5112–5119 *Nucleic Acids Research, 2005, Vol. 33, No. 16*
doi:10.1093/nar/gki824

Extended base pair complementarity between U1 snRNA and the 5′ splice site does not inhibit splicing in higher eukaryotes, but rather increases 5′ splice site recognition

Marcel Freund†, Martin J. Hicks[1,†], Carolin Konermann, Marianne Otte, Klemens J. Hertel[1] and Heiner Schaal*

Institut für Virologie, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, Geb. 22.21, D-40225 Düsseldorf, Germany and [1]Department of Microbiology and Molecular Genetics, School of Medicine, University of California, Irvine, Irvine, CA 92697, USA

Received July 6, 2005; Revised and Accepted August 23, 2005

ABSTRACT

Spliceosome formation is initiated by the recognition of the 5′ splice site through formation of an RNA duplex between the 5′ splice site and U1 snRNA. We have previously shown that RNA duplex formation between U1 snRNA and the 5′ splice site can protect pre-mRNAs from degradation prior to splicing. This initial RNA duplex must be disrupted to expose the 5′ splice site sequence for base pairing with U6 snRNA and to form the active spliceosome. Here, we investigated whether hyperstabilization of the U1 snRNA/5′ splice site duplex interferes with splicing efficiency in human cell lines or nuclear extracts. Unlike observations in *Saccharomyces cerevisiae*, we demonstrate that an extended U1 snRNA/5′ splice site interaction does not decrease splicing efficiency, but rather increases 5′ splice site recognition and exon inclusion. However, low complementarity of the 5′ splice site to U1 snRNA significantly increases exon skipping and RNA degradation. Although the splicing mechanisms are conserved between human and *S.cerevisiae*, these results demonstrate that distinct differences exist in the activation of the spliceosome.

INTRODUCTION

A critical step in splicing is the nucleotide-specific recognition of the exon and intron boundaries. For coding exons, this guarantees maintenance of the open reading frame, for non-coding exons it preserves *cis*-acting sequences that can influence the properties of the mRNA, e.g. their stability (1,2). Thus, it is not surprising that 5′ splice site recognition seems to be subjected to a proofreading mechanism in which the 5′ splice site must be consecutively recognized by two different snRNAs. Initially, an RNA duplex is formed between the 5′ splice site and U1 snRNA, which is primarily thought to promote the association of a U2 snRNP complex with the branchpoint region (3) through the activities of the DExD/H family RNA-dependent ATPase hPrp5 (4). However, *in vitro* binding of the U2 snRNP to the branchpoint region can occur independent of U1 snRNP binding to the 5′ splice site (5).

Subsequent to U2 snRNP binding, the tri-snRNP complex consisting of U4/U6·U5 snRNP carries out its role within the inactive spliceosome. As part of this tri-snRNP complex, the U6 snRNA is associated with U4 snRNA by two RNA helices, which must be unwound prior to base pairing between the 5′ splice site and U6 snRNA (6). Because the U6 snRNA and U1 snRNA binding sites with the 5′ splice site are mutually exclusive, the initial RNA duplex between the 5′ splice site and the U1 snRNA must be disrupted prior to forming the active spliceosome (7). The rearrangements between the pre-mRNA and snRNAs are catalyzed by a large family of conserved, ATP-dependent proteins, termed the DExH/D box helicases (8). The U1 to U6 snRNA transition at the 5′ splice site is a dynamic process, which involves multiple RNA annealing and dissociation activities. The elegant work by Staley and Guthrie (9) demonstrated that ATP and the DEAD Box Protein Prp28p are required for U1 destabilization in yeast. Thus, it was reasonable to assume that an extended

*To whom correspondence should be addressed. Tel: +49 211 81 12393; Fax: +49 211 81 12227; Email: schaal@uni-duesseldorf.de
Correspondence may also be addressed to Klemens Hertel. Tel: +1 949 824 2127; Fax: +1 949 824 8598; Email: khertel@uci.edu

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

interaction between U1 snRNA and the 5′ splice site interfered with the subsequent and mutually exclusive association of U6 snRNA with the 5′ splice site. Indeed, Staley and Guthrie (9) demonstrated that hyperstabilization of the U1 snRNA/5′ splice site interaction by extending base pairing between U1 and the 5′ splice site can lead to temperature-sensitive splicing repression in yeast. The human homolog of Prp28p, a U5 snRNP specific 100 kDa protein (U5-100K), has also been identified (10).

The strength of the U1 snRNA/5′ splice site interaction has also been implicated in the regulation of alternative splicing pattern and transcript stabilization. For example, during the early phase of HIV-1 replication the two essential viral proteins Tat and Rev are translated from multiple spliced viral transcripts. However, during the late phase of the viral life cycle, removal of the intron between the first and second coding exons of *tat* and *rev* has to be restrained since it encodes part of the glycoprotein. Thus, glycoprotein expression is only possible if splicing is inefficient enough to allow Rev-mediated nuclear export of the glycoprotein mRNA. We have previously shown that U1 snRNA binding at HIV-1 5′ splice site #4 protects the unstable glycoprotein pre-mRNA against degradation, independent of splicing (11). The formation of the stable RNA duplex between U1 snRNA and the 5′ splice site allowed efficient glycoprotein expression because the splicing efficiency of this *tat* and *rev* intron was impaired by an inefficient 3′ splice site (12,13) and its associated splicing silencer ESS3 [(14–19) and S. Kammler and H. Schaal, unpublished data].

Although numerous examples are known in which binding of U1 snRNA to a 5′ splice site is assisted by proteins belonging to the SR protein family (5,11,20–23), we have recently correlated the hydrogen bonding pattern of the RNA duplex between U1 snRNA and the 5′ splice site with U1 snRNA binding (5,23). Using an HIV-1 glycoprotein expression vector, we were able to develop an algorithm to describe functional U1 snRNA binding sites (5) in the context of an upstream bidirectional exonic splicing enhancer (23). Because the overall strength of the hydrogen bonding between U1 snRNA and the 5′ splice site correlated with glycoprotein expression only as long as glycoprotein expression was suboptimal (11), we hypothesized that maximum stabilization of the unstable glycoprotein pre-mRNA could be achieved with a subset of the potential complementarity between the 5′ splice site and U1 snRNA. To address the question whether full complementarity would result in hyperstabilization of U1 snRNA binding in higher eukaryotes, thus blocking efficient usage of the 5′ splice site and preventing the replacement of U1 by U6 snRNA, we carried out a series of *in vitro* and *in vivo* experiments.

Here, we investigated whether hyperstabilization of the U1/5′ splice site interaction interferes with the splicing efficiency in human cell lines or nuclear extracts. In contrast to the phenomena observed in *Saccharomyces cerevisiae*, our studies show that an extended U1 snRNA/5′ splice site interaction did not decrease splicing efficiency, but rather improved efficiency and increased 5′ splice site recognition. However, low complementarity increased exon skipping and mRNA degradation. Furthermore, an increase in the potential base pairing interaction between U6 snRNA and the 5′ splice site did not appear to significantly influence exon recognition.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were synthesized and purified as described previously (24). The oligonucleotides are as follows: 814, 5′-CATCAAGCTTCTCTATCAAAGCAGTACGTATTACATGTAATGCA; 1084, CATCCTTAAGCTTCTCTATCTATCAGGTAAGTATTACATGTAATGCA; 1119, 5′-CGGGGTAACATCC; 1120, 5′-TTAAGGATGTTACCCCGAGCT; 1142, 5′-TTAAGGATGCTACCTCGAGCT; 1143, 5′-CGAGGTAGCATCC; 1224, 5′-TCTTCCAGCCTCCCATCAGCGTTTGG; 1225, 5′-CAACAGAAATCCAACCTAGAGCTGCT; 1542, 5′-CACCTTCTTCTTCTATTCCTT; 1544, 5′-CTTGAAAGCGAAAGTAAAGC; 1834, 5′-CGAGAGCTCATCAGAACAGTCAGACTCATCAAGCTGCTCTATCAAAGCAGTAACACGTACATGTAATG; 1979, 5′-CCAGGTTCTACAC; 1980, 5′-TTAAGTGTAGAACCTGGAGCT; 1981, 5′-CCAGGTAAGTAGC; 1982, 5′-TTAAGCTACTTACCTGGAGCT.

Recombinant plasmids

The SV40 early *env* expression vector SV E/X tat$^-$rev$^-$ carrying the EcoRI–XhoI fragment of pNLA1 (25) and the 5′ splice site mutations were constructed as described previously (11). SVcrev was constructed by recloning the EcoRI–XhoI fragment of pUHcrev (26) into pSVT7. Env expression vector SV-env/GAR and the 5′ splice site mutations were constructed as described previously (23). All PCR-amplified sequences were confirmed by DNA sequencing. The sequences of all constructs are available on request.

Cell culture and transfection

HeLa-T4$^+$ cells (27) were propagated and transfected as described previously (11) with FuGENE 6 (Roche Molecular Biochemicals), and the transfection efficiency was monitored by cotransfection of pGL3-control (Promega). The medium was changed after 24 h, and cells were harvested and prepared 48 h after transfection. For RT–PCR, cells were transfected with 1 μg of the respective SV-1-env plasmid and 1 μg of pXGH5 (28) as a control for transfection efficiency, and total RNA was isolated after 30 h.

Western blot analysis

Cells were scraped from six-well plates into the medium, sedimented at 12 000 *g* for 14 s, washed twice in phosphate-buffered saline (PBS) and suspended in 200 μl of SDS–PAGE sample buffer (29). An aliquot from the PBS washing step was analyzed for luciferase activity (luciferase assay system; Promega). The protein concentration was measured by a Bradford protein assay (Bio-Rad) and adjusted in each sample to equal amounts of luciferase activity and protein amount by adding extracts of mock-transfected cells. Samples were subjected to electrophoresis on SDS–7% polyacrylamide gels and transferred to a polyvinylidene difluoride membrane (0.45 μm pore size, Immobillon™ P; Millipore). Protein detection was performed with a monoclonal mouse anti-gp120 antibody and visualized by a chemiluminescence detection system [ECL system and ECL hyperfilm (Amersham)].

RT–PCR assay

Isolation of total RNA was performed using a modified guanidinium isothiocyanate protocol (30). Cells were washed twice with 2 ml of PBS and cell lysis was performed with 500 µl of buffer D [7.6 µl of 2-mercaptoethanol, 50 µl of 3 M sodium acetate (pH 4)], 500 µl of phenol and 100 µl of a chloroform-isoamyl alcohol mixture (24:1) were added and mixed for 15 s. After incubation on ice for 15 min, phases were separated by centrifugation (10 000 r.p.m., Eppendorf centrifuge). RNA was precipitated in 1 vol of isopropanol overnight. After centrifugation (10 000 r.p.m., Eppendorf centrifuge) was washed twice with 70% ethanol and dissolved in 10 µl of DMDC-ddH_2O. Prior to RT, 4 µl of RNA samples were subjected to DNase I digestion using 10 U DNase I (Roche Molecular Biochemicals) with 50 mM Tris, pH 7.5 and 10 mM $MgCl_2$ in a total volume of 10 µl at room temperature for 1.45 h. After DNase I inactivation at 95°C for 10 min, 4.5 µl of the DNase digested RNA samples were reversed transcribed with 200 U SuperScript III RNase H-Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol using 0.375 mM oligo$(dT)_{15}$ (Roche Molecular Biochemicals) as primer. PCR was carried out with 1.25 U AmpliTaq (Applied Biosystems) in a total of 50 µl according to the manufacturer's protocol in a Robocycler Gradient 96 Temperature Cycler (Stratagene). All primers were used at a final concentration of 0.2 µM. Double-spliced and skipped RNA was detected with primer pair 1544/1542. PCR products were separated on 6% non-denaturing polyacrylamide gels, stained with ethidium bromide for 10 min and visualized with the Lumi-Imager F1 (Roche Molecular Biochemicals). For the detection of hGH mRNA in the RT–PCR assay primer pair 1225/1224 was used.

In vitro splicing

Reactions were carried out in 30% nuclear extract, containing T7 generated ^{32}P-labeled RNA transcripts, 1.0 mM ATP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$ and incubated at 30°C for a determined time course indicated in figure legends. Following incubation, reactions were proteinase K digested, phenol–chloroform extracted, ethanol precipitated, resolved on 8% denaturing PAGE and measured by PhosphorImager analysis (Bio-Rad) as described previously (31). Percent spliced is defined as spliced product/(unspliced product + spliced product). To derive kinetic rate constants, time points were fit to a first order rate description for product appearance. Rates of degradation were computed by fitting the total counts per lane to a first order decay rate description. Prior to fitting, the counts detected in each lane were normalized to input (loading volume). The background was determined individually for each lane, as the total number of substrate and product counts varied throughout the time course. Rate determinations were repeated at least three times.

RNAs used for *in vitro* splicing

DNA templates encoding regions of the first two exons (excluding the 5′ splice site of exon 2) of β-globin were cloned downstream of the T7 promoter in the SP73 plasmid. Alternative 5′ splice site cassettes were cloned into the location of the original 5′ splice site of exon 1, and potential cryptic 5′ splice sites (those containing a GT) were mutated within the exonic and intronic regions surrounding the 5′ splice site of exon 1. These β-globin minigenes were subsequently linearized. ^{32}P-labeled RNA transcripts were synthesized by T7 RNA polymerase and resolved and purified on 4% denaturing PAGE.

RESULTS

Hyperstabilization of the U1 snRNA/5′ splice site interaction maintains efficient splicing and promotes RNA stability

To determine how the efficiency of pre-mRNA splicing is influenced by hyper- and hypostabilization of the RNA duplex formed between U1 snRNA and the 5′ splice site, we have made use of HIV-1 *env* expression vectors. In the absence of the viral regulatory protein Rev these vectors are one-intron splicing reporters (Figure 1A). We performed splicing assays for a series of 5′ splice sites with increasing complementary with U1 snRNA (Figure 1B). Following transient transfection of HeLa-T4$^+$ cells, we detected measurable levels of spliced mRNAs only for the substrates with 5′ splice sites that can form 14 or more hydrogen bonds with U1 snRNA (Figure 1C, lanes 2–5, 9 and 10), but not with fewer (Figure 1C, lane 1).

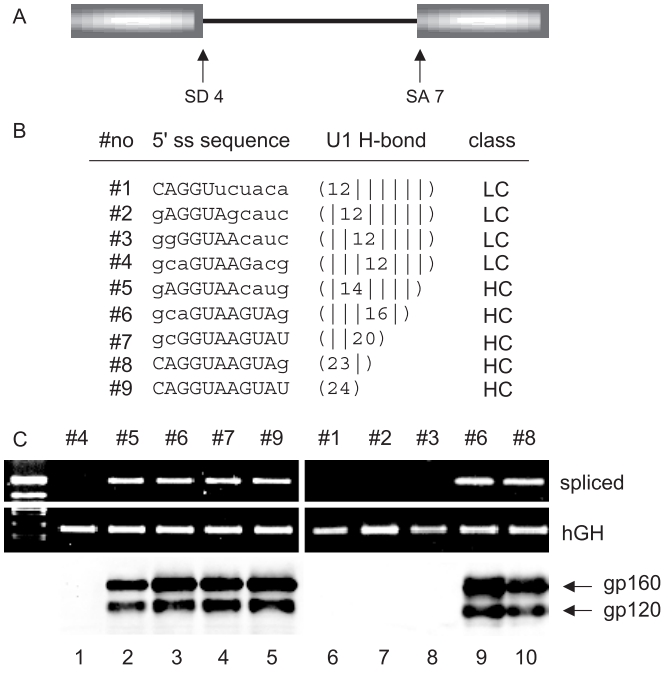

Figure 1. Full complementarity between the 5′ splice site and U1 snRNA maintains efficient splicing and RNA stability. (**A**) Schematic drawing of the HIV-1 pre-mRNA transcripts synthesized from *env* expression vectors. (**B**) Sequences of the 5′ splice sites analyzed. Mismatches to U1 snRNA are given in lower case letters. Possible H-bonds (U1 H-bond) and 'HC'/'LC' classification were calculated using the splicefinder algorithm (www.splicefinder.net) (5). (**C**) RT–PCR assay (upper panel) and immunoblot analysis of glycoproteins (lower panel) of HeLa-T4$^+$ cells transfected with *env* expression vectors carrying mutations in the SD4 5′ splice site and pXGH5 expressing the hGH mRNA as an internal control. For the immunoblot analysis SVc*rev* and pGL3-control instead of pXGH5 was cotransfected as described previously (5). The numbers above each lane indicate the 5′ splice site tested; each lane is numbered below the image.

Significantly, pre-mRNAs capable of maximal base pairing with U1 snRNA (11 bp, 24 hydrogen bonds) were spliced as efficiently as any other pre-mRNA with fewer base pair interactions. Because the splicing analyses were carried out within the linear range of amplification for the viral template, we conclude that hyperstabilization of the U1 snRNA interaction with the 5′ splice site does not lead to reduced efficiency of pre-mRNA splicing under the conditions used here.

To investigate whether the lack of detection of spliced mRNA (Figure 1C, lane 1) was independent of the position of the formed RNA duplex, we analyzed additional 5′ splice site mutations that differed in the position of their continuous stretch of complementary nucleotides while maintaining 12 possible hydrogen bonds (Figure 1C, lane 6–8). Even though computational predictions of these 5′ splice sites suggested significant differences in intrinsic strength [the MaxEntScore of these four 5′ splice site varies from 0.74 to 4.95 (32)], none of them allowed detection of spliced mRNA. These results suggest that irrespective of the predicted duplex strength, a U1 snRNA/5′ splice site interaction formed by only 12 hydrogen bonds is too low to form a stable RNA duplex that initiates pre-mRNA splicing.

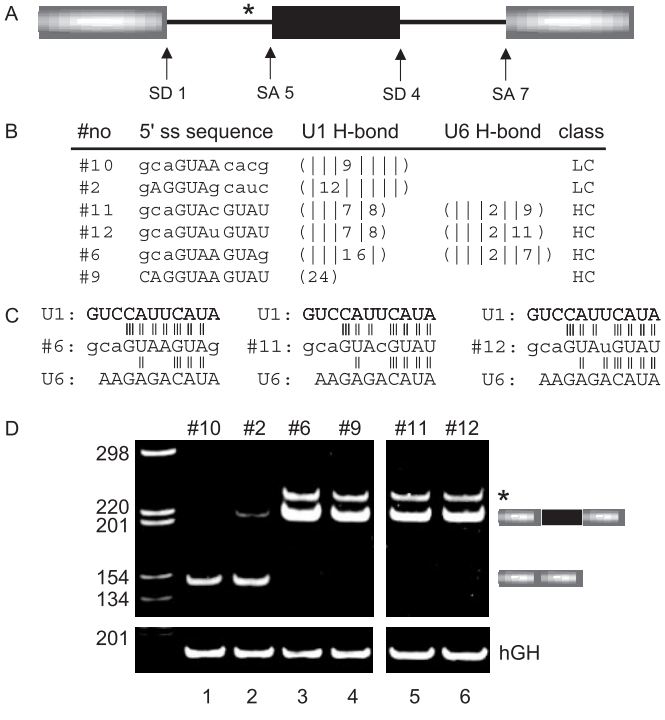

Figure 2. Full complementarity between the 5′ splice site and U1 snRNA does not compromise internal exon inclusion. (**A**) Schematic drawing of the three-exon substrates used. Each substrate contains the HIV-1 genome but lacks the U3 region of the 5′ LTR, the non-coding region of the 3′ LTR and the region downstream of the *gag* start codon, as described previously (5). The asterisk indicates the alternative 3′ splice site SA4a. (**B**) Sequences of the analyzed 5′ splice sites. Mismatches to U1 snRNA are given in lower case letters. Possible H-bonds between the 5′ splice site and U1 snRNA (U1 H-bond) and U6 snRNA (U6 H-bond) are indicated. 'HC'/'LC' classifications were calculated using the splicefinder algorithm (www.splicefinder.net) (5). (**C**) 5′ splice site sequences for substrates #6, #11 and #12. All possible base pairs with U1 snRNA (above) and U6 snRNA (below) are shown. Mismatches to U1 snRNA are given in lower case letters. (**D**) RT–PCR assay of HeLa-T4$^+$ cells transfected with the test 5′ splice sites and with pXGH5 expressing the hGH mRNA as an internal control. The numbers above each lane indicate the 5′ splice site tested; each lane is numbered below the image. The numbers to the left of the image define the sizes of DNA markers. The cartoon on the right of the image defines the migration of exon included- and exon excluded spliced products. The asterisk indicates an alternative exon inclusion product using the alternative 3′ splice site SA4a. All spliced products were verified by sequence analysis.

We have demonstrated previously that in the presence of Rev the stability of the unspliced *env* mRNA depends critically on the hydrogen bonding pattern between the 5′ splice site and U1 snRNA (11). Hence, HIV-1 *env* gene expression allows monitoring U1 snRNA binding at the 5′ splice site. Therefore, we repeated the transfection experiments with cotransfection of a Rev expression vector and analyzed glycoprotein expression. In agreement with the results obtained by RT–PCR, western blot analysis of cells transfected with constructs capable of forming more than 12 consecutive hydrogen bonds with U1 snRNA showed wild-type env expression levels (Figure 1C, lower panel, lane 2–5, 9 and 10). We conclude that at least 14 hydrogen bonds are required to stabilize *env* RNA levels and that hyperstabilization of this interaction does not limit *env* protein expression.

Hyperstabilization of the U1 snRNA/5′ splice site interaction does not cause an inhibition of exon recognition

To examine how exon recognition is influenced by stabilizing the U1 snRNA/5′ splice site interaction in pre-mRNA substrates containing multiple exons, we enlarged the HIV-1 transcription unit to obtain a 2-intron HIV-1 sub-genomic *env* expression construct (Figure 2A). The 5′ splice site of the internal exon was varied as shown in Figure 2B. The splicing patterns of these constructs were analyzed using RT–PCR within the linear range of amplification for the viral template (Figure 2D, upper panel) and with an hGH substrate as a loading control (Figure 2D, lower panel). Identification of the splice junctions was obtained by sequence analysis of gel-isolated PCR fragments. More than 12 hydrogen bonds or 5 consecutive base pairs between the 5′ splice site and U1 snRNA resulted in exclusive inclusion of the internal exon (Figure 2D, lanes 3 and 4). As was observed for the single exon substrate, hyperstabilization of the U1 snRNA/5′ splice site interaction did not cause an inhibition of exon recognition or exclusion. However, reduced complementarity resulted in exon skipping (Figure 2D, lane 2), demonstrating that the recognition of the internal exon is a stochastic process (Figure 2D, lanes 1 and 2). We conclude that hyperstabilization of the interaction between U1 snRNA and the 5′ splice site favors exon inclusion without reducing the efficiency of the splicing reaction.

To assess whether the relative competition between the mutually exclusive interactions of either U1 snRNA with the 5′ splice site or U6 snRNA with the 5′ splice site influences pre-mRNA splicing patterns, additional two-intron test substrates were evaluated for splicing efficiency and RNA stability. Increasing the base pairing between U6 snRNA and the test 5′ splice site from 5 to 6 bp (from 11 to 13 hydrogen bonds), while maintaining the strength of the U1 snRNA/5′ splice site interaction (compare Figure 2B, #11 and #12, and Figure 2C), did not generate detectable changes in the levels of exon inclusion (Figure 2D, compare lane 5 with lane 6) or of the stability of the unspliced *env* mRNA (data not shown).

We conclude that the degree of U6 snRNA interaction with the 5′ splice site does not significantly influence the observed splicing patterns.

Extended base pairing between U1 snRNA and the 5′ splice does not lead to inhibition of pre-mRNA splicing *in vitro*

Hyperstabilization of the interaction between U1 snRNA and the 5′ splice site in yeast was observed to be strikingly inhibitory when assayed at lower temperatures (9). To evaluate the effects that extended base pairing between U1 snRNA and the 5′ splice site has on pre-mRNA splicing at different temperatures an *in vitro* splicing assay was used. Test pre-mRNAs containing variable 5′ splice sites were assayed at different reaction temperatures in HeLa cell nuclear extract to evaluate their kinetics of splicing. As illustrated in Figure 3 and summarized in Table 1, pre-mRNAs containing 5′ splice sites with extended U1 snRNA base pairing were active at the highest and lowest temperatures tested. In contrast, pre-mRNAs containing reduced U1 snRNA base pairing were active only at lower temperatures. At the highest temperature tested (37°C), the efficiency of pre-mRNA splicing was reduced significantly. Furthermore, the data summarized in Table 1 support the above conclusion that the magnitude of U6 snRNA interaction with the 5′ splice site does not influence the efficiency of splicing in the context of the pre-mRNAs tested. While the complementarity between the 5′ splice site and U1 snRNA significantly influenced the steady-state levels of the test substrate (Figure 3), the *in vitro* splicing analysis detected only minor differences in the stability of the test RNA. We conclude that extended base pairing between U1 snRNA and the 5′ splice does not lead to inhibition of pre-mRNA splicing, even at the lowest temperature tested (21°C).

DISCUSSION

The analysis presented here does not support the notion that increased complementarity between the 5′ splice site and U1 snRNA is inhibitory to pre-mRNA splicing in mammalian cell culture or *in vitro* splicing assays. Rather, extended base pair interactions with U1 snRNA led to increased exon inclusion, higher efficiencies of pre-mRNA splicing and greater overall

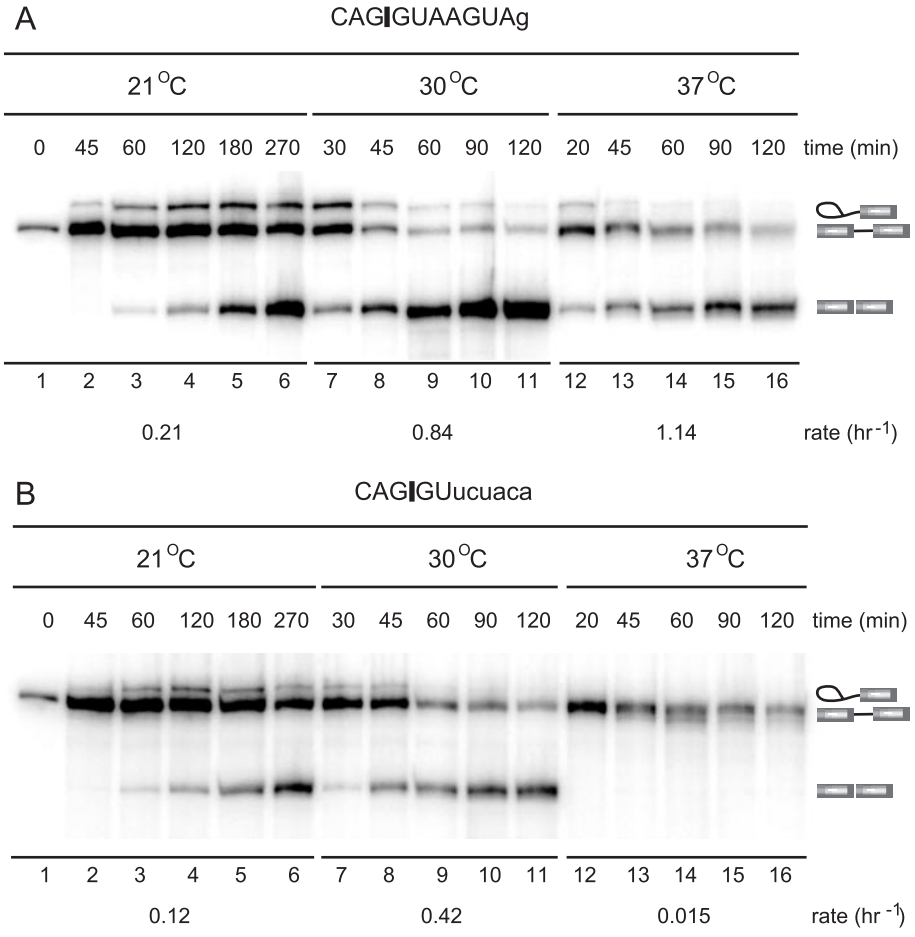

Figure 3. *In vitro* splicing efficiency of pre-mRNAs with hyper- and hypostabilized interaction with U1 snRNA. Representative reaction profiles are shown for pre-mRNAs with an extended U1 snRNA interaction potential (**A**) and a pre-mRNA with a low U1 snRNA interaction potential (**B**). Each time course was used to determine observed rates of splicing at each temperature tested. The average rate of splicing is shown below each time course. Rate determinations were repeated at least three times. The nucleotide sequence above each autoradiogram represents the 5′ splice site sequences of the test pre-mRNA. Capital letters symbolize base pairing with U1 snRNA; small letters symbolize mismatched nucleotides with U1 snRNA. The vertical line (|) indicates the intron/exon junction. The identities of spliced and unspliced RNAs are indicated on the right.

Table 1. Rates of splicing[a] and degradation[a] for pre-mRNA substrates with variable 5′ splice site strengths

5′ Splice site of substrate[b]	U1 Match[c]	U6 Match[c]	21°C (h^{-1})	30°C (h^{-1})	37°C (h^{-1})	Ratio of 30°C/21°C	Degradation 37°C (h^{-1})
CAG|GUAAGUAg	23|	||2||7|	0.21	0.84	1.14	4	0.36
CAG|GUAAGUca	21||	||2||5||	0.34	1.26	0.31	4.75	1
CAG|GUugGUAU	12||8	||3|9	0.3	0.78	0.036	2.6	0.72
CAG|GUucuaca	12|||||	||3||||	0.12	0.42	0.015	3.5	0.6
gcaIGUAcguau	||7|8	||2||9	0.18	0.57	0.06	3.12	0.48
gcaIGUAuguau	||7|8	||2|11	0.12	0.42	0.016	3.5	0.54

[a]Each value was determined from at least two independent splicing reactions. Experimental error for each rate determination is within 20%. Rates determined from different experiments varied within 2-fold.
[b]The vertical line (|) indicates the intron/exon junction. 5′ Splice site sequences are listed from the −3 to the +8 position. Capital letters symbolize base pairing with U1 snRNA; small letters symbolize mismatched nucleotides with U1 snRNA.
[c]The number of hydrogen bonds between the 5′ splice sites and U1 snRNA or U6 snRNA were calculated according to Freund *et al*. (5). The vertical lines (|) designate a mismatched nucleotide position.

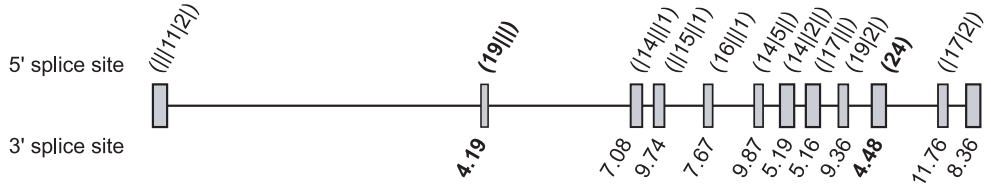

Figure 4. Schematic drawing of the exon–intron structure of PLA2G7 platelet-activating factor acetylhydrolase. H-bond patterns of the 5′ splice sites calculated according to the splicefinder algorithm (5) are shown above the exons (open boxes), the calculated strength of the 3′ splice sites as MaxEntScores (32) below the respective exons.

stability of the pre-mRNA. These conclusions are consistent with a recent analysis of cryptic splice site activation, which demonstrated that the complementarity of the 5′ splice site to the 5′ end of the U1 snRNA was the dominant parameter in determining the extent of splicing at the aberrantly activated 5′ splice site (33). However, 5′ splice site recognition is also influenced by the activities of protein components within U1 snRNP, as demonstrated by functional SELEX experiments in the absence of the 5′ end of U1 snRNA (34–37). It was suggested that other factors such as the U1 snRNP specific protein U1-C contribute to sequence specificity. Similarly, the U5 snRNP specific protein Prp8 is likely to stabilize the interaction of U5 snRNA with the 5′ splice site. These proposals were supported by the observations that U1-C (21,38) as well as Prp8 cross-link to the 5′ splice site (39–44). Despite these additional protein contributions to the recognition of the 5′ splice site, the dominance of the RNA/RNA in the activation of the 5′ splice site has been shown unequivocally in many cases by restoring the functionality of mutated 5′ splice sites in the presence of compensatory changes in U1 snRNAs (5,11,45,46).

A survey of annotated 5′ splice sites of human chromosome 6 revealed the presence of 0.1% naturally occurring 5′ splice sites with a perfect match to all 11 nt of the 5′ end of U1 snRNA (H. Schaal, unpublished data). Interestingly, when calculating the strength of all splice sites of a randomly chosen gene, PLA2G7 [(platelet-activating factor (PAF) acetylhydrolase], we found that the constitutively spliced exons #2 and #10 which carry the weakest 3′ splice sites [MaxEntScore (32) <5.0] are flanked by 5′ splice sites of higher-than-average complementary (Figure 4). Although speculative, these observations suggest that highly complementary 5′ splice sites might compensate for weak upstream 3′ splice sites to ensure exon recognition in higher eukaryotes.

In agreement with previous studies (23), we show that hypostabilized 5′ splice sites are discriminated against. This may be due to insufficient binding of U1 snRNP to the 5′ splice site, thus resulting in reduced recognition of the 5′ splice site and, consequently, to higher frequencies of exon exclusion. As expected from thermodynamic considerations for RNA duplexes, the reduced recognition of a splice site with decreased U1 snRNA complementarity could be overcome by lowering the temperature to 30°C (Figure 3B). These results are consistent with the temperature-dependent splicing of human β-globin (47) and human collagen substrates (48), where correct splicing could be restored by incubating cells at decreased temperature.

The observation that an increased complementarity between the 5′ splice site and U1 snRNA is not inhibitory to pre-mRNA splicing in higher eukaryotes, even at lower temperatures, highlights that the replacement reaction of U1 snRNP with U6 snRNP during spliceosome activation is another distinction between the splicing mechanisms employed by higher eukaryotes and *S.cerevisiae*. In addition, we find that unlike the observation made in *S.cerevisiae*, the degree of U6 snRNA interaction with the 5′ splice site did not influence the splicing patterns significantly in the mammalian system analyzed here. Increasing the base pairing between U6 snRNA and the 5′ splice site, while maintaining the strength of the U1 snRNA/5′ splice site interaction, did not generate detectable changes in the levels of exon inclusion or splicing efficiency (Figure 2). Comparable mutations, allowing additional base pair interactions at position +4 of the 5′ splice site, were also tested in a human β-globin competition assay (33). In agreement with our data, mutations that activated splicing at the cryptic 5′ splice site correlated with increased stability of the RNA duplex between the 5′ splice site and the U1 snRNA, regardless of their level of complementarity to U6 snRNA (33). However, in

a few reported cases U6 snRNA, rather than U1 snRNA, seemed to direct the position of the 5′ splice site. Nevertheless, selection always depended on the vicinity of a U1 binding site (49,50).

The observation that hyperstabilization of the U1 snRNA interaction with the 5′ splice site did not interfere with pre-mRNA splicing efficiency demonstrates that the transition of 5′ splice site interactions with U1 snRNP to interactions with U6 snRNP within the mammalian spliceosome is not at risk of being rate limiting. These considerations suggest that the helicase(s) catalyzing the U1/U6 transition is(are) substantially more efficient in executing its function when compared with other RNA/RNA or RNA/protein rearrangements or catalytic steps of the mammalian splicing reaction. The high efficiency of the U1/U6 transition in mammals could be explained by the involvement of more than one helicase participating in the rearrangements at the 5′ splice site. Indeed, cross-linking experiments revealed an interaction of a 65 kDa protein (p65) with the U1/5′ splice site RNA duplex during spliceosome assembly (51). This p65 protein was later identified as the essential p68 RNA helicase (52). Depletion of endogenous p68 RNA helicase in HeLa nuclear extracts did not affect the binding of U1 snRNP to the 5′ splice site, whereas dissociation of U1 snRNP from the 5′ splice site was largely inhibited (52). These data suggested that at least two RNA helicases (U5-100k, p68) function in destabilizing the U1 snRNA/5′ splice site interaction in the mammalian splicing reaction. Thus, it is possible that the involvement of multiple helicases could explain why the mammalian U1/U6 transition of hyperstabilized 5′ splice sites appears to be more efficient than the yeast counterpart.

Furthermore, Chen and co-workers (53) have shown that specific mutations that alter the U1 associated protein U1-C can bypass the requirements for Prp28p during the U1/U6 transition. These results suggest that Prp28p counteracts the stabilizing effects of U1-C on the U1/5′ splice site RNA duplex (53) rather than disrupting the base pairs between U1 snRNA and the 5′ splice site. In mammalian cells, it is possible that p68 and U5-100K cooperate in destabilizing the interaction between U1 snRNP and the 5′ splice site, one acting on the protein factors and the other acting directly on the RNA duplex.

S.cerevisiae contains relatively few intron-containing genes. It is possible that yeast may have evolved a less flexible and smaller range of variance in the 5′ splice site nucleotide sequence to optimize its initial recognition by U1 snRNP and the subsequent replacement by U6 snRNP. In mammals, the kinetics at this rearrangement step is not rate limiting, thus permitting a greater deviation in recognizable 5′ splice site signals, from weak to strong (according to U1 snRNA binding). In turn, this relaxation may permit a greater range of potential 5′ splice sites, ultimately resulting in more alternative 5′ splice site splicing patterns and increased proteomic divergence.

ACKNOWLEDGEMENTS

The authors thank K. Strebel and M. Martin for providing plasmid pNLA1, M. Gething for pSVT7, R. Axel for HeLa-T4$^+$ cells (through the MRC AIDS Directed Program Reagent Project) and Dade Behring for the anti-gp120 mAb (87–133/026). The authors are also grateful to Imke Meyer for excellent technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft grants to H.S. (SCHA 909/1-2 and SCHA 909/2-2), by support from the Heinz-Ansmann-Stiftung (H.S.), and by a National Institute of Health Grant (GM-62287) to K.J.H. Funding to pay the Open Access publication charges for this article was provided by Deutsche Forschungsgemeinschaft.

Conflict of interest statement. None declared.

REFERENCES

1. Krummheuer,J., Lenz,C., Kammler,S., Scheid,A. and Schaal,H. (2001) Influence of the small leader exons 2 and 3 on human immunodeficiency virus type 1 gene expression. *Virology*, **285**, 276–289.
2. Sureau,A., Gattoni,R., Dooghe,Y., Stevenin,J. and Soret,J. (2001) SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J.*, **20**, 1785–1796.
3. Seraphin,B. and Rosbash,M. (1989) Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell*, **59**, 349–358.
4. Xu,Y.Z., Newnham,C.M., Kameoka,S., Huang,T., Konarska,M.M. and Query,C.C. (2004) Prp5 bridges U1 and U2 snRNPs and enables stable U2 snRNP association with intron RNA. *EMBO J.*, **23**, 376–385.
5. Freund,M., Asang,C., Kammler,S., Konermann,C., Krummheuer,J., Hipp,M., Meyer,I., Gierling,W., Theiss,S., Preuss,T. *et al.* (2003) A novel approach to describe a U1 snRNA binding site. *Nucleic Acids Res.*, **31**, 6963–6975.
6. Staley,J.P. and Guthrie,C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, **92**, 315–326.
7. Madhani,H.D. and Guthrie,C. (1994) Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.*, **28**, 1–26.
8. Luking,A., Stahl,U. and Schmidt,U. (1998) The protein family of RNA helicases. *Crit. Rev. Biochem. Mol. Biol.*, **33**, 259–296.
9. Staley,J.P. and Guthrie,C. (1999) An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell*, **3**, 55–64.
10. Teigelkamp,S., Mundt,C., Achsel,T., Will,C.L. and Luhrmann,R. (1997) The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA*, **3**, 1313–1326.
11. Kammler,S., Leurs,C., Freund,M., Krummheuer,J., Seidel,K., Tange,T.O., Lund,M.K., Kjems,J., Scheid,A. and Schaal,H. (2001) The sequence complementarity between HIV-1 5′ splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. *RNA*, **7**, 421–434.
12. Staffa,A. and Cochrane,A. (1994) The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3′ splice site. *J. Virol.*, **68**, 3071–3079.
13. O'Reilly,M.M., McNally,M.T. and Beemon,K.L. (1995) Two strong 5′ splice sites and competing, suboptimal 3′ splice sites involved in alternative splicing of human immunodeficiency virus type 1 RNA. *Virology*, **213**, 373–385.
14. Lamond,A.I., Konarska,M.M. and Sharp,P.A. (1987) A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. *Genes Dev.*, **1**, 532–543.
15. Staffa,A. and Cochrane,A. (1995) Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol. Cell. Biol.*, **15**, 4597–4605.
16. Si,Z.H., Rauch,D. and Stoltzfus,C.M. (1998) The exon splicing silencer in human immunodeficiency virus type 1 Tat exon 3 is bipartite and acts early in spliceosome assembly. *Mol. Cell. Biol.*, **18**, 5404–5413.
17. Amendt,B.A., Si,Z.H. and Stoltzfus,C.M. (1995) Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: evidence for inhibition mediated by cellular factors. *Mol. Cell. Biol.*, **15**, 4606–4615.
18. Tange,T.O. and Kjems,J. (2001) SF2/ASF binds to a splicing enhancer in the third HIV-1 tat exon and stimulates U2AF binding independently of the RS domain. *J. Mol. Biol.*, **312**, 649–662.
19. Marchand,V., Mereau,A., Jacquenet,S., Thomas,D., Mougin,A., Gattoni,R., Stevenin,J. and Branlant,C. (2002) A Janus splicing regulatory

element modulates HIV-1 tat and rev mRNA production by coordination of hnRNP A1 cooperative binding. *J. Mol. Biol.*, **323**, 629–652.
20. Puig,O., Gottschalk,A., Fabrizio,P. and Seraphin,B. (1999) Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection. *Genes Dev.*, **13**, 569–580.
21. Zhang,D. and Rosbash,M. (1999) Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. *Genes Dev.*, **13**, 581–592.
22. Del Gatto-Konczak,F., Bourgeois,C.F., Le Guiner,C., Kister,L., Gesnel,M.C., Stevenin,J. and Breathnach,R. (2000) The RNA-binding protein TIA-1 is a novel mammalian splicing regulator acting through intron sequences adjacent to a 5′ splice site. *Mol. Cell. Biol.*, **20**, 6287–6299.
23. Caputi,M., Freund,M., Kammler,S., Asang,C. and Schaal,H. (2004) A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 rev, env, vpu, and nef gene expression. *J. Virol.*, **78**, 6517–6526.
24. Schaal,H., Klein,M., Gehrmann,P., Adams,O. and Scheid,A. (1995) Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. *J. Virol.*, **69**, 3308–3314.
25. Strebel,K., Daugherty,D., Clouse,K., Cohen,D., Folks,T. and Martin,M.A. (1987) The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature*, **328**, 728–730.
26. Schaal,H., Pfeiffer,P., Klein,M., Gehrmann,P. and Scheid,A. (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.
27. Maddon,P.J., Dalgleish,A.G., McDougal,J.S., Clapham,P.R., Weiss,R.A. and Axel,R. (1986) The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell*, **47**, 333–348.
28. Selden,R.F., Howie,K.B., Rowe,M.E., Goodman,H.M. and Moore,D.D. (1986) Human growth hormone as a reporter gene in regulation studies employing transient gene expression. *Mol. Cell. Biol.*, **6**, 3173–3179.
29. Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
30. Chomczynski,P. and Sacchi,N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.*, **162**, 156–159.
31. Lim,S.R. and Hertel,K.J. (2004) Commitment to splice site pairing coincides with A complex formation. *Mol. Cell*, **15**, 477–483.
32. Yeo,G. and Burge,C.B. (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.*, **11**, 377–394.
33. Roca,X., Sachidanandam,R. and Krainer,A.R. (2005) Determinants of the inherent strength of human 5′ splice sites. *RNA*, **11**, 683–698.
34. Lund,M. and Kjems,J. (2002) Defining a 5′ splice site by functional selection in the presence and absence of U1 snRNA 5′ end. *RNA*, **8**, 166–179.
35. Crispino,J.D., Blencowe,B.J. and Sharp,P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866–1869.
36. Crispino,J.D. and Sharp,P.A. (1995) A U6 snRNA:pre-mRNA interaction can be rate-limiting for U1-independent splicing. *Genes Dev.*, **9**, 2314–2323.
37. Tarn,W.Y. and Steitz,J.A. (1994) SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. *Genes Dev.*, **8**, 2704–2717.
38. Rossi,F., Forne,T., Antoine,E., Tazi,J., Brunel,C. and Cathala,G. (1996) Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5′ splice site-U1 snRNP interaction. *J. Biol. Chem.*, **271**, 23985–23991.
39. Wyatt,J.R., Sontheimer,E.J. and Steitz,J.A. (1992) Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. *Genes Dev.*, **6**, 2542–2553.
40. Teigelkamp,S., Whittaker,E. and Beggs,J.D. (1995) Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res.*, **23**, 320–326.
41. Reyes,J.L., Gustafson,E.H., Luo,H.R., Moore,M.J. and Konarska,M.M. (1999) The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site. *RNA*, **5**, 167–179.
42. Siatecka,M., Reyes,J.L. and Konarska,M.M. (1999) Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes Dev.*, **13**, 1983–1993.
43. Collins,C.A. and Guthrie,C. (1999) Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev.*, **13**, 1970–1982.
44. Maroney,P.A., Romfo,C.M. and Nilsen,T.W. (2000) Functional recognition of 5′ splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol. Cell*, **6**, 317–328.
45. Zhuang,Y. and Weiner,A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell*, **46**, 827–835.
46. Lu,X.B., Heimer,J., Rekosh,D. and Hammarskjold,M.L. (1990) U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proc. Natl Acad. Sci. USA*, **87**, 7598–7602.
47. Gemignani,F., Sazani,P., Morcos,P. and Kole,R. (2002) Temperature-dependent splicing of beta-globin pre-mRNA. *Nucleic Acids Res.*, **30**, 4592–4598.
48. Weil,D., D'Alessio,M., Ramirez,F., Steinmann,B., Wirtz,M.K., Glanville,R.W. and Hollister,D.W. (1989) Temperature-dependent expression of a collagen splicing defect in the fibroblasts of a patient with Ehlers-Danlos syndrome type VII. *J. Biol. Chem.*, **264**, 16804–16809.
49. Hwang,D.Y. and Cohen,J.B. (1996) U1 snRNA promotes the selection of nearby 5′ splice sites by U6 snRNA in mammalian cells. *Genes Dev.*, **10**, 338–350.
50. Brackenridge,S., Wilkie,A.O. and Screaton,G.R. (2003) Efficient use of a 'dead-end' GA 5′ splice site in the human fibroblast growth factor receptor genes. *EMBO J.*, **22**, 1620–1631.
51. Liu,Z.R., Sargueil,B. and Smith,C.W. (1998) Detection of a novel ATP-dependent cross-linked protein at the 5′ splice site-U1 small nuclear RNA duplex by methylene blue-mediated photo-cross-linking. *Mol. Cell. Biol.*, **18**, 6910–6920.
52. Liu,Z.R. (2002) p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5′ splice site duplex. *Mol. Cell. Biol.*, **22**, 5443–5450.
53. Chen,J.Y., Stands,L., Staley,J.P., Jackups,R.R.Jr, Latus,L.J. and Chang,T.H. (2001) Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. *Mol. Cell*, **7**, 227–232.