:703-711.

360. **Melton, A. A., J. Jackson, J. Wang, and K. W. Lynch.** 2007. Combinatorial control of signal-induced exon repression by hnRNP L and PSF. *Molecular and cellular biology* **27**:6972-6984.
361. **Merendino, L., S. Guth, D. Bilbao, C. Martinez, and J. Valcarcel.** 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402**:838-841.
362. **Mermoud, J. E., P. Cohen, and A. I. Lamond.** 1992. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic acids research* **20**:5263-5269.
363. **Mermoud, J. E., P. T. Cohen, and A. I. Lamond.** 1994. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *The EMBO journal* **13**:5679-5688.
364. **Merrill, B. M., S. F. Barnett, W. M. LeSturgeon, and K. R. Williams.** 1989. Primary structure differences between proteins C1 and C2 of HeLa 40S nuclear ribonucleoprotein particles. *Nucleic acids research* **17**:8441-8449.
365. **Michael, N. L., P. Morrow, J. Mosca, M. Vahey, D. S. Burke, and R. R. Redfield.** 1991. Induction of human immunodeficiency virus type 1 expression in chronically infected cells is associated primarily with a shift in RNA splicing patterns. *Journal of virology* **65**:7084.
366. **Michaud, S., and R. Reed.** 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & development* **5**:2534-2546.
367. **Michlewski, G., J. R. Sanford, and J. F. Cáceres.** 2008. The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. *Molecular cell* **30**:179-189.
368. **Mikula, M., J. Karczmarski, A. Dzwonek, T. Rubel, E. Hennig, M. Dadlez, J. M. Bujnicki, K. Bomsztyk, and J. Ostrowski.** 2006. Casein kinases phosphorylate multiple residues spanning the entire hnRNP K length. *Biochimica et biophysica acta* **1764**:299-306.
369. **Mili, S., H. J. Shu, Y. Zhao, and S. Pinol-Roma.** 2001. Distinct RNP complexes of shuttling hnRNP proteins with pre-mRNA and mRNA: candidate intermediates in formation and export of mRNA. *Molecular and cellular biology* **21**:7307-7319.

-
370. **Min, H., R. C. Chan, and D. L. Black.** 1995. The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. *Genes & development* **9**:2659-2671.
371. **Misteli, T., J. F. Caceres, J. Q. Clement, A. R. Krainer, M. F. Wilkinson, and D. L. Spector.** 1998. Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *The Journal of cell biology* **143**:297-307.
372. **Misteli, T., J. F. Caceres, and D. L. Spector.** 1997. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**:523-527.
373. **Modafferi, E. F., and D. L. Black.** 1997. A complex intronic splicing enhancer from the c-src pre-mRNA activates inclusion of a heterologous exon. *Molecular and cellular biology* **17**:6537-6545.
374. **Moroy, T., and F. Heyd.** 2007. The impact of alternative splicing in vivo: mouse models show the way. *RNA* **13**:1155-1171.
375. **Motta-Mena, L. B., F. Heyd, and K. W. Lynch.** 2010. Context-dependent regulatory mechanism of the splicing factor hnRNP L. *Molecular cell* **37**:223-234.
376. **Mullen, M. P., C. W. Smith, J. G. Patton, and B. Nadal-Ginard.** 1991. Alpha-tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines default exon choice. *Genes & development* **5**:642-655.
377. **Mulligan, G. J., W. Guo, S. Wormsley, and D. M. Helfman.** 1992. Polypyrimidine tract binding protein interacts with sequences involved in alternative splicing of beta-tropomyosin pre-mRNA. *The Journal of biological chemistry* **267**:25480-25487.
378. **Munis, J. R., R. S. Kornbluth, J. C. Guatelli, and D. D. Richman.** 1992. Ordered appearance of human immunodeficiency virus type 1 nucleic acids following high multiplicity infection of macrophages. *The Journal of general virology* **73 (Pt 8)**:1899-1906.
379. **Murphy, E. L., A. C. Collier, L. A. Kalish, S. F. Assmann, M. F. Para, T. P. Flanagan, P. N. Kumar, L. Mintz, F. R. Wallach, and G. J. Nemo.** 2001. Highly active antiretroviral therapy decreases mortality and morbidity in patients with advanced HIV disease. *Annals of internal medicine* **135**:17-26.

-
380. **Nakahata, S., and S. Kawamoto.** 2005. Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic acids research* **33**:2078-2089.
381. **Nasim, F. U., S. Hutchison, M. Cordeau, and B. Chabot.** 2002. High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA* **8**:1078-1089.
382. **Nelissen, R. L., V. Heinrichs, W. J. Habets, F. Simons, R. Luhrmann, and W. J. van Venrooij.** 1991. Zinc finger-like structure in U1-specific protein C is essential for specific binding to U1 snRNP. *Nucleic acids research* **19**:449-454.
383. **Nelissen, R. L., C. L. Will, W. J. van Venrooij, and R. Luhrmann.** 1994. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *The EMBO journal* **13**:4113-4125.
384. **Nelson, K. K., and M. R. Green.** 1989. Mammalian U2 snRNP has a sequence-specific RNA-binding activity. *Genes & development* **3**:1562-1571.
385. **Neubauer, G., A. Gottschalk, P. Fabrizio, B. Seraphin, R. Luhrmann, and M. Mann.** 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **94**:385-390.
386. **Ni, J. Z., L. Grate, J. P. Donohue, C. Preston, N. Nobida, G. O'Brien, L. Shiue, T. A. Clark, J. E. Blume, and M. Ares, Jr.** 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & development* **21**:708-718.
387. **Nichols, R. C., X. W. Wang, J. Tang, B. J. Hamilton, F. A. High, H. R. Herschman, and W. F. Rigby.** 2000. The RGG domain in hnRNP A2 affects subcellular localization. *Experimental cell research* **256**:522-532.
388. **Nilsen, T. W., and B. R. Graveley.** 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**:457-463.
389. **Norton, P. A.** 1994. Polypyrimidine tract sequences direct selection of alternative branch sites and influence protein binding. *Nucleic acids research* **22**:3854-3860.

390. **Nott, A., S. H. Meislin, and M. J. Moore.** 2003. A quantitative analysis of intron effects on mammalian gene expression. *RNA* **9**:607-617.
391. **O'Reilly, M. M., M. T. McNally, and K. L. Beemon.** 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.
392. **Oberstrass, F. C., S. D. Auweter, M. Erat, Y. Hargous, A. Henning, P. Wenter, L. Reymond, B. Amir-Ahmady, S. Pitsch, D. L. Black, and F. H. Allain.** 2005. Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science* **309**:2054-2057.
393. **Oh, Y. L., B. Hahm, Y. K. Kim, H. K. Lee, J. W. Lee, O. Song, K. Tsukiyama-Kohara, M. Kohara, A. Nomoto, and S. K. Jang.** 1998. Determination of functional domains in polypyrimidine-tract-binding protein. *The Biochemical journal* **331 (Pt 1)**:169-175.
394. **Ohkura, N., M. Takahashi, H. Yaguchi, Y. Nagamura, and T. Tsukada.** 2005. Coactivator-associated arginine methyltransferase 1, CARM1, affects pre-mRNA splicing in an isoform-specific manner. *The Journal of biological chemistry* **280**:28927-28935.
395. **Ohtaka, H., A. Schon, and E. Freire.** 2003. Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations. *Biochemistry* **42**:13659-13666.
396. **Okamoto, Y., H. Onogi, R. Honda, H. Yasuda, T. Wakabayashi, Y. Nimura, and M. Hagiwara.** 1998. cdc2 kinase-mediated phosphorylation of splicing factor SF2/ASF. *Biochemical and biophysical research communications* **249**:872-878.
397. **Okunola, H. L., and A. R. Krainer.** 2009. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Molecular and cellular biology* **29**:5620-5631.
398. **Orengo, J. P., and T. A. Cooper.** 2007. Alternative splicing in disease. *Advances in experimental medicine and biology* **623**:212-223.
399. **Ott, M., M. Geyer, and Q. Zhou.** 2011. The control of HIV transcription: keeping RNA polymerase II on track. *Cell host & microbe* **10**:426-435.
400. **Otte, M.** Heinrich-Heine-Universität Düsseldorf, 2006, PhD thesis. Identifizierung von cis-wirkenden Sequenzen in den alternativen HIV-1 Leaderexons und ihre funktionelle Bedeutung für die Spleißregulation.

-
401. **Pacheco, T. R., M. B. Coelho, J. M. Desterro, I. Mollet, and M. Carmo-Fonseca.** 2006. In vivo requirement of the small subunit of U2AF for recognition of a weak 3' splice site. *Molecular and cellular biology* **26**:8183-8190.
402. **Pacheco, T. R., A. Q. Gomes, N. L. Barbosa-Morais, V. Benes, W. Ansorge, M. Wollerton, C. W. Smith, J. Valcarcel, and M. Carmo-Fonseca.** 2004. Diversity of vertebrate splicing factor U2AF35: identification of alternatively spliced U2AF1 mRNAs. *The Journal of biological chemistry* **279**:27039-27049.
403. **Pacheco, T. R., L. F. Moita, A. Q. Gomes, N. Hacohen, and M. Carmo-Fonseca.** 2006. RNA interference knockdown of hU2AF35 impairs cell cycle progression and modulates alternative splicing of Cdc25 transcripts. *Molecular biology of the cell* **17**:4187-4199.
404. **Pagani, F., and F. E. Baralle.** 2004. Genomic variants in exons and introns: identifying the splicing spoilers. *Nature reviews. Genetics* **5**:389-396.
405. **Palella, F. J., Jr., K. M. Delaney, A. C. Moorman, M. O. Loveless, J. Fuhrer, G. A. Satten, D. J. Aschman, and S. D. Holmberg.** 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *The New England journal of medicine* **338**:853-860.
406. **Pan, Q., O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe.** 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature genetics* **40**:1413-1415.
407. **Park, S. H., G. H. Park, H. Gu, W. I. Hwang, I. K. Lim, W. K. Paik, and S. Kim.** 1997. Heterogeneous nuclear RNP protein A1-arginine methylation during HCT-48 cell cycle. *Biochemistry and molecular biology international* **42**:657-666.
408. **Parker, R., P. G. Siliciano, and C. Guthrie.** 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**:229-239.
409. **Pastuszak, A. W., M. P. Joachimiak, M. Blanchette, D. C. Rio, S. E. Brenner, and A. D. Frankel.** 2011. An SF1 affinity model to identify branch point sequences in human introns. *Nucleic acids research* **39**:2344-2356.

-
410. **Patel, N. A., S. Kaneko, H. S. Apostolatos, S. S. Bae, J. E. Watson, K. Davidowitz, D. S. Chappell, M. J. Birnbaum, J. Q. Cheng, and D. R. Cooper.** 2005. Molecular and genetic studies imply Akt-mediated signaling promotes protein kinase C β 1 alternative splicing via phosphorylation of serine/arginine-rich splicing factor SRp40. *The Journal of biological chemistry* **280**:14302-14309.
411. **Peng, X., and S. M. Mount.** 1995. Genetic enhancement of RNA-processing defects by a dominant mutation in B52, the *Drosophila* gene for an SR protein splicing factor. *Molecular and cellular biology* **15**:6273-6282.
412. **Perez, I., C. H. Lin, J. G. McAfee, and J. G. Patton.** 1997. Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo. *RNA* **3**:764-778.
413. **Perez, I., J. G. McAfee, and J. G. Patton.** 1997. Multiple RRM domains contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein. *Biochemistry* **36**:11881-11890.
414. **Perriman, R., and M. Ares, Jr.** 2010. Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing. *Molecular cell* **38**:416-427.
415. **Perriman, R. J., and M. Ares, Jr.** 2007. Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. *Genes & development* **21**:811-820.
416. **Petersen-Mahrt, S. K., C. Estmer, C. Ohmalm, D. A. Matthews, W. C. Russell, and G. Akusjarvi.** 1999. The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *The EMBO journal* **18**:1014-1024.
417. **Pinol-Roma, S., Y. D. Choi, M. J. Matunis, and G. Dreyfuss.** 1988. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes & development* **2**:215-227.
418. **Pinol-Roma, S., and G. Dreyfuss.** 1993. hnRNP proteins: localization and transport between the nucleus and the cytoplasm. *Trends in cell biology* **3**:151-155.
419. **Pinol-Roma, S., M. S. Swanson, J. G. Gall, and G. Dreyfuss.** 1989. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *The Journal of cell biology* **109**:2575-2587.

-
420. **Pollard, V. W., and M. H. Malim.** 1998. The HIV-1 Rev protein. Annual review of microbiology **52**:491-532.
421. **Pomeranz Krummel, D. A., C. Oubridge, A. K. Leung, J. Li, and K. Nagai.** 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. Nature **458**:475-480.
422. **Ponthier, J. L., C. Schluepen, W. Chen, R. A. Lersch, S. L. Gee, V. C. Hou, A. J. Lo, S. A. Short, J. A. Chasis, J. C. Winkelmann, and J. G. Conboy.** 2006. Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1R alternative exon 16. The Journal of biological chemistry **281**:12468-12474.
423. **Popielarz, M., Y. Cavaloc, M. G. Mattei, R. Gattoni, and J. Stevenin.** 1995. The gene encoding human splicing factor 9G8. Structure, chromosomal localization, and expression of alternatively processed transcripts. The Journal of biological chemistry **270**:17830-17835.
424. **Prasad, J., K. Colwill, T. Pawson, and J. L. Manley.** 1999. The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. Molecular and cellular biology **19**:6991-7000.
425. **Purcell, D. F., and M. A. Martin.** 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. Journal of virology **67**:6365-6378.
426. **Qi, J., S. Su, and W. Mattox.** 2007. The doublesex splicing enhancer components Tra2 and Rbp1 also repress splicing through an intronic silencer. Molecular and cellular biology **27**:699-708.
427. **Query, C. C., M. J. Moore, and P. A. Sharp.** 1994. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes & development **8**:587-597.
428. **Rain, J. C., Z. Rafi, Z. Rhani, P. Legrain, and A. Kramer.** 1998. Conservation of functional domains involved in RNA binding and protein-protein interactions in human and *Saccharomyces cerevisiae* pre-mRNA splicing factor SF1. RNA **4**:551-565.
429. **Rasheva, V. I., D. Knight, P. Bozko, K. Marsh, and M. V. Frolov.** 2006. Specific role of the SR protein splicing factor B52 in cell cycle control in *Drosophila*. Molecular and cellular biology **26**:3468-3477.

-
430. **Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, and et al.** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313**:277-284.
431. **Reed, R.** 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes & development* **3**:2113-2123.
432. **Ring, H. Z., and J. T. Lis.** 1994. The SR protein B52/SRp55 is essential for *Drosophila* development. *Molecular and cellular biology* **14**:7499-7506.
433. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and cellular biology* **10**:84-94.
434. **Robert-Guroff, M., M. Popovic, S. Gartner, P. Markham, R. C. Gallo, and M. S. Reitz.** 1990. Structure and expression of tat-, rev-, and nef-specific transcripts of human immunodeficiency virus type 1 in infected lymphocytes and macrophages. *Journal of virology* **64**:3391-3398.
435. **Romani, B., R. H. Glashoff, and S. Engelbrecht.** 2010. Functional integrity of naturally occurring mutants of HIV-1 subtype C Vpr. *Virus research* **153**:288-298.
436. **Romano, M.** 2010. G runs in cystathionine beta-synthase c.833C/c.844_845ins68 mRNA are splicing silencers of pathogenic 3' splice sites. *Biochimica et biophysica acta* **1799**:568-574.
437. **Romfo, C. M., C. J. Alvarez, W. J. van Heeckeren, C. J. Webb, and J. A. Wise.** 2000. Evidence for splice site pairing via intron definition in *Schizosaccharomyces pombe*. *Molecular and cellular biology* **20**:7955-7970.
438. **Rooke, N., V. Markovtsov, E. Cagavi, and D. L. Black.** 2003. Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. *Molecular and cellular biology* **23**:1874-1884.
439. **Ropers, D., L. Ayadi, R. Gattoni, S. Jacquenet, L. Damier, C. Branlant, and J. Stevenin.** 2004. 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. *The Journal of biological chemistry* **279**:29963-29973.
440. **Roscigno, R. F., and M. A. Garcia-Blanco.** 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* **1**:692-706.

441. **Rosel, T. D., L. H. Hung, J. Medenbach, K. Donde, S. Starke, V. Benes, G. Ratsch, and A. Bindereif.** 2011. RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. *The EMBO journal* **30**:1965-1976.
442. **Rosin, S.** Heinrich-Heine-Universität Düsseldorf, 2008, diploma thesis. Die Rekrutierung spleißregulatorischer Proteindomänen in die Nähe von HIV-1 Spleißstellen und ihre Wirkung auf die Spleißstellennutzung.
443. **Rossi, F., E. Labourier, T. Forne, G. Divita, J. Derancourt, J. F. Riou, E. Antoine, G. Cathala, C. Brunel, and J. Tazi.** 1996. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature* **381**:80-82.
444. **Roth, M. B., A. M. Zahler, and J. A. Stolk.** 1991. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *The Journal of cell biology* **115**:587-596.
445. **Rousseau, M. N., L. Vergne, B. Montes, M. Peeters, J. Reynes, E. Delaporte, and M. Segondy.** 2001. Patterns of resistance mutations to antiretroviral drugs in extensively treated HIV-1-infected patients with failure of highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* **26**:36-43.
446. **Rudner, D. Z., R. Kanaar, K. S. Breger, and D. C. Rio.** 1998. Interaction between subunits of heterodimeric splicing factor U2AF is essential in vivo. *Molecular and cellular biology* **18**:1765-1773.
447. **Rudner, D. Z., R. Kanaar, K. S. Breger, and D. C. Rio.** 1996. Mutations in the small subunit of the *Drosophila* U2AF splicing factor cause lethality and developmental defects. *Proceedings of the National Academy of Sciences of the United States of America* **93**:10333-10337.
448. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52**:207-219.
449. **Rutz, B., and B. Seraphin.** 1999. Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA* **5**:819-831.
450. **Ryo, A., Y. Suzuki, M. Arai, N. Kondoh, T. Wakatsuki, A. Hada, M. Shuda, K. Tanaka, C. Sato, M. Yamamoto, and N. Yamamoto.** 2000. Identification and characterization of differentially expressed mRNAs in HIV type 1-infected human T cells. *AIDS research and human retroviruses* **16**:995-1005.

451. **Saguez, C., J. R. Olesen, and T. H. Jensen.** 2005. Formation of export-competent mRNP: escaping nuclear destruction. *Current opinion in cell biology* **17**:287-293.
452. **Sakabe, N. J., and S. J. de Souza.** 2007. Sequence features responsible for intron retention in human. *BMC genomics* **8**:59.
453. **Saliou, J. M., C. F. Bourgeois, L. Ayadi-Ben Mena, D. Ropers, S. Jacquenet, V. Marchand, J. Stevenin, and C. Branlant.** 2009. Role of RNA structure and protein factors in the control of HIV-1 splicing. *Frontiers in bioscience : a journal and virtual library* **14**:2714-2729.
454. **Sambrook, J., Fritsch, E.F. and Maniatis, T.** 1989. *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
455. **Samson, M. L.** 2008. Rapid functional diversification in the structurally conserved ELAV family of neuronal RNA binding proteins. *BMC genomics* **9**:392.
456. **Sanford, J. R., N. K. Gray, K. Beckmann, and J. F. Caceres.** 2004. A novel role for shuttling SR proteins in mRNA translation. *Genes & development* **18**:755-768.
457. **Sanford, J. R., X. Wang, M. Mort, N. Vanduyne, D. N. Cooper, S. D. Mooney, H. J. Edenberg, and Y. Liu.** 2009. Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. *Genome research* **19**:381-394.
458. **Sapra, A. K., M. L. Anko, I. Grishina, M. Lorenz, M. Pabis, I. Poser, J. Rollins, E. M. Weiland, and K. M. Neugebauer.** 2009. SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. *Molecular cell* **34**:179-190.
459. **Sato, H., N. Hosoda, and L. E. Maquat.** 2008. Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. *Molecular cell* **29**:255-262.
460. **Schaal, H., P. Pfeiffer, M. Klein, P. Gehrman, and A. Scheid.** 1993. Use of DNA end joining activity of a *Xenopus laevis* egg extract for construction of deletions and expression vectors for HIV-1 Tat and Rev proteins. *Gene* **124**:275-280.

461. **Schaal, T. D., and T. Maniatis.** 1999. Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. *Molecular and cellular biology* **19**:261-273.
462. **Schaal, T. D., and T. Maniatis.** 1999. Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. *Molecular and cellular biology* **19**:1705-1719.
463. **Schaub, M. C., S. R. Lopez, and M. Caputi.** 2007. Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. *The Journal of biological chemistry* **282**:13617-13626.
464. **Scherer, S.** 2008. *A Short Guide to the Human Genome*. Cold Spring Harbor, NY. Cold Spring Harbor Laboratory Press.
465. **Schneider, M., C. L. Will, M. Anokhina, J. Tazi, H. Urlaub, and R. Luhrmann.** 2010. Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes. *Molecular cell* **38**:223-235.
466. **Schöneweis, K.** Heinrich-Heine-Universität Düsseldorf, 2010, diploma thesis. U1 snRNA – vermittelter Gentherapieansatz zur Behandlung humaner 5' Spleißstellenmutationen.
467. **Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo, and G. N. Pavlakis.** 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *Journal of virology* **64**:2519-2529.
468. **Schwartz, S., B. K. Felber, E. M. Fenyo, and G. N. Pavlakis.** 1990. Env and Vpu proteins of human immunodeficiency virus type 1 are produced from multiple bicistronic mRNAs. *Journal of virology* **64**:5448-5456.
469. **Schwartz, S., B. K. Felber, and G. N. Pavlakis.** 1991. Expression of human immunodeficiency virus type 1 vif and vpr mRNAs is Rev-dependent and regulated by splicing. *Virology* **183**:677-686.
470. **Schwerk, C., and K. Schulze-Osthoff.** 2005. Regulation of apoptosis by alternative pre-mRNA splicing. *Molecular cell* **19**:1-13.
471. **Screaton, G. R., J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, and A. R. Krainer.** 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *The EMBO journal* **14**:4336-4349.

-
472. **Selden, R. F., K. B. Howie, M. E. Rowe, H. M. Goodman, and D. D. Moore.** 1986. Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Molecular and cellular biology* **6**:3173-3179.
473. **Selenko, P., G. Gregorovic, R. Sprangers, G. Stier, Z. Rhani, A. Kramer, and M. Sattler.** 2003. Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Molecular cell* **11**:965-976.
474. **Shan, J., T. P. Munro, E. Barbarese, J. H. Carson, and R. Smith.** 2003. A molecular mechanism for mRNA trafficking in neuronal dendrites. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **23**:8859-8866.
475. **Shao, W., H. S. Kim, Y. Cao, Y. Z. Xu, and C. C. Query.** 2012. A U1-U2 snRNP interaction network during intron definition. *Molecular and cellular biology* **32**:470-478.
476. **Sharkey, M. E., I. Teo, T. Greenough, N. Sharova, K. Luzuriaga, J. L. Sullivan, R. P. Bucy, L. G. Kostrikis, A. Haase, C. Veryard, R. E. Davaro, S. H. Cheeseman, J. S. Daly, C. Bova, R. T. Ellison, 3rd, B. Mady, K. K. Lai, G. Moyle, M. Nelson, B. Gazzard, S. Shaunak, and M. Stevenson.** 2000. Persistence of episomal HIV-1 infection intermediates in patients on highly active anti-retroviral therapy. *Nature medicine* **6**:76-81.
477. **Sharma, S., L. A. Kohlstaedt, A. Damianov, D. C. Rio, and D. L. Black.** 2008. Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nature structural & molecular biology* **15**:183-191.
478. **Sharma, S., C. Maris, F. H. Allain, and D. L. Black.** 2011. U1 snRNA directly interacts with polypyrimidine tract-binding protein during splicing repression. *Molecular cell* **41**:579-588.
479. **Sharova, N., C. Swingler, M. Sharkey, and M. Stevenson.** 2005. Macrophages archive HIV-1 virions for dissemination in trans. *The EMBO journal* **24**:2481-2489.
480. **Shaw, S. D., S. Chakrabarti, G. Ghosh, and A. R. Krainer.** 2007. Deletion of the N-terminus of SF2/ASF permits RS-domain-independent pre-mRNA splicing. *PloS one* **2**:e854.

481. **Shen, H., and M. R. Green.** 2004. A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Molecular cell* **16**:363-373.
482. **Shen, H., and M. R. Green.** 2006. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes & development* **20**:1755-1765.
483. **Shen, M., and W. Mattox.** 2012. Activation and repression functions of an SR splicing regulator depend on exonic versus intronic-binding position. *Nucleic acids research* **40**:428-437.
484. **Shepard, P. J., and K. J. Hertel.** 2008. Conserved RNA secondary structures promote alternative splicing. *RNA* **14**:1463-1469.
485. **Shepard, P. J., and K. J. Hertel.** 2009. The SR protein family. *Genome biology* **10**:242.
486. **Shepard, S., M. McCreary, and A. Fedorov.** 2009. The peculiarities of large intron splicing in animals. *PloS one* **4**:e7853.
487. **Shi, H., and R. M. Xu.** 2003. Crystal structure of the Drosophila Mago nashi-Y14 complex. *Genes & development* **17**:971-976.
488. **Shi, J., Z. Hu, K. Pabon, and K. W. Scotto.** 2008. Caffeine regulates alternative splicing in a subset of cancer-associated genes: a role for SC35. *Molecular and cellular biology* **28**:883-895.
489. **Shibata, H., D. P. Huynh, and S. M. Pulst.** 2000. A novel protein with RNA-binding motifs interacts with ataxin-2. *Human molecular genetics* **9**:1303-1313.
490. **Shitashige, M., Y. Naishiro, M. Idogawa, K. Honda, M. Ono, S. Hirohashi, and T. Yamada.** 2007. Involvement of splicing factor-1 in beta-catenin/T-cell factor-4-mediated gene transactivation and pre-mRNA splicing. *Gastroenterology* **132**:1039-1054.
491. **Shitashige, M., R. Satow, K. Honda, M. Ono, S. Hirohashi, and T. Yamada.** 2007. Increased susceptibility of Sf1(+/-) mice to azoxymethane-induced colon tumorigenesis. *Cancer science* **98**:1862-1867.
492. **Shukla, S., F. Del Gatto-Konczak, R. Breathnach, and S. A. Fisher.** 2005. Competition of PTB with TIA proteins for binding to a U-rich cis-element determines tissue-specific splicing of the myosin phosphatase targeting subunit 1. *RNA* **11**:1725-1736.

-
493. **Si, Z., B. A. Amendt, and C. M. Stoltzfus.** 1997. 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 research* **25**:861-867.
494. **Si, Z. H., D. Rauch, and C. M. Stoltzfus.** 1998. The exon splicing silencer in human immunodeficiency virus type 1 Tat exon 3 is bipartite and acts early in spliceosome assembly. *Molecular and cellular biology* **18**:5404-5413.
495. **Sickmier, E. A., K. E. Frato, H. Shen, S. R. Paranawithana, M. R. Green, and C. L. Kielkopf.** 2006. Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. *Molecular cell* **23**:49-59.
496. **Siegal, F. P., C. Lopez, G. S. Hammer, A. E. Brown, S. J. Kornfeld, J. Gold, J. Hassett, S. Z. Hirschman, C. Cunningham-Rundles, B. R. Adelsberg, and et al.** 1981. Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. *The New England journal of medicine* **305**:1439-1444.
497. **Silva, A. L., and L. Romao.** 2009. The mammalian nonsense-mediated mRNA decay pathway: to decay or not to decay! Which players make the decision? *FEBS letters* **583**:499-505.
498. **Simpson, P. J., T. P. Monie, A. Szendroi, N. Davydova, J. K. Tyzack, M. R. Conte, C. M. Read, P. D. Cary, D. I. Svergun, P. V. Konarev, S. Curry, and S. Matthews.** 2004. Structure and RNA interactions of the N-terminal RRM domains of PTB. *Structure* **12**:1631-1643.
499. **Singh, K. K., S. Erkelenz, S. Rattay, A. K. Dehof, A. Hildebrandt, K. Schulze-Osthoff, H. Schaal, and C. Schwerk.** 2010. Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *RNA* **16**:2442-2454.
500. **Singh, R., and J. Valcarcel.** 2005. Building specificity with nonspecific RNA-binding proteins. *Nature structural & molecular biology* **12**:645-653.
501. **Singh, R., J. Valcarcel, and M. R. Green.** 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* **268**:1173-1176.
502. **Sloan, R. D., and M. A. Wainberg.** 2011. The role of unintegrated DNA in HIV infection. *Retrovirology* **8**:52.

503. **Smith, C. W., T. T. Chu, and B. Nadal-Ginard.** 1993. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Molecular and cellular biology* **13**:4939-4952.
504. **Smith, C. W., and J. Valcarcel.** 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in biochemical sciences* **25**:381-388.
505. **Smith, D. J., C. C. Query, and M. M. Konarska.** 2008. "Nought may endure but mutability": spliceosome dynamics and the regulation of splicing. *Molecular cell* **30**:657-666.
506. **Solis, A. S., N. Shariat, and J. G. Patton.** 2008. Splicing fidelity, enhancers, and disease. *Frontiers in bioscience : a journal and virtual library* **13**:1926-1942.
507. **Sommer, P., and U. Nehrbass.** 2005. Quality control of messenger ribonucleoprotein particles in the nucleus and at the pore. *Current opinion in cell biology* **17**:294-301.
508. **Sonza, S., H. P. Mutimer, K. O'Brien, P. Ellery, J. L. Howard, J. H. Axelrod, N. J. Deacon, S. M. Crowe, and D. F. Purcell.** 2002. Selectively reduced tat mRNA heralds the decline in productive human immunodeficiency virus type 1 infection in monocyte-derived macrophages. *Journal of virology* **76**:12611-12621.
509. **Soret, J., M. Gabut, C. Dupon, G. Kohlhagen, J. Stevenin, Y. Pommier, and J. Tazi.** 2003. Altered serine/arginine-rich protein phosphorylation and exonic enhancer-dependent splicing in Mammalian cells lacking topoisomerase I. *Cancer research* **63**:8203-8211.
510. **Soret, J., and J. Tazi.** 2003. Phosphorylation-dependent control of the pre-mRNA splicing machinery. *Progress in molecular and subcellular biology* **31**:89-126.
511. **Spector, D. L.** 1993. Macromolecular domains within the cell nucleus. *Annual review of cell biology* **9**:265-315.
512. **Spector, D. L.** 1996. Nuclear organization and gene expression. *Experimental cell research* **229**:189-197.
513. **Spellman, R., A. Rideau, A. Matlin, C. Gooding, F. Robinson, N. McGlincy, S. N. Grellscheid, J. Southby, M. Wollerton, and C. W. Smith.** 2005. Regulation of alternative splicing by PTB and associated factors. *Biochemical Society transactions* **33**:457-460.

514. **Spellman, R., and C. W. Smith.** 2006. Novel modes of splicing repression by PTB. *Trends in biochemical sciences* **31**:73-76.
515. **Sperling, J., M. Azubel, and R. Sperling.** 2008. Structure and function of the Pre-mRNA splicing machine. *Structure* **16**:1605-1615.
516. **Staffa, A., and A. Cochrane.** 1995. Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Molecular and cellular biology* **15**:4597-4605.
517. **Staffa, A., and A. Cochrane.** 1994. The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3' splice site. *Journal of virology* **68**:3071-3079.
518. **Staley, J. P., and C. Guthrie.** 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**:315-326.
519. **Stark, H., P. Dube, R. Luhrmann, and B. Kastner.** 2001. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* **409**:539-542.
520. **Stickeler, E., F. Kittrell, D. Medina, and S. M. Berget.** 1999. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene* **18**:3574-3582.
521. **Stoltzfus, C. M.** 2009. Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. *Advances in virus research* **74**:1-40.
522. **Stoltzfus, C. M., and J. M. Madsen.** 2006. Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. *Current HIV research* **4**:43-55.
523. **Strebel, K., D. Daugherty, K. Clouse, D. Cohen, T. Folks, and M. A. Martin.** 1987. The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature* **328**:728-730.
524. **Sugnet, C. W., K. Srinivasan, T. A. Clark, G. O'Brien, M. S. Cline, H. Wang, A. Williams, D. Kulp, J. E. Blume, D. Haussler, and M. Ares, Jr.** 2006. Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLoS computational biology* **2**:e4.
525. **Suhasini, M., and T. R. Reddy.** 2009. Cellular proteins and HIV-1 Rev function. *Current HIV research* **7**:91-100.

-
526. **Sun, S., Z. Zhang, O. Fregoso, and A. R. Krainer.** 2012. Mechanisms of activation and repression by the alternative splicing factors RBFOX1/2. *RNA* **18**:274-283.
527. **Sun, S., Z. Zhang, R. Sinha, R. Karni, and A. R. Krainer.** 2010. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nature structural & molecular biology* **17**:306-312.
528. **Sureau, A., R. Gattoni, Y. Dooghe, J. Stevenin, and J. Soret.** 2001. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *The EMBO journal* **20**:1785-1796.
529. **Swanson, C. M., N. M. Sherer, and M. H. Malim.** 2010. SRp40 and SRp55 promote the translation of unspliced human immunodeficiency virus type 1 RNA. *Journal of virology* **84**:6748-6759.
530. **Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss.** 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Molecular and cellular biology* **7**:1731-1739.
531. **Swartz, J. E., Y. C. Bor, Y. Misawa, D. Rekosh, and M. L. Hammarskjold.** 2007. The shuttling SR protein 9G8 plays a role in translation of unspliced mRNA containing a constitutive transport element. *The Journal of biological chemistry* **282**:19844-19853.
532. **Szymczynska, B. R., J. Bowman, S. McCracken, A. Pineda-Lucena, Y. Lu, B. Cox, M. Lambermon, B. R. Graveley, C. H. Arrowsmith, and B. J. Blencowe.** 2003. Structure and function of the PWI motif: a novel nucleic acid-binding domain that facilitates pre-mRNA processing. *Genes & development* **17**:461-475.
533. **Tacke, R., Y. Chen, and J. L. Manley.** 1997. Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. *Proceedings of the National Academy of Sciences of the United States of America* **94**:1148-1153.
534. **Tacke, R., and J. L. Manley.** 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. *The EMBO journal* **14**:3540-3551.

-
535. **Tacke, R., M. Tohyama, S. Ogawa, and J. L. Manley.** 1998. Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. *Cell* **93**:139-148.
536. **Takahashi, K., S. L. Wesselingh, D. E. Griffin, J. C. McArthur, R. T. Johnson, and J. D. Glass.** 1996. Localization of HIV-1 in human brain using polymerase chain reaction/in situ hybridization and immunocytochemistry. *Annals of neurology* **39**:705-711.
537. **Tanackovic, G., and A. Kramer.** 2005. Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. *Molecular biology of the cell* **16**:1366-1377.
538. **Tang, J., N. Abovich, M. L. Fleming, B. Seraphin, and M. Rosbash.** 1997. Identification and characterization of a yeast homolog of U1 snRNP-specific protein C. *The EMBO journal* **16**:4082-4091.
539. **Tang, Z. Z., S. Zheng, J. Nikolic, and D. L. Black.** 2009. Developmental control of CaV1.2 L-type calcium channel splicing by Fox proteins. *Molecular and cellular biology* **29**:4757-4765.
540. **Tange, T. O., C. K. Damgaard, S. Guth, J. Valcarcel, and J. Kjems.** 2001. The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *The EMBO journal* **20**:5748-5758.
541. **Tange, T. O., T. H. Jensen, and J. Kjems.** 1996. In vitro interaction between human immunodeficiency virus type 1 Rev protein and splicing factor ASF/SF2-associated protein, p32. *The Journal of biological chemistry* **271**:10066-10072.
542. **Thickman, K. R., E. A. Sickmier, and C. L. Kielkopf.** 2007. Alternative conformations at the RNA-binding surface of the N-terminal U2AF(65) RNA recognition motif. *Journal of molecular biology* **366**:703-710.
543. **Tian, Q., M. Streuli, H. Saito, S. F. Schlossman, and P. Anderson.** 1991. A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. *Cell* **67**:629-639.
544. **Tollervey, J. R., T. Curk, B. Rogelj, M. Briese, M. Cereda, M. Kayikci, J. Konig, T. Hortobagyi, A. L. Nishimura, V. Zupunski, R. Patani, S. Chandran, G. Rot, B. Zupan, C. E. Shaw, and J. Ule.** 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nature neuroscience* **14**:452-458.

-
545. **Tranell, A., E. M. Fenyo, and S. Schwartz.** 2010. Serine- and arginine-rich proteins 55 and 75 (SRp55 and SRp75) induce production of HIV-1 vpr mRNA by inhibiting the 5'-splice site of exon 3. *The Journal of biological chemistry* **285**:31537-31547.
546. **Tranell, A., S. Tingsborg, E. M. Fenyo, and S. Schwartz.** 2011. Inhibition of splicing by serine-arginine rich protein 55 (SRp55) causes the appearance of partially spliced HIV-1 mRNAs in the cytoplasm. *Virus research* **157**:82-91.
547. **Tripathi, V., J. D. Ellis, Z. Shen, D. Y. Song, Q. Pan, A. T. Watt, S. M. Freier, C. F. Bennett, A. Sharma, P. A. Bubulya, B. J. Blencowe, S. G. Prasanth, and K. V. Prasanth.** 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular cell* **39**:925-938.
548. **Tsuruno, C., K. Ohe, M. Kuramitsu, T. Kohma, Y. Takahama, Y. Hamaguchi, I. Hamaguchi, and K. Okuma.** 2011. HMGA1a is involved in specific splice site regulation of human immunodeficiency virus type 1. *Biochemical and biophysical research communications* **406**:512-517.
549. **Tuerk, C., and L. Gold.** 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**:505-510.
550. **Ule, J., G. Stefani, A. Mele, M. Ruggiu, X. Wang, B. Taneri, T. Gaasterland, B. J. Blencowe, and R. B. Darnell.** 2006. An RNA map predicting Nova-dependent splicing regulation. *Nature* **444**:580-586.
551. **UNAIDS.** 2009. AIDS Epidemic Update 2009. <http://www.unaids.org/en/KnowledgeCentre/HIVData/EpiUpdate/EpiUpdArchive/2009/default.asp,:1-100>.
552. **UNAIDS.** 2008. Report on the global AIDS epidemic. http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/2008_Global_report.asp,:1-360.
553. **Underwood, J. G., P. L. Boutz, J. D. Dougherty, P. Stoilov, and D. L. Black.** 2005. Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Molecular and cellular biology* **25**:10005-10016.

-
554. **Valcarcel, J., R. K. Gaur, R. Singh, and M. R. Green.** 1996. Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. *Science* **273**:1706-1709.
555. **Valentine, C. R.** 1998. The association of nonsense codons with exon skipping. *Mutation research* **411**:87-117.
556. **van der Houven van Oordt, W., M. T. Diaz-Meco, J. Lozano, A. R. Krainer, J. Moscat, and J. F. Caceres.** 2000. The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *The Journal of cell biology* **149**:307-316.
557. **van Eekelen, C. A., T. Riemen, and W. J. van Venrooij.** 1981. Specificity in the interaction of hnRNA and mRNA with proteins as revealed by in vivo cross linking. *FEBS letters* **130**:223-226.
558. **Van Lint, C., S. Emiliani, M. Ott, and E. Verdin.** 1996. Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *The EMBO journal* **15**:1112-1120.
559. **Varbanov, M., L. Espert, and M. Biard-Piechaczyk.** 2006. Mechanisms of CD4 T-cell depletion triggered by HIV-1 viral proteins. *AIDS reviews* **8**:221-236.
560. **Vassileva, M. T., and M. J. Matunis.** 2004. SUMO modification of heterogeneous nuclear ribonucleoproteins. *Molecular and cellular biology* **24**:3623-3632.
561. **Venables, J. P.** 2007. Downstream intronic splicing enhancers. *FEBS letters* **581**:4127-4131.
562. **Venables, J. P., D. J. Elliott, O. V. Makarova, E. M. Makarov, H. J. Cooke, and I. C. Eperon.** 2000. RBMY, a probable human spermatogenesis factor, and other hnRNP G proteins interact with Tra2beta and affect splicing. *Human molecular genetics* **9**:685-694.
563. **Venables, J. P., C. S. Koh, U. Froehlich, E. Lapointe, S. Couture, L. Inkel, A. Bramard, E. R. Paquet, V. Watier, M. Durand, J. F. Lucier, J. Gervais-Bird, K. Tremblay, P. Prinos, R. Klinck, S. A. Elela, and B. Chabot.** 2008. Multiple and specific mRNA processing targets for the major human hnRNP proteins. *Molecular and cellular biology* **28**:6033-6043.

-
564. **Verdin, E., P. Paras, Jr., and C. Van Lint.** 1993. Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *The EMBO journal* **12**:3249-3259.
565. **Victoriano, A. F., K. Imai, H. Togami, T. Ueno, K. Asamitsu, T. Suzuki, N. Miyata, K. Ochiai, and T. Okamoto.** 2011. Novel histone deacetylase inhibitor NCH-51 activates latent HIV-1 gene expression. *FEBS letters* **585**:1103-1111.
566. **Vitali, F., A. Henning, F. C. Oberstrass, Y. Hargous, S. D. Auweter, M. Erat, and F. H. Allain.** 2006. Structure of the two most C-terminal RNA recognition motifs of PTB using segmental isotope labeling. *The EMBO journal* **25**:150-162.
567. **Wahl, M. C., C. L. Will, and R. Luhrmann.** 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**:701-718.
568. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alizon.** 1985. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**:9-17.
569. **Wang, E., W. F. Mueller, K. J. Hertel, and F. Cambi.** 2011. G Run-mediated recognition of proteolipid protein and DM20 5' splice sites by U1 small nuclear RNA is regulated by context and proximity to the splice site. *The Journal of biological chemistry* **286**:4059-4071.
570. **Wang, G. S., and T. A. Cooper.** 2007. Splicing in disease: disruption of the splicing code and the decoding machinery. *Nature reviews. Genetics* **8**:749-761.
571. **Wang, H. Y., W. Lin, J. A. Dyck, J. M. Yeakley, Z. Songyang, L. C. Cantley, and X. D. Fu.** 1998. SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. *The Journal of cell biology* **140**:737-750.
572. **Wang, J., and J. L. Manley.** 1995. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. *RNA* **1**:335-346.
573. **Wang, X., S. Bruderer, Z. Rafi, J. Xue, P. J. Milburn, A. Kramer, and P. J. Robinson.** 1999. Phosphorylation of splicing factor SF1 on Ser20 by cGMP-dependent protein kinase regulates spliceosome assembly. *The EMBO journal* **18**:4549-4559.

-
574. **Wang, Z., and C. B. Burge.** 2008. Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**:802-813.
575. **Webb, C. J., S. Lakhe-Reddy, C. M. Romfo, and J. A. Wise.** 2005. Analysis of mutant phenotypes and splicing defects demonstrates functional collaboration between the large and small subunits of the essential splicing factor U2AF in vivo. *Molecular biology of the cell* **16**:584-596.
576. **Webb, C. J., and J. A. Wise.** 2004. The splicing factor U2AF small subunit is functionally conserved between fission yeast and humans. *Molecular and cellular biology* **24**:4229-4240.
577. **Weber, G., S. Trowitzsch, B. Kastner, R. Luhrmann, and M. C. Wahl.** 2010. Functional organization of the Sm core in the crystal structure of human U1 snRNP. *The EMBO journal* **29**:4172-4184.
578. **Weinstock, H. S., I. Zaidi, W. Heneine, D. Bennett, J. G. Garcia-Lerma, J. M. Douglas, Jr., M. LaLota, G. Dickinson, S. Schwarcz, L. Torian, D. Wendell, S. Paul, G. A. Goza, J. Ruiz, B. Boyett, and J. E. Kaplan.** 2004. The epidemiology of antiretroviral drug resistance among drug-naïve HIV-1-infected persons in 10 US cities. *The Journal of infectious diseases* **189**:2174-2180.
579. **Weissenhorn, W., A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley.** 1997. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**:426-430.
580. **Wentz-Hunter, K., and J. Potashkin.** 1996. The small subunit of the splicing factor U2AF is conserved in fission yeast. *Nucleic acids research* **24**:1849-1854.
581. **Whitson, S. R., W. M. LeSturgeon, and A. M. Krezel.** 2005. Solution structure of the symmetric coiled coil tetramer formed by the oligomerization domain of hnRNP C: implications for biological function. *Journal of molecular biology* **350**:319-337.
582. **Widera, M.** Heinrich-Heine-Universität Düsseldorf, 2010, diploma thesis. Identifizierung von intronischen Spleißregulationselementen und deren funktionelle Bedeutung für die HIV-1 vif-/vpr-mRNA-Prozessierung.
583. **Will, C. L., and R. Luhrmann.** 2011. Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* **3**.

584. **Will, C. L., S. Rumpler, J. Klein Gunnewiek, W. J. van Venrooij, and R. Luhrmann.** 1996. In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic acids research* **24**:4614-4623.
585. **Witten, J. T., and J. Ule.** 2011. Understanding splicing regulation through RNA splicing maps. *Trends in genetics : TIG* **27**:89-97.
586. **Wodrich, H., and H. G. Krausslich.** 2001. Nucleocytoplasmic RNA transport in retroviral replication. *Results and problems in cell differentiation* **34**:197-217.
587. **Wonderlich, E. R., J. A. Leonard, and K. L. Collins.** 2011. HIV immune evasion disruption of antigen presentation by the HIV Nef protein. *Advances in virus research* **80**:103-127.
588. **Wu, J., and J. L. Manley.** 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes & development* **3**:1553-1561.
589. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**:1061-1070.
590. **Wu, S., C. M. Romfo, T. W. Nilsen, and M. R. Green.** 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **402**:832-835.
591. **Xiao, S. H., and J. L. Manley.** 1998. Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. *The EMBO journal* **17**:6359-6367.
592. **Xiao, S. H., and J. L. Manley.** 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes & development* **11**:334-344.
593. **Xiao, X., Z. Wang, M. Jang, R. Nutiu, E. T. Wang, and C. B. Burge.** 2009. Splice site strength-dependent activity and genetic buffering by poly-G runs. *Nature structural & molecular biology* **16**:1094-1100.
594. **Xie, J., J. A. Lee, T. L. Kress, K. L. Mowry, and D. L. Black.** 2003. Protein kinase A phosphorylation modulates transport of the polypyrimidine tract-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **100**:8776-8781.

-
595. **Xu, X., D. Yang, J. H. Ding, W. Wang, P. H. Chu, N. D. Dalton, H. Y. Wang, J. R. Bermingham, Jr., Z. Ye, F. Liu, M. G. Rosenfeld, J. L. Manley, J. Ross, Jr., J. Chen, R. P. Xiao, H. Cheng, and X. D. Fu.** 2005. ASF/SF2-regulated CaMKII δ alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell* **120**:59-72.
596. **Xu, Y. Z., C. M. Newnham, S. Kameoka, T. Huang, M. M. Konarska, and C. C. Query.** 2004. Prp5 bridges U1 and U2 snRNPs and enables stable U2 snRNP association with intron RNA. *The EMBO journal* **23**:376-385.
597. **Xu, Y. Z., and C. C. Query.** 2007. Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly. *Molecular cell* **28**:838-849.
598. **Xue, Y., Y. Zhou, T. Wu, T. Zhu, X. Ji, Y. S. Kwon, C. Zhang, G. Yeo, D. L. Black, H. Sun, X. D. Fu, and Y. Zhang.** 2009. Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Molecular cell* **36**:996-1006.
599. **Yang, G., S. C. Huang, J. Y. Wu, and E. J. Benz, Jr.** 2008. Regulated Fox-2 isoform expression mediates protein 4.1R splicing during erythroid differentiation. *Blood* **111**:392-401.
600. **Yang, X., M. R. Bani, S. J. Lu, S. Rowan, Y. Ben-David, and B. Chabot.** 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **91**:6924-6928.
601. **Yang, Y. Y., G. L. Yin, and R. B. Darnell.** 1998. The neuronal RNA-binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. *Proceedings of the National Academy of Sciences of the United States of America* **95**:13254-13259.
602. **Yeo, G., and C. B. Burge.** 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of computational biology : a journal of computational molecular cell biology* **11**:377-394.
603. **Yeo, G. W., N. G. Coufal, T. Y. Liang, G. E. Peng, X. D. Fu, and F. H. Gage.** 2009. An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nature structural & molecular biology* **16**:130-137.

-
604. **Yeo, G. W., E. L. Van Nostrand, and T. Y. Liang.** 2007. Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLoS genetics* **3**:e85.
605. **Yu, Y., P. A. Maroney, J. A. Denker, X. H. Zhang, O. Dybkov, R. Luhrmann, E. Jankowsky, L. A. Chasin, and T. W. Nilsen.** 2008. Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. *Cell* **135**:1224-1236.
606. **Zahler, A. M., C. K. Damgaard, J. Kjems, and M. Caputi.** 2004. 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. *The Journal of biological chemistry* **279**:10077-10084.
607. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes & development* **6**:837-847.
608. **Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth.** 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science* **260**:219-222.
609. **Zamore, P. D., and M. R. Green.** 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proceedings of the National Academy of Sciences of the United States of America* **86**:9243-9247.
610. **Zamore, P. D., J. G. Patton, and M. R. Green.** 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **355**:609-614.
611. **Zapp, M. L., and M. R. Green.** 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature* **342**:714-716.
612. **Zavanelli, M. I., J. S. Britton, A. H. Igel, and M. Ares, Jr.** 1994. Mutations in an essential U2 small nuclear RNA structure cause cold-sensitive U2 small nuclear ribonucleoprotein function by favoring competing alternative U2 RNA structures. *Molecular and cellular biology* **14**:1689-1697.
613. **Zhang, C., M. A. Frias, A. Mele, M. Ruggiu, T. Eom, C. B. Marney, H. Wang, D. D. Licatalosi, J. J. Fak, and R. B. Darnell.** 2010. Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* **329**:439-443.

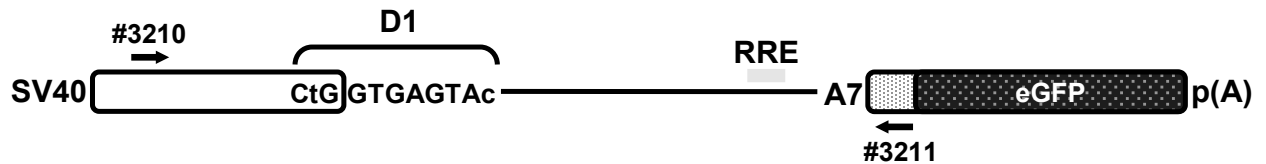
-
614. **Zhang, C., Z. Zhang, J. Castle, S. Sun, J. Johnson, A. R. Krainer, and M. Q. Zhang.** 2008. Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes & development* **22**:2550-2563.
615. **Zhang, L., C. Chung, B. S. Hu, T. He, Y. Guo, A. J. Kim, E. Skulsky, X. Jin, A. Hurley, B. Ramratnam, M. Markowitz, and D. D. Ho.** 2000. Genetic characterization of rebounding HIV-1 after cessation of highly active antiretroviral therapy. *The Journal of clinical investigation* **106**:839-845.
616. **Zhang, X. H., M. A. Arias, S. Ke, and L. A. Chasin.** 2009. Splicing of designer exons reveals unexpected complexity in pre-mRNA splicing. *RNA* **15**:367-376.
617. **Zhang, X. H., and L. A. Chasin.** 2004. Computational definition of sequence motifs governing constitutive exon splicing. *Genes & development* **18**:1241-1250.
618. **Zhang, Z., and A. R. Krainer.** 2004. Involvement of SR proteins in mRNA surveillance. *Molecular cell* **16**:597-607.
619. **Zhang, Z., and A. R. Krainer.** 2007. Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. *Proceedings of the National Academy of Sciences of the United States of America* **104**:11574-11579.
620. **Zheng, Y. H., H. F. Yu, and B. M. Peterlin.** 2003. Human p32 protein relieves a post-transcriptional block to HIV replication in murine cells. *Nature cell biology* **5**:611-618.
621. **Zhou, H. L., and H. Lou.** 2008. Repression of prespliceosome complex formation at two distinct steps by Fox-1/Fox-2 proteins. *Molecular and cellular biology* **28**:5507-5516.
622. **Zhu, H., R. A. Hasman, K. M. Young, N. L. Kedersha, and H. Lou.** 2003. U1 snRNP-dependent function of TIAR in the regulation of alternative RNA processing of the human calcitonin/CGRP pre-mRNA. *Molecular and cellular biology* **23**:5959-5971.
623. **Zhu, J., and A. R. Krainer.** 2000. Pre-mRNA splicing in the absence of an SR protein RS domain. *Genes & development* **14**:3166-3178.

624. **Zhu, J., A. Mayeda, and A. R. Krainer.** 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Molecular cell* **8**:1351-1361.
625. **Zhuang, Y., and A. M. Weiner.** 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes & development* **3**:1545-1552.
626. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**:827-835.
627. **Zorio, D. A., and T. Blumenthal.** 1999. Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **402**:835-838.
628. **Zuccato, E., E. Buratti, C. Stuani, F. E. Baralle, and F. Pagani.** 2004. An intronic polypyrimidine-rich element downstream of the donor site modulates cystic fibrosis transmembrane conductance regulator exon 9 alternative splicing. *The Journal of biological chemistry* **279**:16980-16988.
629. **Zuo, P., and T. Maniatis.** 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes & development* **10**:1356-1368.
630. **Zuo, P., and J. L. Manley.** 1993. Functional domains of the human splicing factor ASF/SF2. *The EMBO journal* **12**:4727-4737.

VII. Appendix

VII.1 Experimental Data

A



B

Splicing reporter	amplicon size [in bp] #3210/#3211
Fig.III-2	
SV GAR ⁻ ESE ⁻ D1-env/eGFP	180
SV GAR ⁻ D1-env/eGFP	180
SV SRSF7(2x) D1-env/eGFP	195
SV GAR ⁻ D1 SRSF7(2x)-env/eGFP	180
SV TIA-1 (2x) D1-env/eGFP	195
SV GAR ⁻ D1 TIA-1(2x)-env/eGFP	180
Fig.III-3	
SV neutral(3x) D1-env/eGFP	179
SV neutral(3x) D1 neutral(5x)-env/eGFP	179
SV SRSF7(2x) D1-env/eGFP	195
SV neutral(3x) D1 SRSF7(2x)-env/eGFP	179
SV TIA-1 (2x) D1-env/eGFP	195
SV neutral(3x) D1 TIA-1(2x)-env/eGFP	179
Fig.III-4	
SV GAR ⁻ ESE ⁻ D1-env/eGFP	180
SV neutral (3x) D1 GAR ⁻ ESE ⁻ -env/eGFP	180
SV GAR ⁻ D1-env/eGFP	180
SV neutral(3x) D1 GAR ⁻ -env/eGFP	180
SV SRSF2(2x) D1-env/eGFP	195
SV GAR ⁻ D1 SRSF2(2x)-env/eGFP	180
SV SRSF7(2x) D1-env/eGFP	195
SV GAR ⁻ D1 SRSF7(2x)-env/eGFP	180
SV TIA-1(2x) D1-env/eGFP	195
SV GAR ⁻ D1 TIA-1(2x)-env/eGFP	180

Appendix

SV IAS-1(2x) D1-env/eGFP	203	
SV GAR ⁻ D1 IAS-1(2x)-env/eGFP	180	
SV Fox1/2(2x) D1-env/eGFP	167	
SV GAR ⁻ D1 Fox1/2(2x)-env/eGFP	180	
SV hnRNP F(2x) D1-env/eGFP	165	
SV GAR ⁻ D1 hnRNP F/H(2x)-env/eGFP	180	
SV hnRNP A1(2x) D1-env/eGFP	167	
SV GAR ⁻ D1 hnRNP A1(2x)-env/eGFP	180	
SV PTB(2x) D1-env/eGFP	167	
SV GAR ⁻ D1 PTB(2x)-env/eGFP	180	
Fig.III-5		
SV neutral(3x) D1 neutral(5x)-env/eGFP	179	
SV SRSF2(1x) D1-env/eGFP	175	
SV SRSF2(2x) D1-env/eGFP	195	
SV SRSF2(4x) D1-env/eGFP	235	
SV neutral(3x) D1 TIA-1(1x) -env/eGFP	179	
SV neutral(3x) D1 TIA-1(2x) -env/eGFP	179	
SV neutral(3x) D1 TIA-1(4x) -env/eGFP	179	
Fig.III-6		
SV TIA-1(2x) D1 SRSF7(2x)-env/eGFP	195	
SV SRSF7(2x) D1 TIA-1(2x)-env/eGFP	195	
SV SRSF7(2x) L1 TIA-1(2x)-env/eGFP	195	
SV SRSF7(2x) L2 TIA-1(2x)-env/eGFP	195	
SV SRSF7(2x) L3 TIA-1(2x)-env/eGFP	195	
SV SRSF7(4x) L2 TIA-1(4x)-env/eGFP	235	
SV SRSF7(4x) L3 TIA-1(4x)-env/eGFP	235	
Fig.III-7		
SV 2MS2 D1-env/eGFP	223	
SV GAR ⁻ D1 2MS2-env/eGFP	180	
Fig.III-8		
	distal	proximal
SV D1 2xMS2 D1-env/eGFP	138	236
SV D1 GAR ⁻ ESE ⁻ D1-env/eGFP	138	194
SV D1 neutral(3x) D1-env/eGFP	138	193
SV D1 neutral(5x) D1-env/eGFP	138	209
SV D1 SRSF7(2x) D1-env/eGFP	138	209
SV D1 TIA-1(2x) D1-env/eGFP	138	209
Fig.III-9		
SV SRSF7(2x) neutral(3x) D1-env/eGFP	219	

SV SRSF7(2x) TIA-1(1x) D1-env/eGFP	215
SV SRSF7(2x) neutral(5x) D1-env/eGFP	235
SV SRSF7(2x) TIA-1(2x) D1-env/eGFP	235
SV SRSF2(2x) D1-env/eGFP	195
SV TIA-1(2x) SRSF7(2x) D1-env/eGFP	235
SV neutral(3x) SRSF7(2x) D1-env/eGFP	219
SV GAR ⁻ ESE ⁻ D1-env/eGFP	180
SV GAR ⁻ D1-env/eGFP	180
SV SRSF7(2x) D1-env/eGFP	195
SV GAR ⁻ D1 SRSF7(2x)-env/eGFP	180
SV TIA-1 (2x) D1-env/eGFP	195
SV GAR ⁻ D1 TIA-1(2x)-env/eGFP	180
SV SRSF7(2x) GAR ⁻ ESE ⁻ D1-env/eGFP	219
SV SRSF7(2x) TIA-1(2x) D1-env/eGFP	235
Fig.III-9	
SV 2MS2 D1 2PP7-env/eGFP	223

Fig.VII-1.1 Primer positions and PCR product sizes for SV-env/eGFP derived constructs

(A) Primer positions of #3210 and #3211 within the HIV-1 based splicing reporter SV-env/eGFP. **(B)** PCR product sizes after amplification using primer pair #3210/#3211.

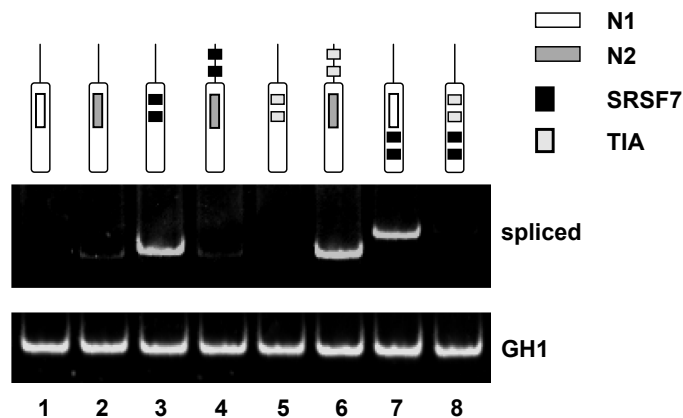
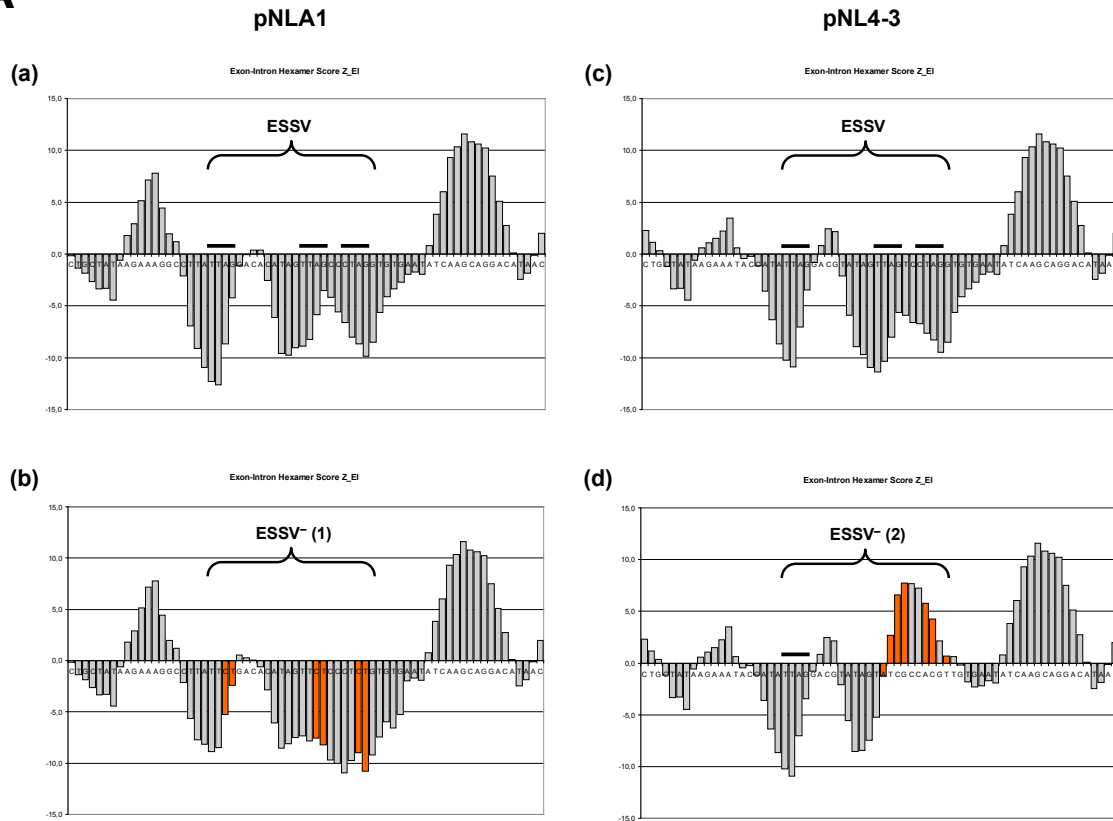


Fig.VII-1.2: PDEs act as silencers of splicing outside their activating position

RT-PCR analysis was performed as described in Fig.III-9 using the neutral sequence N1 for spacing (lane 7).

A



B

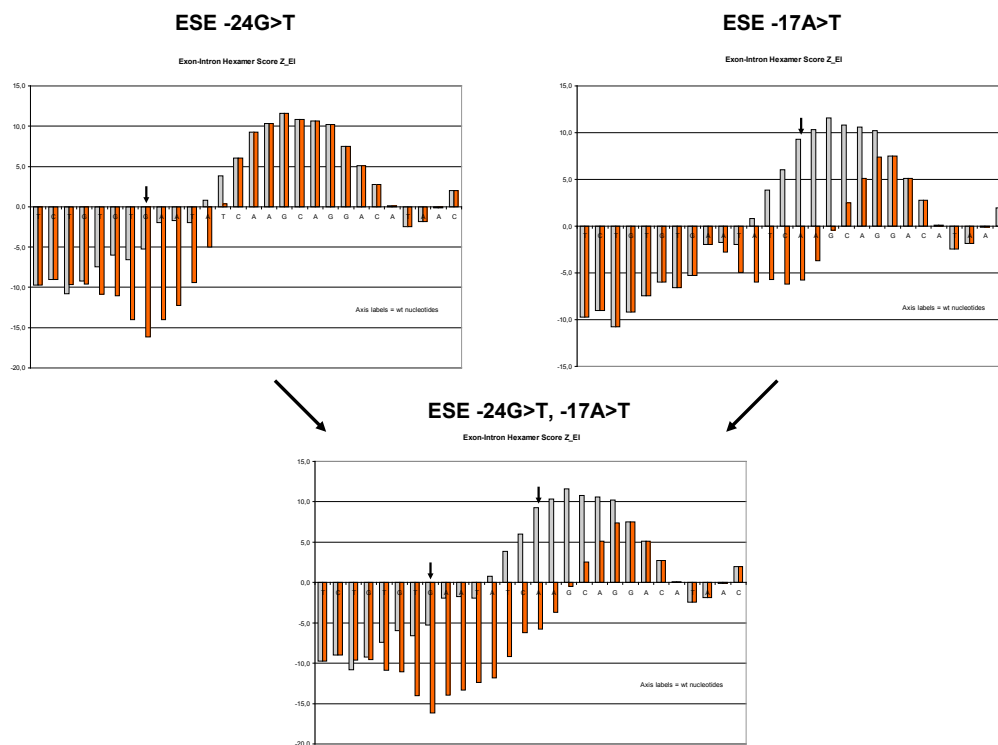
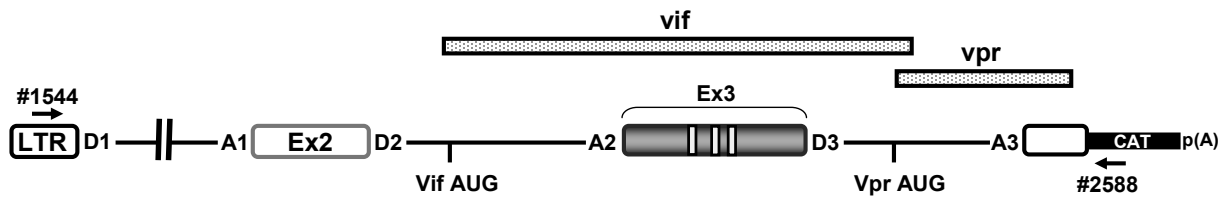


Fig.VII-1.3: Hexamer Score changes following ESSV and ESE mutations

(A) Z_{EI}-Score plot. The X axis shows the exon 3 nucleotide sequence of pNLA-1 and pNL4-3. The Y axis depicts the variations in score along the sequence. Mutated nucleotide residues within the ESSV are indicated (orange). **(B)** Each plot is showing the Z_{EI}-Scores of the wildtype (grey) versus mutant ESE sequence (orange). The mutated nucleotides are indicated by arrows.

A



B

LTR ex2 ex3	amplicon size [in bp] #1544/#2588
Tat1 [1.4]	267
Tat2 [1.2.4]	317
Tat3 [1.3.4]	341
Tat4 [1.2.3.4]	391
Vpr1 [1.3E]	654
Vpr2 [1.2.3E]	704
Vif1 [1.2E]	1131

Fig.VII-1.4: Primer positions and PCR product sizes for LTR ex2 ex3

(A) Primer positions of #1544 and #2588 within the HIV-1 based splicing reporter SV-env/eGFP. **(B)** PCR product sizes for the alternatively spliced reporter mRNAs after amplification using primer pair #1544/#2588.

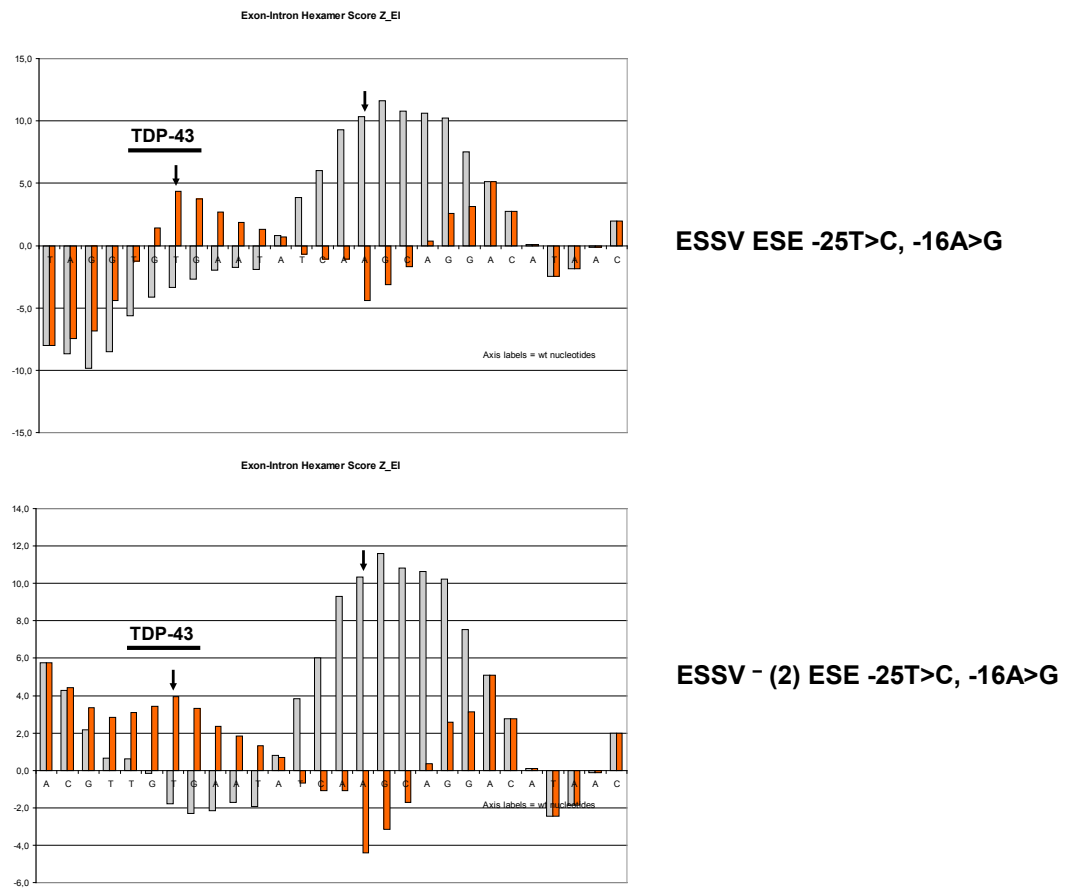
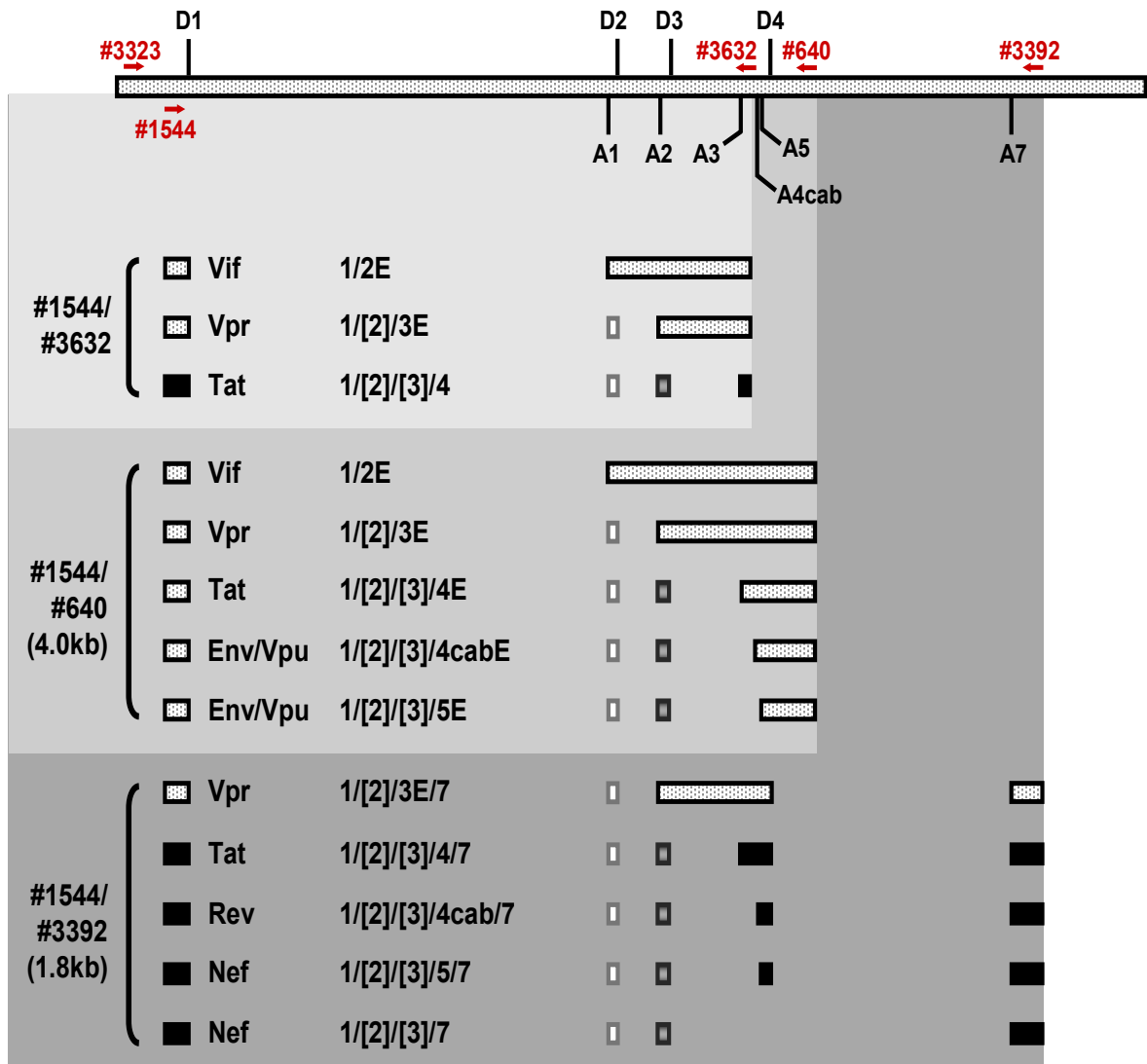


Fig.VII-1.5: Hexamer Score changes following insertion of silent ESE_{vpr} mutations in the context of ESSV-positive and -negative sequences

Z_{EI}-Score plots of the wildtype (grey) versus mutant ESE_{vpr} sequence (orange) in the context of an intact or disrupted ESSV. The mutated nucleotides are indicated by arrows.

A

[continued on next page]

B

Viral isoform	tat-mRNA class		Viral isoform	1.8kb class		Viral isoform	4.0kb class	
	#1544/ #3632	#3323/ #3632		#1544/ #3392	#3323/ #3392		#1544/ #640	#3323/ #640
Tat1 [1.4]	186	362	Nef1 [1.7]	212	388	Env1 [1.5E]	602	778
Tat2 [1.2.4]	236	412	Nef2 [1.5.7]	281	457	Env2 [1.4bE]	618	794
Tat3 [1.3.4]	260	436	Rev1 [1.4b.7]	297	473	Env3 [1.4aE]	624	800
Tat4 [1.2.3.4]	310	486	Rev2 [1.4a.7]	303	479	Env4 [1.4cE]	642	818
Vpr1 [1.3E]	573	749	Rev3 [1.4c.7]	321	497	Env5 [1.2.5E]	652	828
Vpr2 [1.2.3E]	623	799	Nef3 [1.2.5.7]	331	507	Env6 [1.2.4bE]	668	844
Vif1 [1.2E]	1050	1226	Rev4 [1.2.4b.7]	347	523	Env7 [1.2.4aE]	674	850
			Rev5 [1.2.4a.7]	353	529	Env8 [1.3.5E]	676	852
			Nef4 [1.3.5.7]	355	531	Env9 [1.2.4cE]	692	868
			Rev6 [1.2.4c.7]	371	547	Env10 [1.3.4bE]	692	868
			Rev7 [1.3.4b.7]	371	547	Env11 [1.3.4aE]	698	874
			Rev8 [1.3.4a.7]	377	553	Env12 [1.3.4cE]	718	894
			Rev9 [1.3.4c.7]	395	571	Env13 [1.2.3.5E]	726	902
			Nef5 [1.2.3.5.7]	405	581	Env14 [1.2.3.4bE]	742	918
			Rev10 [1.2.3.4b.7]	421	597	Env15 [1.2.3.4aE]	728	904
			Rev11 [1.2.3.4a.7]	429	603	Env16 [1.2.3.4cE]	760	936
			Rev12 [1.2.3.4c.7]	445	621	Tat5 [1.4E]	795	971
			Tat1 [1.4.7]	480	656	Tat6 [1.2.4E]	845	1021
			Tat2 [1.2.4.7]	530	706	Tat7 [1.3.4E]	869	1045
			Tat3 [1.3.4.7]	554	730	Tat8 [1.2.3.4E]	919	1095
			Tat4 [1.2.3.4.7]	604	780	Vpr3 [1.3E]	1180	1356
			Vpr1 [1.3E.7]	867	1043	Vpr4 [1.2.3E]	1230	1406

Fig.VII-1.6: Primer positions and PCR products sizes for viral mRNAs

(A) Positions of primers within pNL4-3. (B) PCR product sizes of viral mRNA isoforms after amplification with the respective primer pairs.

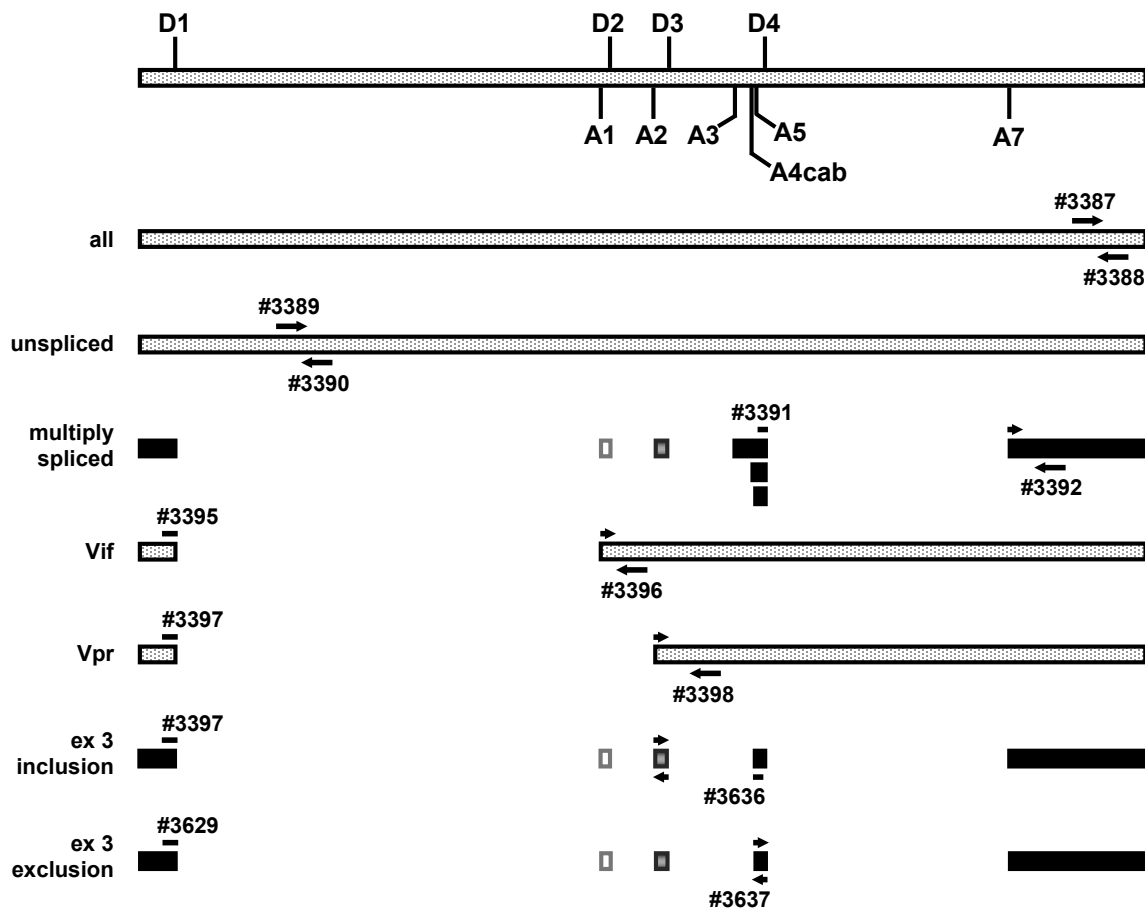


Fig.VII-1.7: Schematic representation of the primers used in quantitative RT-PCR assays (indicated by arrows)

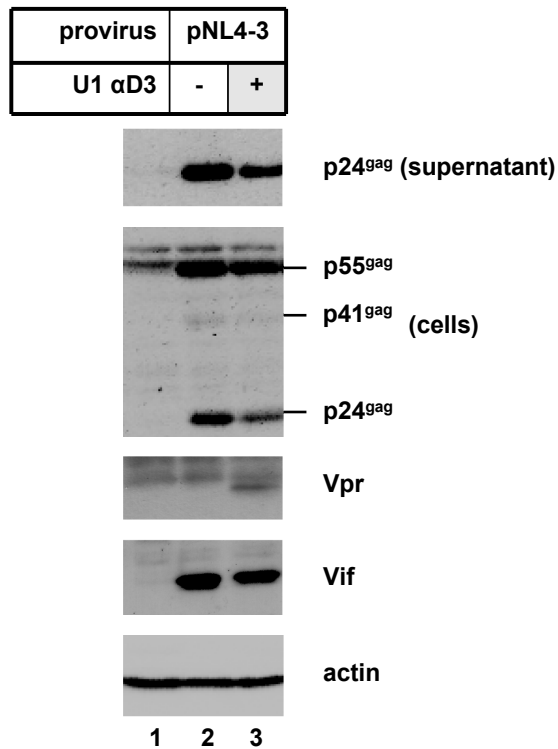


Fig.VII-1.8: Coexpression of a 5'-end modified U1 snRNA promotes Vpr expression

Western Blot analysis of cellular lysates and supernatants of HEK 293T cells either expressing pNL4-3 alone or together with the modified U1 snRNA. Samples were separated by 12% SDS-PAGE, electroblotted to a nitrocellulose membrane and probed with antibodies specifically detecting viral proteins p24^{gag}, Vif and Vpr and cellular β -actin as internal standard.

VII.2 Abbreviations and units

VII.2.1 Abbreviations

3' ss	3' splice site
5' ss	5' splice site
Ac	Acetate
AIDS	Acquired Immunodeficiency Syndrome
Amp	Ampicillin
ATP	adenosine-5'-triphosphat
BSA	bovine serum albumin
CTP	cytidine-5'-triphosphate
d	<i>downstream</i>
DAPI	4',6-diamidino-2-phenylindole
ddH ₂ O	deionised and distilled water
DMDC	dimethyl-dicarbonate
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
<i>env</i>	gene for the viral membrane protein (envelope)
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
EtBr	ethidium bromide (3,8-Diamino-6-ethyl-5-phenylphenatridiumbromid)
FCS	fetal calf serum
<i>gag</i>	gene for the viral structural proteins (group specific antigen)
	gp glycoprotein
GTP	guanosine-5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hGH	human growth hormone

Appendix

HIV-1	Human Immunodeficiency Virus Type 1
hnRNP	heterogeneous nuclear ribonucleoprotein particle
LB	Luria Broth base
LTR	long terminal repeat
mRNA	messenger ribonucleic acid
ORF	open reading frame
ori	origin of replication
pA	polyadenylation signal
PBS	phosphate buffered saline
PBS _{def}	Dulbecco's phosphate buffered saline deficient in Ca ²⁺ and Mg ²⁺
PDE	position-dependent <i>cis</i> -acting splicing enhancer element
PCR	polymerase chain reaction
PMSF	phenylmethane-sulfonyl-fluoride
POD	peroxidase
<i>pol</i>	gene for the viral enzymes (polymerase)
poly(A)+	polyadenylated
<i>rev</i>	gene for the viral protein Rev (regulator of viral protein expression)
RNA	ribonucleic acid
RNase	ribonuclease
RRE	Rev-responsive element
RRM	RNA recognition motif
RS	arginine/serine-rich
SA	splice acceptor
SD	splice donor
SELEX	Systematic Evolution of Ligands by Exponential enrichment
SDS	sodium dodecyl sulfate
SR	serine/arginine-rich
SRE	splicing regulatory element
SRSF	serine/arginine-rich splicing factor
SU	viral surface envelope protein
SV40	Simian Virus 40

<i>tat</i>	gene for the viral protein Tat (Trans-Activator of Transcription)
TE	Tris-EDTA buffer
TM	viral transmembrane envelope protein
Tris	Tris-(hydroxymethyl)-aminomethane
TTP	thymidine-5'-triphosphate
u	<i>upstream</i>
UV	ultraviolet
<i>vif</i>	gene for the viral protein Vif (Viral Infectivity Factor)
<i>vpr</i>	gene for the viral protein Vpr (Viral Protein R)
v/v	volume per volume
w/v	weight per volume

VII.2.2 Units

A	ampere
bp	base pairs m meter
°C	degree Celsius
min	minutes
M	molar
g	gram
n	nano (10^{-9})
h	hour nt nucleotide
kb	kilobase
kDa	kilodalton
rpm	rotations per minute
l	liter
sec	second
μ	micro (10^{-6})
U	unit
m	milli (10^{-3})
V	volt

VII.3 Publications

The results of the work presented here were in part published in scientific journals or presented at international conferences.

List of Publications

Erkelenz, S*, W.F. Mueller*, M.S. Evans, A. Busch, K. Schöneweis, K.J. Hertel, and H. Schaal. 2013. Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA* **19**:96-102

*These authors contributed equally to this work

Erkelenz, S., G. Poschmann, S. Theiss, A. Stefanski, F. Hillebrand, M. Otte, K. Stühler, and H. Schaal. 2012. Tra2-mediated recognition of HIV-1 5'ss D3 as a key factor in processing vpr-mRNA. *Journal of virology* [Epub ahead of print]

Widera, M., S. Erkelenz, F. Hillebrand, K. Krikoni, D. Widera, W. Kaisers, C. Deenen, M. Gombert, R. Dellen, T. Pfeiffer, B. Kaltschmidt, C. Münk, V. Bosch, K. Köhrer, and H. Schaal. 2012. An intronic G-run within HIV-1 intron 2 is critical for splicing regulation of *vif*-mRNA. *Journal of Virology* [Epub ahead of print]

Asang, C., S. Erkelenz, and H. Schaal. 2012. The HIV-1 major splice donor D1 is activated by splicing enhancer elements within the leader region and the p17-inhibitory sequence. *Virology* **432**:133-145.

Voelker, R.B., S. Erkelenz, V. Reynoso, H. Schaal, and J.A. Berglund. 2012. Frequent gain and loss of intronic splicing enhancers during the evolution of vertebrates. *Genome Biology and Evolution* **4**: 659-674.

Singh, K. K.*, S. Erkelenz*, S. Rattay, A. K. Dehof, A. Hildebrandt, K. Schulze-Osthoff, H. Schaal, and C. Schwert. 2010. Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *RNA* **16**:2442-2454.

*These authors contributed equally to this work

Zychlinski, D., S. Erkelenz, V. Mehlhorn, C. Baum, H. Schaal, and J. Bohne.

2009. Limited complementarity between U1 snRNA and a retroviral 5' splice site permits its attenuation via RNA secondary structure. *Nucleic acids research*

37(22):7429-7440.

Meeting abstracts

Erkelenz, S., S. Rosin, and H. Schaal. 2008. Re-enaction of HIV-1 alternative splicing regulation by simultaneous application of an enhanced MS2- and PP7-tethering system. 18th Annual Meeting Society of Virology. Heidelberg, March 2008. (Talk)

VII.4 Curriculum Vitae

Steffen Erkelenz

PERSONAL INFORMATION

DATE OF BIRTH 16th July 1980
 PLACE OF BIRTH Krefeld (GER)
 NATIONALITY German
 MARITAL Single

CONTACT

Küperstr.2
 47798 Krefeld
 Germany
 mobile: 0176-30410485
 email: steffen.erkelenz@gmx.de

EDUCATION

11/2006 – today	PhD Thesis – Institute for Virology , HHU Duesseldorf, GER Topic: Regulation of human and HIV-1 splice sites.
12/2005 – 10/2006	Diploma Thesis – Institute for Virology , HHU Duesseldorf, GER Topic: Exon Signals in Exon Recognition Degree: 1.0 ≈ excellent
10/2001 – 10/2006	Heinrich-Heine Universität (HHU) Duesseldorf , GER Final degree Ø 1.16 ≈ excellent Subject: Biology Major: Microbiology, Virology, Immunology
06/1999	A-levels , 'Abitur', Ø 1.9 ≈ very good

Travel Grant

04/2011 –06/2011	Department of Microbiology and Molecular Genetics, School of Medicine University of California, Irvine, Irvine, CA 92697, USA, Prof. Dr. Klemens J. Hertel Boehringer Ingelheim Fonds
------------------	---

VII.5 Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der "Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Steffen Erkelenz

Düsseldorf, den

Danksagung

Diese Arbeit wäre ohne die Unterstützung etlicher Menschen nicht möglich gewesen, weshalb ich an dieser Stelle die Gelegenheit ergreifen möchte, diesen zu danken.

In besonderer Weise möchte ich mich bei **Prof. Dr. Heiner Schaal** für seine stete Hilfe und Geduld bedanken, sowie die nötigen wissenschaftlichen Freiräume, auch eigene Ideen umsetzen zu können.

Danken möchte ich auch **Prof. Dr. Dieter Willbold** und **Prof. Dr. Albrecht Bindereif** für die bereitwillige Übernahme des Referats.

Ferner möchte ich mich bei **Prof. Dr. Hartmut Hengel** bedanken, für die Möglichkeit die Arbeit am Institut für Virologie fertigzustellen.

Meinen Kollegen **Corinna Asang, Alizée Borsche, Anna-Lena Brillen, Nora Diehl, Anja Flößer, Linda Hartmann, Isabel Heide, Frank Hillebrand, Claus Lenski, Tomek Ochman, Jan Otto Peter, Katrin Schöneweis, Björn Wefers und Marek Widera** möchte ich für die Hilfe und Unterstützung danken, aber allem voran die grandiose Arbeitsatmosphäre. Besonders will ich Frank und Tomek dafür danken, mich mit jeder Menge erfrischenden Unsinns, aber auch ihrer uneingeschränkten Hilfsbereitschaft in allen Lebenslagen, viele Male in die gute Laune zurückgeführt zu haben.

Weiterhin möchte ich mich bei **Jens Bauer, Daniel Degrandi, Henning Hofmann** und **Sebastian Howe** dafür bedanken, es auch ausserhalb der eigenen Arbeitsgruppe nie weit gehabt zu haben, um sich auf ein (zwei, drei...;) Bierchen treffen zu können.

Danken möchte ich ferner **Stephan Theiss, Dr. Werner Bouschen** und **Dr. Sabine Metzger** für die konstruktive Zusammenarbeit.

Allen Mitarbeiterinnen und Mitarbeitern des Institutes für Virologie und des Instituts für Mikrobiologie für die angenehme Arbeitsatmosphäre.

Ich möchte auch besonders **Prof. Dr. Klemens J. Hertel** und dem gesamten Hertel Lab für eine tolle Zeit in Kalifornien danken.

Dank gebührt auch all meinen Freunden ausserhalb des Instituts, für die ich mir hier gerne noch einmal Zeit nehme, sie einzeln aufzuführen: **Christian Adam, Heike Bucker, Tina Dienst, Elvis Dolff, Peter Hambloch, Christian Hartmann, Jens Kirchhoff, Lars Krafft, William F. Mueller, Tobias Rosing, Alexander Schmidt, Kerstin Seidel, Jan Sporing, Yvonne Stratemann, Kathleen Weise, Andreas Wilden**. Selbstverständlich möchte ich mich auch bei all jenen bedanken, die ich eigentlich hoffe, hier nicht vergessen zu haben;)

Zuletzt, aber vor allem, will ich meiner ganzen Familie danken:

Meinen Eltern, ohne deren bedingungslosen Rückhalt diese Zeit nicht zu schaffen gewesen wäre.

Meinem Bruder, zu dem ich aufschaue.

Meiner ganzen restlichen Familie, die man sich selber nicht besser zusammenstellen könnte.