Multiple *cis* elements regulate an alternative splicing event at 4.1R pre-mRNA during erythroid differentiation

Mireille Deguillien, Shu-Ching Huang, Madeleine Morinière, Natacha Dreumont, Edward J. Benz Jr, and Faouzi Baklouti

The inclusion of exon 16 in the mature protein 4.1R messenger RNA (mRNA) is a critical event in red blood cell membrane biogenesis. It occurs during late erythroid development and results in inclusion of the 10-kd domain needed for stabilization of the spectrin/actin lattice. In this study, an experimental model was established in murine erythroleukemia cells that reproduces the endogenous exon 16 splicing patterns from a transfected minigene. Exon 16 was excluded in predifferentiated and predominantly included after induction. This suggests that the minigene contained exon and abutting intronic sequences sufficient for splicing regulation. A systematic analysis of the *cis*-acting regulatory sequences that reside within the exon and flanking introns was performed. Results showed that (1) the upstream intron of 4.1R pre-mRNA is required for exon recognition and it displays 2 enhancer elements, a distal element acting in differentiating cells and a proximal constitutive enhancer that resides within the 25 nucleotides preceding the acceptor site; (2) the exon itself contains a strong constitutive splicing silencer; (3) the exon has a weak 5′ splice site; and (4) the downstream intron contains at least 2 splicing enhancer elements acting in differentiating cells, a proximal element at the vicinity of the 5′ splice site, and a distal element containing 3 copies of the UGCAUG motif. These results suggest that the interplay between negative and positive elements may determine the inclusion or exclusion of exon 16. The activation of the enhancer elements in late erythroid differentiation may play an important role in the retention of exon 16. (Blood. 2001;98: 3809-3816)

© 2001 by The American Society of Hematology

Introduction

Protein 4.1R is a critical 80-kd cytoskeletal protein found in circulating red blood cells (RBCs). It mediates the formation and maintenance of spectrin/actin complex and anchors the cytoskeleton to the overlying lipid bilayer.[1] Human 4.1R is encoded by a single genomic locus over 200 kb in length,[2] and is expressed as multiple isoforms resulting from complex alternative pre–messenger RNA (pre-mRNA) splicing pathways. Previous studies have shown that inclusion of a 21-amino acid sequence motif at the N-terminus of the 10-kd spectrin/actin-binding (SAB) domain is required to promote cytoskeletal junctional complex stability.[3,4] Genomic cloning of both the mouse and human genes confirmed that the 63-nucleotide (nt) motif is encoded by an individual exon, exon 16.[2,5] The splicing of this exon is therefore regulated in a cassette fashion. Exon 16 is omitted from much of the 4.1R mRNA of pre-erythroid cells but is included in most of the mRNA produced in late erythroid cells (Figure 1).[6,7] Splicing of exon 16 is thus highly regulated in a differentiation stage-specific manner, and this regulated event is critical for production of 4.1R isoforms that sustain the function of 4.1R in circulating RBCs.

Alternative splicing of pre-mRNA is a fundamental mechanism for regulating eukaryotic gene expression.[8] In many cases, alternative RNA splicing contributes to developmentally regulated and cell type-specific patterns of gene expression. Although a great deal of information is available concerning the general constitutive splicing reactions of simple splicing units, much less is known about the molecular mechanisms and cellular factors involved in regulated alternative splicing.[8,9] Studies of vertebrate genes aimed at understanding the regulation of alternative splicing have led to identification of a number of pre-mRNA features that influence alternative splice selection. These include the sequence of the 5′ and 3′ splice sites, the branch point sequence, the upstream polypyrimidine (poly[Y]) tract, and specific RNA sequences (enhancers and silencers) located in exons or introns.[10] Elements that stimulate splicing to an adjacent splice site are known collectively as splicing enhancers, whereas elements that repress the use of splice site are known as splicing silencers.

In addition to *cis*-acting elements present in pre-mRNA, the general pre-mRNA splicing mechanism is governed by small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP splicing factors.[9,10] Multiple factors are believed to be involved in the regulation of alternative splicing through a complex network of interactions between the factors and pre-mRNA elements. Factors exert both positive and negative effects on the exon recognition. Moreover, several studies have shown that tissue-specific distribution of "constitutive" factors (SR proteins) varies to a certain degree and that overexpression of SR proteins in vivo alters the
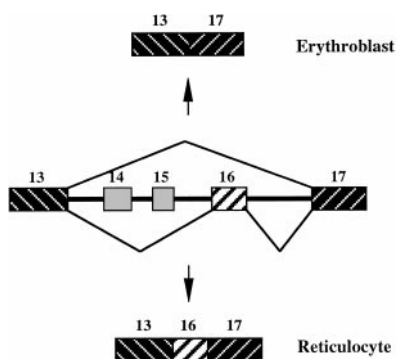
Figure 1. Alternative splicing of 4.1R exon 16 during erythroid differentiation. Exons 14 and 15 are skipped in most tissues. Exon 16 is skipped in lymphoid cells and early erythroblasts and retained in terminally differentiated reticulocytes.

splicing pattern.[11] Hence, differences in the activities or amounts of general non-snRNP factors, as well as the presence of "specific" factors, could modulate alternative splicing pathways.

We have shown that exon 16 is preferentially included in mature 4.1R mRNA in induced mouse erythroleukemia (MEL) cells, and that the protein isoform bearing exon 16-encoding peptide is preferentially and rapidly recruited in the membrane.[7] We present in this report the identification of sequences in the *4.1R* gene that regulate the extent of exon 16 inclusion during differentiation stage-specific splicing events occurring in maturing erythroid cells. We developed a stable transfection system that reproduces endogenous 4.1R splicing patterns, using inducible MEL cells and transfected minigene constructs bearing exon 16. Furthermore, we identified multiple splicing enhancers and silencers that lie within exon 16 and close-vicinity intronic sequences. Small changes in the wild-type (WT) sequences altered exon 16 splicing patterns, but only certain of these elements seem to modulate exon recognition in a stage-specific manner. Our results suggest that exon sequences and abutting intronic sequences contain both positive and negative elements that mediate control of alternative splicing of exon 16. These elements must serve coordinately as primary targets to promote exon skipping in predifferentiated MEL cells and to allow exon inclusion as the cells differentiate.

Materials and methods

Genomic cloning and DNA sequencing

A λCharon human genomic library was screened using a protein 4.1R complementary DNA (cDNA) probe pLym.[12] Screening, hybridization, and probe-labeling procedures have been previously described.[2] Positive clones containing the SAB domain-encoding exons were isolated and mapped using a nonradioactive Southern blot analysis (Boehringer Mannheim, Indianapolis, IN). Selected exon-containing fragments were subcloned into pBluescript II vector (Stratagene Cloning Systems, La Jolla, CA). Exon-intron boundaries were determined by double-strand sequencing of recombinant plasmids.

Minigene constructs

A splicing cassette (p(13,17)/cytomegalovirus [CMV]) was designed to contain the 2 adjacent constitutive exons 13 and 17 with their downstream and upstream flanking intron sequences, respectively. A polycloning site was included at the intron junction. This chimeric DNA fragment was generated by a 2-step polymerase chain reaction (PCR) amplification,[13] and subsequently inserted at the *Not*I/*Xba*I site of pRc/CMV expression vector (Invitrogen, San Diego, CA). Exon 16 was inserted at the *Bst*EII/*Nhe*I site

of the cassette as a 0.7-kb PCR-generated genomic fragment. The resulting construct will be referred to as wild-type or WT minigene.

Mutant constructs were also prepared as *Bst*EII/*Nhe*I fragments and subcloned in p(13,17)/CMV. Large sequence substitutions were generated by the same 2-step PCR method.[13] The first PCR step was performed using chimeric primers overlapping exon 16 and exon 7 or exon 16 and globin exon 2, to generate precise insertions of predefined fragments. To construct point-mutated minigenes, primers were designed to contain the sequence changes. Each mutant was then generated by the 2-step PCR method.[13] Alternatively, the mutation was incorporated by a PCR site-directed mutagenesis procedure using complementary mutant primers and the WT construct as template (QuickChange Site-Directed Mutagenesis Kit, Stratagene). The *Bst*EII/*Nhe*I mutated fragments were further subcloned in p(13,17)/CMV cassette.

In all constructs with base substitutions, except construct "AA," the nucleotide changes created or abolished new restriction sites, which allowed rapid miniprep selection of the mutant clones. These mutant constructs were named after the corresponding restriction endonuclease: B for *Bgl*II; L for *Sal*I; UI-Xm for upstream intron, *Xma*I; DI-Pm for downstream intron, *Pml*I; and so forth. The sequence substitutions are indicated in Figure 2 and listed in Table 1. The cassette, as well as the WT and the mutant inserts, were fully sequenced to ascertain the absence of any additional mismatch, resulting from PCR or cloning errors, which would interfere with the splicing process or the RNase protection assays (RPAs).

Cell culture, induction, and transfection

Several MEL cell subclones were used. In most of the transfection experiments described here, we used subclone 745-A. Cells were cultured and induced to erythroid differentiation as previously described,[7] on fibronectin substratum or in suspension in the absence of fibronectin. For stable expression, proliferative cells (2×10^5 cells) were transfected prior to dimethyl sulfoxide (DMSO) induction with 0.5 to 1 μg plasmid DNA using LipofectAmine (Life Technologies, Gaithersburg, MD) or Escort (Sigma, St Louis, MO) transfection reagents in 35-mm tissue culture plates, according to the manufacturer's procedures. Transfected cells were cultured in complete Dulbecco modified Eagle medium containing 800 μg G-418/mL. Selection of stably transfected clones was performed by incubation of individual cell colonies in 96-well microtiter plates.

Semiquantitative reverse transcription–PCR

Total RNA was obtained by ultracentrifugation of guanidinium isothiocyanate cell lysate on a CsCl cushion, as previously described,[7] or using

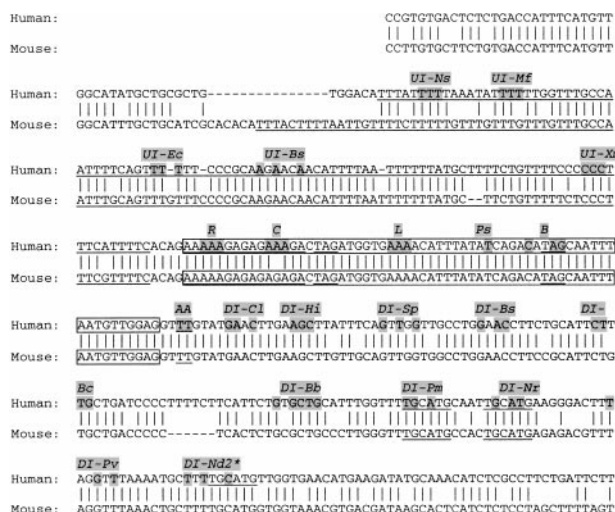**Figure 2. Genomic sequences of exon 16 and surrounding introns in human and mouse.** Exon sequences are boxed. The long poly(Y)-rich sequence and the particular conserved motifs are underlined (see text). Minigene construct designation and the corresponding targeted nucleotides are shaded. *In DI-Nd2 construct, the mutated UGCAUG was inserted in duplicate as ascertained by sequencing analysis.

Table 1. Minigene construct mutations

Construct	WT sequence	Mutated sequence
UI-Ns	ATTTATTTTTAAATA	ATTTAT<u>GCA</u>TAAATA
UI-Mf	AAATATTTTTTGGTT	AAATAT<u>CAA</u>TTGGTT
UI-Ec	TTCAGTTTTTTCCCG	TTCAGT<u>GAA</u>TTCCCG
UI-Bs	CCGCAAGAACAACAT	CCGCA<u>C</u>G<u>TAC</u>GACAT
UI-Xm	TTTCCCCCCTTTCAT	TTTCCC<u>GGG</u>TTTCAT
R	ACAGAAAAAGAGAGA	ACAGAA<u>TTC</u>GAGAGA
C	AGAGAGAAAGACTAG	AGAGAG<u>CTC</u>GACTAG
L	ATGGTGAAAACATTT	ATGGTGT<u>CG</u>ACATTT
Ps	ATTTATATCAGACAT	ATTTATA<u>A</u>CAGACAT
B	TCAGACATAGCAATT	TCAGAGAT<u>CT</u>CAATT
AA	GGAGGTTTGTATGAA	GGAGGT<u>AAG</u>TATGAA
DI-Cl	TGTATGAACTTGAAG	TGTATC<u>GA</u>TTTGAAG
DI-Hi	ACTTGAAGCTTATTT	ACTTGA<u>TAA</u>TTATTT
DI-Sp	TTTCAGTTGGTTGCC	TTTCA<u>C</u>TAG<u>T</u>TTGCC
DI-Bs	GCCTGGAACCTTCTG	GCCTGTA<u>CAC</u>TTCTG
DI-Bc	GCATTCTTTGCTGAT	GCATT<u>GAT</u>CACTGAT
DI-Bb	TTCTGTGCTGCATTT	TTCT<u>TTCGAA</u>CATTT
DI-Pm	GGTTTTGCATGCAAT	GGTTT<u>CAC</u>GTGCAAT
DI-Nr	CAATTGCATGAAGGG	CAATT<u>CGC</u>GAAGGG
DI-Pv	GACTTTAGGTTTAAA	GACTT<u>CAGCT</u>GTAAA
DI-Nd2	ATGCTTTTGCATGTT	ATGC<u>AT</u>AT<u>GG</u>ATGTT

Deviations from the WT sequences are underlined.

RNeasy Total RNA Kit (Qiagen, Hilden, Germany). The reverse transcription (RT) step was performed on 1μg total RNA for 20 minutes at 42°C, using random hexamers and 25 U Moloney murine leukemia virus (M-MLV) reverse transcriptase (Perkin Elmer-Cetus, Norwalk, CT). After heat inactivation of the reverse transcriptase, the PCR reagents were added. The resulting master PCR mixture contained in final concentrations: 1 × PCR buffer, 2 mmol/L MgCl$_2$, 0.2 mmol/L of each dNTP, 0.2 μmol/L of each primer, and 1.2 U *Taq* polymerase. Two sets of primers were used to amplify specifically the transcripts deriving from either the endogenous murine or the transfected human genes. The semiquantitative PCR amplification was carried out for the cycle numbers discussed later, and the PCR products quantified as previously described.[14] Quantification measurements were obtained using ImageQuantMac v1.2 software.

RPA

Total RNA (5 μg) from transfected MEL cells was analyzed by RPA as previously detailed,[7] using H(13,16) a human-specific probe. H(13,16) was generated by RT-PCR of reticulocyte RNA, using the forward primer of the cassette (5′-AAGGTGGCGTCCTAGATGCCTCTGCT-3′) and a reverse primer within exon 16 (5′-ACATTAAATTGCTATGTCTG-3′). The resulting PCR fragment was subcloned into the pCR II vector (Invitrogen). It is noteworthy that the forward primer of the cassette (which is also used as a sense primer in the PCR amplification of the transfection products [see above] and as a forward primer to design the RPA probe), contains a 12-nt human-specific sequence at its 3′ end. This sequence is absent from the mouse 4.1 cDNA.[5] This feature allowed us to distinguish (RPA) or select (RT-PCR) the RNA products derived from the transfected minigene from those arising from the endogenous mouse *4.1R* genes.

Results

Potential *cis*-elements within and surrounding exon 16

We isolated a human genomic clone containing exon 16 and adjacent introns.[2] The corresponding genomic fragment in the mouse has also been isolated.[5] Intron sequences were determined (Figure 2). In addition to a highly conserved exon sequence, several characteristics, remarkably conserved in both human and mouse, have been found within and surrounding the exon (underlined

sequences in Figure 2): (1) an exceptionally long poly(Y)-rich tract of 105 and 121 nt in human and mouse, respectively, upstream of the 3′ splice site; (2) an exclusive purine sequence containing GAR repeats at the 5′ third of the exon—similar GAR repeat elements have been described as splicing enhancers[15]; (3) 2 UAG exonic triplets—this motif is known to act as a splicing silencer in other systems through its binding to hnRNPA1[16]; (4) a weak 5′ splice site—although it shows the highly invariable dinucleotide/GT, this splice site diverges from the consensus 5′ splice site at positions +3 and +4 where pyrimidines (TT) are present instead of the usually encountered purines (A[or G] A)[17]; and (5) a TGCATG motif, which has been found in other pre-mRNA alternative splicing,[18-20] present in 3 copies in the downstream intron of both human and mouse sequences. Each of these particular sequences has been targeted to determine their potential role in exon 16 alternative splicing.

Alternative splicing of minigene pre-mRNAs is regulated in induced MEL cells

Exon 16 with about 300 nt flanking intron sequences on each side was inserted in a splicing cassette as detailed in "Materials and methods" (Figure 3). We postulated that the resulting minigene would contain the minimum sequence for a proper and regulated splicing and would therefore mimic the behavior of the endogenous gene. Indeed, evidence suggests that the sequence elements required for splicing of most pre-mRNA species are located within ~300-nt region containing the branch point, the exon itself, and the 5′ and 3′ splice sites on both sides of the exon.[17]

The MEL cells stably transfected with construct WT were induced to erythroid differentiation and mRNA collected after 5 days of induction. As shown in Figure 3, RPA and semiquantitative RT-PCR revealed a switch from a predominant exon skipping to a predominant exon retention after cell induction. The switch remains partial but very similar to the pattern of endogenous pre-mRNA arising from the mouse genes. Reticulocytes isolated from peripheral blood produce almost exclusively an mRNA isoform that includes exon 16. As previously discussed,[7] the presence of some mRNA lacking exon 16 in induced cells is likely due to the heterogeneity of the cell population, within which only a fraction of the cells attain terminal differentiation. These data suggest that the minigene contains sequences that are sufficient for splicing regulation of that particular exon in an erythroid context. In these experiments on WT sequences, as well as in experiments carried out on most mutant constructs (see below), we tested at least 2 different stable clones per construct. We noted differences in the level of total pre-mRNA transcription but not in the qualitative pattern (splicing) expression. Exon inclusion was also measured in different transfection experiments using either a pool of G-418–resistant cells, or individual stably transfected clones. No significant difference in the pattern was observed (not shown). All of these data argued against any artifacts arising from the integration site or transcriptional efficiency of the minigene.[21] We also conducted a comparative analysis of RPA and RT-PCR as methods of tracking the changes in splicing patterns. We found that the changes in exon inclusion were basically of the same magnitude in both cases (Figure 3).

Upstream intron enhancer elements are required for exon recognition and its splicing regulation

The sequence upstream of the 3′ splicing site displays an unusually long stretch of poly(Y). Substitution of this poly(Y)-rich region by
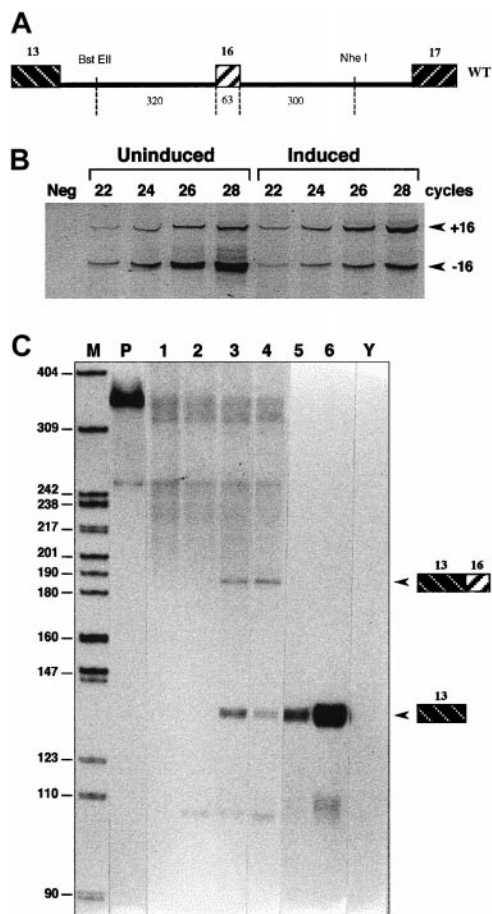
Figure 3. Experimental model for exon 16 erythroid-regulated splicing. (A) Minigene construct. Exon 16 is inserted as an internal exon, together with its flanking intron sequences in the 4.1R cassette p(13,17)/CMV (WT construct). DNA fragment sizes are given in base pairs. (B-C) Differential splicing of transfected exon 16. (B) Semiquantitative RT-PCR analysis on polyacrylamide gel electrophoresis. The +16 and −16 indicate the inclusion or exclusion of exon 16, respectively. Neg indicates negative. (C) RPA using probe H(13,16). Mock uninduced (1) and induced (2) MEL cells; uninduced (3) and induced (4) MEL cells transfected with WT construct; uninduced (5) and induced (6) MEL cells transfected with the p(13,17)/CMV cassette without the 0.7-kb exon 16 fragment. Y, yeast RNA; M, size markers; P, undigested probe. The lower diffused band corresponds to the endogenous 4.1R mRNA.

heterologous intron sequences, which are efficient for constitutive splicing of β-globin exon 2 (construct GI, Figure 4), led to complete skipping of the exon in both uninduced and induced cells. We asked whether this constitutive skipping results only from intronic sequence changes or a cooperative interaction with the weak 5′ splice site by testing 2 mutant constructs where either one or both elements were altered (Figure 4). Substitution of the WT pyrimidines at positions +3 and +4 by purines (AA construct) strengthened the 5′ splice site and led to constitutive exon inclusion in both uninduced and induced cells (Figure 4). Surprisingly, a minigene incorporating both changes (eg, globin intron and AA changes; GI/AA construct in Figure 4), behaved like the GI construct; exon recognition during induction was not rescued by the strengthening of the 5′ splice site. In keeping with our preliminary observations,[22] these data suggest that exon recognition requires the upstream intronic sequences, both in WT context as well as in the presence of a strong 5′ splice site.

To locate more precisely the upstream intronic region that exerts a stimulatory and dominant effect on exon inclusion, we designed 4 constructs by substituting 4.1R intronic sequences for their homo-

logues on globin intron 1. Three regions have been targeted (Figure 4): (1) the proximal region, "a," from −25 to the 3′ splice site; (2) the branch point region, "b," from −55 to −25 (this sequence contains either the WT or the heterologous branch point); and (3) the distal region, "c," from the BstEII site to −55 upstream of the 3′ splice site. Sequence changes at the branch point region, construct GI(b), dramatically inhibited exon inclusion. The exon inclusion is completely abolished when both the branch point region "b" and the distal region "c" were replaced by heterologous sequences. However, the only change of the proximal sequence, construct GI(a), led to complete exon skipping in predifferentiated cells and throughout cell induction.

Thus, the upstream intronic sequences seem to act positively on exon 16 splicing. To know whether this stimulatory effect abolished in constructs GI is due to suppression of a 4.1R splicing enhancer(s) or insertion of globin silencer, we tested 5 additional constructs where the poly(Y)-rich sequence was altered by single nucleotide substitutions (constructs UI, Figure 4). As was true with the constructs containing large fragment substitutions, exon inclusion was virtually abolished in construct UI-Xm (region a) in both predifferentiated and differentiated cells, whereas it was dramatically reduced in constructs UI-Bs (region b), and UI-Ec, UI-Mf, and UI-Ns (region c). In the 2 latter constructs, the reduction occurred most specifically in induced cells (Figure 4). These data suggest that upstream poly(Y)-rich intron sequences contain 2 enhancer elements: a distal element within region c, which stimulates exon inclusion after induction, and a proximal element that resides within the first 25 nt

Construct	Exon 16 inclusion (%)	
	Uninduced	Induced
WT	17±5	57±14
GI	~0	~0
AA	~100	~100
GI/AA	~0	~0
GI(a)	~0	~0
GI(b)	5±2	7±5
GI(b,c)	~0	~0
GI(a,c)	~0	~0
UI-Xm	~0	~0
UI-Bs	5±1	13±4
UI-Ec	7±2	10±2
UI-Mf	17±5	22±5
UI-Ns	16±4	26±6

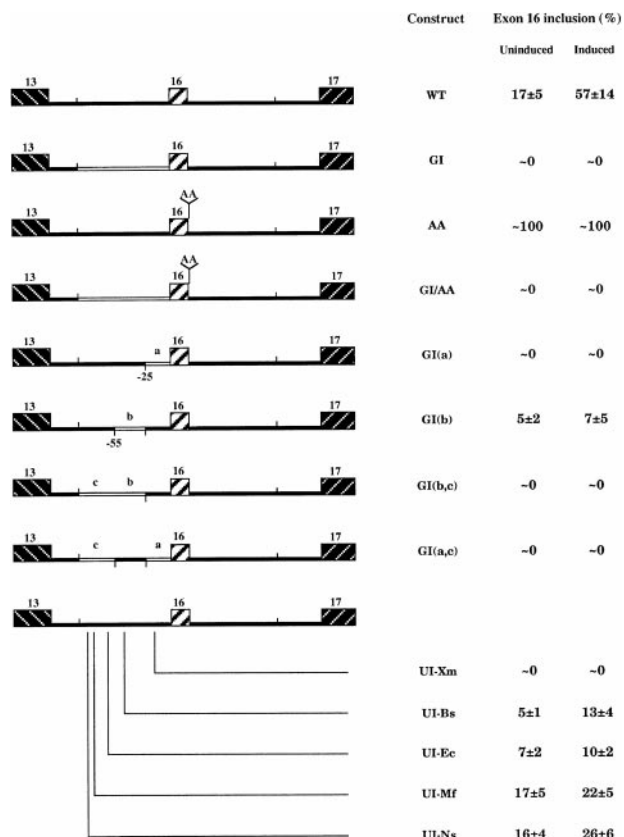**Figure 4. Effect of upstream intron mutations on exon 16 splicing in induced MEL cells.** In construct GI, as well as in GI(a), GI(b), GI(b,c), and GI(a,c), globin intronic fragments (open boxes) were substituted for 4.1R sequences. Data are expressed as percentage of exon 16 inclusion in both uninduced and induced cells. These percentages were averaged from 3 independent RT-PCR experiments ± SD.

upstream of the exon (region a) and that activates exon inclusion in a constitutive manner.

A silencer element within the exon acts in a constitutive manner

We next searched for a possible role for exon sequences. Exon 16 is expected to be an intrinsically poor splicing substrate because of the weak 5′ splice site (see above). To address the contribution of putative exonic regulatory elements that would enhance exon inclusion after cell induction, we first tested 2 different constitutive exon sequences. To do so, exon 16 was precisely substituted by human β-globin exon 2 (construct GE, Figure 5) or 4.1R exon 7 (construct E7). In the E7 minigene, the whole 5′ consensus splice site including the last codon at the 3′ end of the exon was derived from exon 16; however, the donor 5′ splice site in construct GE displayed a consensus (AGG/guuugu) that is potentially weaker than the WT consensus (GAG/guuugu). Surprisingly, the internal exon 7 was excluded in both uninduced and induced cells, whereas globin exon 2 was constitutively and fully included (Figure 5). These results suggest that exon sequences are essential for splicing regulation, at least in this particular intronic nucleotide context.

Exon 16 contains a purine-rich sequence in its 5′ half (Figure 2). Such purine-rich sequences within exons have been described as splicing enhancers in both constitutive and alternative exons (see "Discussion"). Also, the exon sequence contains several AAA trinucleotide stretches, which are found at a relative high frequency in alternatively spliced exons.[23] In construct 15GE, the whole purine-rich tract was substituted by the 5′ 15 nt of β-globin exon 2. In construct 42GE, this purine stretch was kept intact but the 3′ 42 nt of the exon were replaced by the 3′ 42 nt of globin exon 2. The chimeric exon sizes and reading frames were unchanged. In 15GE, substitution dramatically enhanced the exon selection in uninduced cells. This effect was even more pronounced in the 42GE mutant; exon skipping was completely de-repressed both in uninduced cells and throughout differentiation (Figure 5). These data suggest the presence of either a negative element in WT exon 16 or an
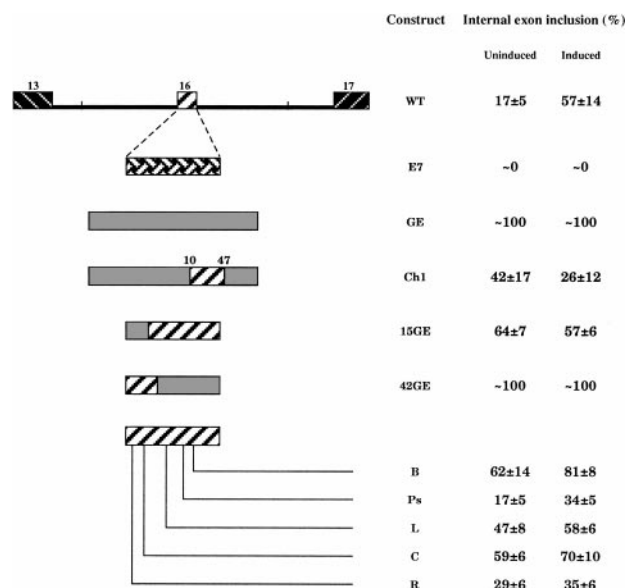
interference effect by an enhancer element within the heterologous globin sequences.

To distinguish between these 2 possibilities, 5 mutant constructs were made (Figure 5); in R and C constructs, A triplets were mutated to disrupt the putative purine-rich enhancer element at the 5′ end of the exon. Similarly, the downstream part of the exon was also targeted by triplet substitutions; in constructs L and B, an A triplet and a UAG element were targeted, respectively. Construct Ps derived from a single base substitution between these upstream and downstream parts of the exon. Again, care was taken not to generate sequences known to affect splicing, such as cryptic splice sites and stop codons. Exon substitution in construct R seemed to have a slight stimulatory effect in uninduced cells and a slight inhibition effect on the regulated exon inclusion after cell induction. In construct Ps, only a little inhibitory effect, if any, was observed after cell induction. In contrast, in the C, L, and B pre-mRNA substrates, the internal exon was preferentially incorporated into the mature mRNAs in predifferentiated cells. Interestingly, a regulatory effect of induction is partly preserved in the C, L, and B mRNAs; indeed, the isoform bearing exon 16 accumulates at a higher amount in differentiating cells than in proliferating cells. On the other hand, it is noteworthy that construct 15GE appeared to combine the effects observed in R and C constructs; exon selection was increased in uninduced cells (C effect) at a level that was not further enhanced following induction (R effect).

Finally, the CLB region of exon 16 was inserted in the β-globin exon 2 (construct Ch1, Figure 5), and the splicing of the chimeric exon was tested in transfected cells. Consistent with the negative effect observed in the WT minigene, this sequence also exhibited a silencing effect in a heterologous sequence context. Inclusion of the internal exon was dramatically reduced in construct Ch1, as compared with construct GE (Figure 5).

Altogether, these results suggest the presence of an inhibitor element encompassing the region mutated in constructs C, L, and B and containing the 2 UAG elements. This latter sequence inhibits the exon selection in uninduced cells, acting therefore as a constitutive exonic splicing silencer.

Downstream intron sequences enhance exon inclusion in induced cells

To determine the contribution of the downstream intron sequences in the regulated splicing event, we tested a series of single mutants (DI constructs, Figure 6), 4 of which were designed to alter the UGCAUG elements. All these mutations led to a significant decrease of exon inclusion in induced cells. In construct DI-Pm/Nr, 2 of the 3 UGCAUG elements were targeted by mean of cumulating DI-Pm and DI-Nr mutations. This led to a more pronounced inhibiting effect. In agreement with previous data, these results suggest a dose-enhancing effects of the UGCAUG elements.[19]

Moreover, DI-Bs and most remarkably DI-Bc and DI-Bb showed a strong splicing inhibition of the exon at both predifferentiated and differentiated stages. Interestingly, Bc mutation alters a UCUU sequence. Such a pyrimidine sequence functions in most cases through a competitive binding to PTB over U2AF upstream of the 3′ splice site and prevents selection of the downstream exon.[24,25]

In any case, our data suggest the presence of enhancer elements spanning a region from the 5′ consensus splice site to the last UGCAUG repeat element downstream. Among these elements, the sequence encompassed by DI-Bc and DI-Bb seems to have a strong impact at the constitutive exon recognition, whereas the regulated

Figure 5. Effect of exon mutations on exon 16 splicing during MEL cell induction. In constructs E7 and GE, exon 16 was substituted by the constitutive 4.1R exon 7 or β-globin exon 2, respectively. In construct Ch1, the putative exon splicing silencer from exon 16 was inserted in the globin exon. In all these constructs, the flanking intronic sequences are identical to the WT intronic context. See also the legend to Figure 4.

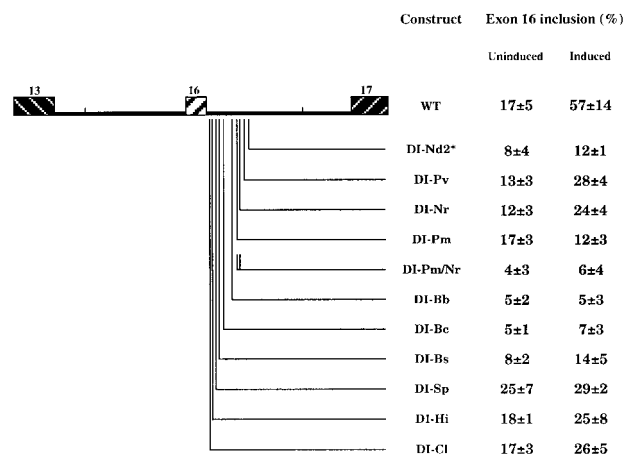| Construct	Exon 16 inclusion (%)	
	Uninduced	Induced
WT	17±5	57±14
DI-Nd2*	8±4	12±1
DI-Pv	13±3	28±4
DI-Nr	12±3	24±4
DI-Pm	17±3	12±3
DI-Pm/Nr	4±3	6±4
DI-Bb	5±2	5±3
DI-Bc	5±1	7±3
DI-Bs	8±2	14±5
DI-Sp	25±7	29±2
DI-Hi	18±1	25±8
DI-Cl	17±3	26±5

Figure 6. Effect of downstream intron mutations on exon 16 splicing in induced MEL cells. In constructs DI-Nd2, DI-Pv, and DI-Nr, the 3 UGCAUG motifs were targeted. *Construct DI-Nd2 contains the mutated UGCAUG in duplicate. Note that the effect on exon inclusion occurs more specifically in induced cells for most of the mutants. For constructs DI-Bb and DI-Bc, exon selection is clearly inhibited in both uninduced and induced cells. See also the legend to Figure 4.

splicing appears to be under the control of sequences present on both sides of this region.

Discussion

Several studies have shown that the expression of exon 16 of protein 4.1R during late erythroid differentiation is required for its interaction with spectrin and actin.[3,4,6,7] However, little is known about the developmental regulated alternative splicing mechanism that governs the selection of exon 16 during erythroid differentiation. In this study, we established an exon 16–minigene system in MEL cells that mimics endogenous exon 16 splicing patterns during DMSO-induced erythroid differentiation. Moreover, we used mutagenesis analysis and demonstrated that a complex interplay of positive and negative regulatory signals present in the RNA controls the developmental stage-specific splicing of exon 16.

Protein 4.1R exon 16 splicing has been examined previously in *Xenopus* oocytes and the HeLa system.[26] Transcripts from an exon 16 minigene resulted in exon 16 inclusion when microinjected into *Xenopus* oocytes and exclusion in HeLa nuclear extracts. This difference suggests that specific alternative splicing factors are present differently in oocytes and HeLa. These 2 splicing systems may or may not reflect the same mechanism responsible for alternative splicing of exon 16 during erythroid differentiation. MEL cells have been shown to successfully reproduce the erythroid pattern of regulated pre-mRNA splicing of β-spectrin[27] and 4.1R[7,22] (and present data) albeit with a partial shift from a predifferentiated to a differentiated splicing pattern after cell induction due to a heterogeneous cell population.[28] Our results showed that when ~60% of the induced MEL population is benzidine positive, 4.1R mRNA includes exon 16 to a high degree in induced cells (57% versus 17% before induction). This makes MEL a useful model system for the study of regulated alternative splicing of exon 16. Indeed, the minigene bearing exon 16 and its flanking upstream and downstream intronic regions mimicked endogenous exon 16 splicing patterns in transfected MEL cells, suggesting that elements needed for exon 16 selection are present within these sequences. In addition to the identification of the *cis*-elements, the MEL experimental model led us to distinguish between 2 classes of regulatory

sequences, that is, elements acting constitutively and others acting more specifically in induced cells.

The analysis of *cis*-elements required for exon 16 splicing revealed several important features within the exon and its surrounding intron sequences. In exon 16, the 3′ splice site is in a sequence context similar to the consensus.[17] However, the 5′ splice site is a weak splice site; the critical positions +3 and +4 are pyrimidine residues that deviate from the consensus purines.[17] Strengthening of this 5′ splice site to consensus sequence as in AA construct led to the selection of the upstream exon 16 in a constitutive manner (Figure 4 and Baklouti and colleagues[22]). Both the 5′ and 3′ splice sites are required for spliceosome complex formation in which the 5′ splice site is recognized by U1 snRNA base pairing and associated with SR splicing factors, and the 3′ splice site is recognized by U2 auxiliary factor (U2AF). Mutation of a weak 5′ splice site toward U1 snRNA complementarity has been shown to enhance exon selection.[29] The ability of SR proteins to direct the use of alternative 5′ splice site by promoting the binding of U1 snRNP to 5′ splice site has also been documented.[30,31] It seems that the weak 5′ splice site contributes to the omission of exon 16 prior to induction. In the induced cell, conditions must be created that permit the 5′ splice site to be used efficiently.

The *cis*-acting enhancing elements that direct the splicing machinery to proximal, usually suboptimal, splice sites fall into 2 general groups, intronic and exonic. A number of intronic splicing enhancers from both vertebrates and invertebrates that influence alternative splicing in a cell-specific manner have been described.[18-20,32-36] In the majority of these, 2 recurring types of sequence elements mediate alternative splice usage: poly(Y) tracts[32,36] and short sequence repeats.[18-20] The upstream intron sequences of exon 16 contain an unusually long poly(Y) stretch (Figure 2) in both mouse and human 4.1R. Similar features have been found in the α-and β-tropomyosin (*TM*) genes (Gallego and coworkers[32] and references therein). We examined the effect of the long poly(Y) sequences on exon 16 splicing by replacing the upstream intron with globin intron 1 (GI construct). The replacement prevented exon 16 inclusion in both uninduced and induced MEL cells. Furthermore, splicing of exon 16 could not be rescued by concomitant strengthening of the WT 5′ splice site (Figure 4 GI/AA) when the upstream intron was replaced with globin intron 1. These data suggest that the presence of upstream positive intronic elements are required for exon 16 recognition regardless of induction status.

Another recurring downstream intronic enhancing element is the hexamer TGCATG, which are found near exon 16. The sequence has been demonstrated to function as an enhancer in the alternative splicing of exon IIIB of the *fibronectin* gene,[18,19] and the neuron-specific 18-base N1 exon of the mouse *c-src* gene.[20] These elements are well conserved near exon 16 between mouse and human (Figure 2). There are 3 UGCAUG elements; the alteration of these inhibited exon inclusion in induced cells. These UGCAUG motifs seem also to function in tissue- and developmental-specific manners.[18,20]

Exon size and sequence play critical roles in the selection of splice sites.[37,38] The different splicing patterns observed with GE and E7 constructs (Figure 5), that is, with 2 different constitutive exons positioned within the same flanking intron sequences, further support this view.

Most exonic enhancers are composed of purine-rich sequences, which are thought to mediate their activity mainly through the binding of constitutive SR proteins.[11] Purine-rich sequences at the

5′ end of exon 16 match most of the consensus octamer RGAA-GAAC proposed as an ASF/SF2 enhancer core element.[39] Strikingly, disruption of this element in constructs C and 15GE, led to a dramatic increase in exon recognition rather than exon skipping (Figure 5). In addition to this 5′ purine-rich sequence, exon 16 contains 2 copies of UAG triplet, which is known as a core element for hnRNPA1-binding exon splicing silencer.[16,40] The 42 base pairs at the 3′ of exon 16 is conserved across many species, including human, mouse, rat, and *Xenopus* suggesting their importance in regulated splicing. The substitution of this region with corresponding sequences from β-globin exon 2 results in the inclusion of exon 16. It has been demonstrated in a number of test cases that SR proteins and hnRNPA1 antagonistically affect 5′ splice site recognition and that subtle changes in the relative concentrations of the 2 types of factors can be determinative for 5′ splice site recognition.[41] Hou and colleagues[42] have recently reported in in vitro splicing assays that selection of exon 16 correlates with hnRNPA1 in a concentration-dependent manner. Addition of excess hnRNPA1 leads to exon exclusion, and partial depletion of hnRNPA1 results in exon inclusion.

The MEL cell model used in this work suggests that exon 16 skipping is the default pathway, occurring in predifferentiated erythroid cells. It would result from a predominant, and possibly cooperative, effect of the exon splicing silencer and the weak downstream 5′ splice site. The regulatory pathway providing for splicing in differentiating cells could result from a stimulatory effect that acts through interactions between the upstream and the downstream intron enhancers and possibly the exonic purine-rich element. These positive interactions would overcome constitutive inhibition following cell commitment.

Exon selection resulting from the coordinate activation of both flanking introns has been described in the neuron-specific and developmental splicing regulation of the 24-nt exon of the GABA(A) receptor γ2 pre-mRNA.[43] Our study further supports the view that an exon is recognized as a whole unit and that the sum of the strength of the splicing elements is the key determinant of exon definition. In addition to the *cis*-elements, the basis for differentiation stage-specific splicing of exon 16 probably also involves exon 16-specific splicing factors that are uniquely present (or increased) in the late erythroid cells and absent (or reduced) in early stage.[42,44] We hypothesize that such differentiation-specific factors would interact with both stimulatory and inhibitory *cis* elements to promote exon 16 splicing.

Several erythroid differentiation stage–specific alternative splicing phenomena have been reported. The third enzyme of the heme synthesis pathway, porphobilinogen deaminase, is expressed both as a "housekeeping" form present in all cell types, and a second isoform expressed only in erythroid cells.[45] Erythroid-specific splice forms involving the first 3 exons of the second (5-aminolevulinate dehydratase),[46] and the fourth (uroporphyringen III synthase)[47] heme biosynthesis enzymes have also been described. These data suggest that erythroid-specific cell development depends in part on regulated splicing. Many mutations result in splicing defects by destroying splice sites, creating cryptic splice sites, or altering splicing enhancer or silencer elements. Erythroid genes whose expression is affected by such mutations include protein 4.1R.[14] The elucidation of the mechanisms controlling alternative splicing of 4.1R pre-mRNA will enhance our understanding of the role of regulated splicing in RBC membrane biogenesis, and possibly in the expression of other key erythroid protein forms during late erythropoiesis.

Acknowledgments

References

1. Bennett V, Gilligan DM. The spectrin-based membrane skeleton and micron-scale organization of the plasma membrane. Annu Rev Cell Biol. 1993; 9:27-66.

2. Baklouti F, Huang SC, Vulliamy TJ, Delaunay J, Benz EJ Jr. Organization of the human protein 4.1 genomic locus: new insights into the tissue-specific alternative splicing of the pre-mRNA. Genomics. 1997;39:289-302.

3. Horne WC, Huang SC, Becker PS, Tang TK, Benz EJ Jr. Tissue-specific alternative splicing of protein 4.1 inserts an exon necessary for formation of the ternary complex with erythrocyte spectrin and F-actin. Blood. 1993;82:2558-2563.

4. Discher DE, Winardi R, Schischmanoff PO, Parra M, Conboy JG, Mohandas N. Mechanochemistry of protein 4.1's spectrin-actin-binding domain: ternary complex interactions, membrane binding, network integration, structural strengthening. J Cell Biol. 1995;130:897-907.

5. Huang JP, Tang CJ, Kou GH, Marchesi VT, Benz EJ Jr, Tang TK. Genomic structure of the locus encoding protein 4.1. Structural basis for complex combinational patterns of tissue-specific alternative RNA splicing. J Biol Chem. 1993;268:3758-3766.

6. Chasis JA, Coulombel L, Conboy J, et al. Differentiation-associated switches in protein 4.1 expression. Synthesis of multiple structural isoforms during normal human erythropoiesis. J Clin Invest. 1993;91:329-338.

7. Baklouti F, Huang SC, Tang TK, Delaunay J, Marchesi VT, Benz EJ Jr. Asynchronous regulation of splicing events within protein 4.1 pre-

mRNA during erythroid differentiation. Blood. 1996;87:3934-3941.

8. Lopez AJ. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. Annu Rev Genet. 1998;32:279-305.

9. Smith CW, Valcarcel J. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci. 2000;25:381-388.

10. Hastings ML, Krainer AR. Pre-mRNA splicing in the new millennium. Curr Opin Cell Biol. 2001;13: 302-309.

11. Graveley BR. Sorting out the complexity of SR protein functions. RNA. 2000;6:1197-1211.

12. Tang TK, Qin Z, Leto T, Marchesi VT, Benz EJ Jr. Heterogeneity of mRNA and protein products arising from the protein 4.1 gene in erythroid and nonerythroid tissues. J Cell Biol. 1990;110:617-624.

13. Maillet P, Delaunay J, Baklouti F. Chimeric probe-mediated ribonuclease protection assay for molecular diagnosis of mRNA deficiencies. Hum Mutat. 1996;7:61-64.

14. Morinière M, Ribeiro L, Dalla Venezia N, et al. Elliptocytosis in patients with C-terminal domain mutations of protein 4.1 correlates with encoded messenger RNA levels rather than with alterations in primary protein structure. Blood. 2000; 95:1834-1841.

15. Tanaka K, Watakabe A, Shimura Y. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol Cell Biol. 1994;14:1347-1354.

16. Del Gatto-Konczak F, Olive M, Gesnel MC,

Breathnach R. hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. Mol Cell Biol. 1999;19:251-260.

17. Reed R. Initial splice-site recognition and pairing during pre-mRNA splicing. Curr Opin Genet Dev. 1996;6:215-220.

18. Huh GS, Hynes RO. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. Genes Dev. 1994;8:1561-1574.

19. Lim LP, Sharp PA. Alternative splicing of the fibronectin EIIIB exon depends on specific TG-CATG repeats. Mol Cell Biol. 1998;18:3900-3906.

20. Modafferi EF, Black DL. A complex intronic splicing enhancer from the c-src pre-mRNA activates inclusion of a heterologous exon. Mol Cell Biol. 1997;17:6537-6545.

21. Cramer P, Caceres JF, Cazalla D, et al. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol Cell. 1999;4: 251-258.

22. Baklouti F, Zhou J, Delaunay J, Huang SC, Benz EJ Jr. Characterization of cis-elements modulating alternative splicing of exons encoding the spectrin/actin binding domain in protein 4.1 pre-mRNA [abstract]. Blood. 1994;84(suppl 1):361a.

23. Stamm S, Zhang MQ, Marr TG, Helfman DM. A sequence compilation and comparison of exons that are alternatively spliced in neurons. Nucleic Acids Res. 1994;22:1515-1526.

24. Chan RC, Black DL. Conserved intron elements repress splicing of a neuron-specific c-src exon in

vitro [published erratum appears in Mol Cell Biol 1997 May;17(5):2970]. Mol Cell Biol. 1995;15: 6377-6385.

25. Gooding C, Roberts GC, Smith CW. Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon. RNA. 1998;4:85-100.

26. Gee SL, Aoyagi K, Lersch R, Hou V, Wu M, Conboy JG. Alternative splicing of protein 4.1R exon 16: ordered excision of flanking introns ensures proper splice site choice. Blood. 2000;95:692-699.

27. Chu ZL, Wickrema A, Krantz SB, Winkelmann JC. Erythroid-specific processing of human beta spectrin I pre-mRNA. Blood. 1994;84:1992-1999.

28. Orkin SH, Harosi FI, Leder P. Differentiation in erythroleukemic cells and their somatic hybrids. Proc Natl Acad Sci U S A. 1975;72:98-102.

29. Dirksen WP, Hampson RK, Sun Q, Rottman FM. A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. J Biol Chem. 1994;269:6431-6436.

30. Eperon IC, Ireland DC, Smith RA, Mayeda A, Krainer AR. Pathways for selection of 5′ splice sites by U1 snRNPs and SF2/ASF. EMBO J. 1993;12:3607-3617.

31. Zahler AM, Neugebauer KM, Lane WS, Roth MB. Distinct functions of SR proteins in alternative pre-mRNA splicing. Science. 1993;260:219-222.

32. Gallego ME, Gattoni R, Stevenin J, Marie J, Expert-Bezancon A. The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-tropomyosin alternative exon 6A. EMBO J. 1997; 16:1772-1784.

33. Carstens RP, McKeehan WL, Garcia-Blanco MA. An intronic sequence element mediates both activation and repression of rat fibroblast growth factor receptor 2 pre-mRNA splicing. Mol Cell Biol. 1998;18:2205-2217.

34. Cooper TA. Muscle-specific splicing of a heterologous exon mediated by a single muscle-specific splicing enhancer from the cardiac troponin T gene. Mol Cell Biol. 1998;18:4519-4525.

35. Kosaki A, Nelson J, Webster NJ. Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. J Biol Chem. 1998;273:10331-10337.

36. Markovtsov V, Nikolic JM, Goldman JA, Turck CW, Chou MY, Black DL. Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. Mol Cell Biol. 2000;20:7463-7479.

37. Reed R, Maniatis T. A role for exon sequences and splice-site proximity in splice-site selection. Cell. 1986;46:681-690.

38. Watakabe A, Tanaka K, Shimura Y. The role of exon sequences in splice site selection. Genes Dev. 1993;7:407-418.

39. Tacke R, Manley JL. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. EMBO J. 1995;14:3540-3551.

40. Si Z, Amendt BA, Stoltzfus CM. Splicing efficiency of human immunodeficiency virus type 1 tat RNA is determined by both a suboptimal 3′ splice site and a 10 nucleotide exon splicing silencer element located within tat exon 2. Nucleic Acids Res. 1997;25:861-867.

41. Caceres JF, Stamm S, Helfman DM, Krainer AR.

Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science. 1994;265:1706-1709.

42. Hou VC, Gee SL, Wu M, Turck CW, Mayeda A, Conboy JG. Developmentally-regulated expression of hnRNPA1 can regulate alternative splicing of protein 4.1R exon 16 through an interaction with an exonic splicing silencer [abstract]. Blood. 2000;96:441a.

43. Ashiya M, Grabowski PJ. A neuron-specific splicing switch mediated by an array of pre-mRNA repressor sites: evidence of a regulatory role for the polypyrimidine tract binding protein and a brain-specific PTB counterpart. RNA. 1997;3: 996-1015.

44. Huang SC, Baklouti F, Cheng H, Benz EJ Jr. Differential expression of the SR protein splicing factors and alternative splicing of protein 4.1R exon 16 during erythroid maturation [abstract]. Blood. 2000;96:592a.

45. Chretien S, Dubart A, Beaupain D, et al. Alternative transcription and splicing of the human porphobilinogen deaminase gene result either in tissue-specific or in housekeeping expression. Proc Natl Acad Sci U S A. 1988;85:6-10.

46. Kaya AH, Plewinska M, Wong DM, Desnick RJ, Wetmur JG. Human delta-aminolevulinate dehydratase (ALAD) gene: structure and alternative splicing of the erythroid and housekeeping mRNAs. Genomics. 1994;19:242-248.

47. Aizencang GI, Bishop DF, Forrest D, Astrin KH, Desnick RJ. Uroporphyrinogen III synthase. An alternative promoter controls erythroid-specific expression in the murine gene. J Biol Chem. 2000;275:2295-2304.