<table>
<thead>
<tr>
<th>Title</th>
<th>Purine-rich exon sequences are not necessarily splicing enhancer sequence in the dystrophin gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Nakamura, Hajime / Ito, Toshiyuki / Takeshima, Yasuhiro / Sakamoto, Hiroshi / Matsuo, Masafumi</td>
</tr>
<tr>
<td>Citation</td>
<td>The Kobe journal of the medical sciences, 47(5):193-202</td>
</tr>
<tr>
<td>Issue date</td>
<td>2001-10</td>
</tr>
<tr>
<td>Resource Type</td>
<td>Departmental Bulletin Paper / 紀要論文</td>
</tr>
<tr>
<td>Resource Version</td>
<td>publisher</td>
</tr>
<tr>
<td>DOI</td>
<td>10.24546/00069787</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://www.lib.kobe-u.ac.jp/handle_kernel/00069787">http://www.lib.kobe-u.ac.jp/handle_kernel/00069787</a></td>
</tr>
</tbody>
</table>
Purine-Rich Exon Sequences Are Not Necessarily Splicing Enhancer Sequence in the Dystrophin Gene

TOSHIYUKI ITO¹, YASUHIRO TAKESHIMA²*, HIROSHI SAKAMOTO³, HAJIME NAKAMURA², and MASAFUMI MATSUO¹

¹Division of Molecular Medicine and ²Department of Pediatrics, Kobe University Graduate School of Medicine and ³Department of Biology, Kobe University Graduate School of Science and Technology, Kobe 657-8501, Japan

Received 23 August 2001/ Accepted 4 October 2001

Key words: dystrophin; splicing; splicing enhancer sequence; purine-rich sequence

Purine-rich sequences within exons are proposed to promote proper splicing as splicing enhancers. In order to test this supposition in the dystrophin gene, which is characterized by the large size of its introns, purine-rich sequences from three exons (exons 43, 46 and 53) were examined for their splicing enhancer activity by using a Drosophila doublesex pre-mRNA. The most powerful activating effect on upstream intron splicing was seen with a sequence from exon 43. A sequence from exon 53 showed relatively low activity, whilst that from exon 46 had little effect. To characterize the splicing enhancer sequences in exons 53 and 46 further, entire exons were divided into 30nt fragments that were examined separately for their splicing enhancer activity. In exon 53, two fragments located at the 5' and 3' ends, respectively, had strong splicing enhancer activity, although they were not the most purine-rich regions of the exon. In contrast, all of the fragments derived from exon 46 had little activity. These results suggest that different exons in the dystrophin transcript are subject to different mechanisms of control by the splicing machinery.

Splicing removes introns from mRNA precursors, allowing exons to be joined to generate the mature mRNA. One of the most remarkable features of splicing is its ability to select correct pairs of splice sites precisely from among a myriad of potential, but inappropriate splice site. Sequences that are essential for intron removal, the consensus sequences at the 5' and 3' splice sites and the branch site, are present at the exon/intron borders. Splice sites are recognized by small nuclear ribonucleoprotein particles (snRNPs) and non-snRNP protein factors that engage to assemble the active splicing complex, known as a spliceosome.

Large numbers of pseudoexons flanked by pseudosplice sites with good matches to the consensus sequences can be easily identified in intron sequences (8). Real exons are recognized and spliced cotranscriptionally (1,27) and the pseudo sites are ignored. Splice site consensus sequences alone seem to be insufficient to enable the recognition of real splice sites and other factors, such as additional signals, must contribute to their selection from among the many potential splice sites within long transcripts. These additional recognition elements could act either positively or negatively. Positive elements or splicing enhancers were first recognized as purine-rich sequences that promote splicing when present within some exons (26). Among such sequences, the trinucleotide AAG has been proposed as a potential splicing enhancer consensus sequence (18,25).

One of the best-characterized exonic splicing enhancers is in the fourth exon of the Drosophila doublesex (dsx) gene. This purine-rich enhancer promotes female-specific
splicing of exon 4. Interactions between the splicing enhancer sequence and nuclear proteins caused recruitment of essential SR proteins in HeLa cell nuclear extracts into a complex that promotes $dsx$ enhancer activity (10). The prevalence of exonic splicing enhancers is unknown and it is not yet clear whether they play an important role in constitutive splicing.

The human dystrophin gene, which is defective in patients with Duchenne or Becker muscular dystrophy (DMD/BMD), spans approximately 3,000 kb of the X-chromosome and encodes a 14 kb transcript consisting of 79 exons. Consequently, more than 99% of the gene sequence is composed of introns. Several introns exceed 100 kb and one of the longest, intron 44, is over 180 kb (4). In addition, several exons are known to be alternatively spliced in a tissue or development-specific manner (6). The extraordinary nature of the dystrophin gene with respect to size and the number of exons points to several potential problems in the production and processing of dystrophin transcripts (21,22). Among the 79 exons, we have shown that two (exons 19 and 27) contain a purine-rich exonic splicing enhancer sequence (18,24). The 31 mer purine-rich sequence from exon 19 exhibits a high capacity to promote upstream intron splicing in the mini-dystrophin gene product (24). These findings strongly suggest every constitutive exon of the dystrophin gene could have a splicing enhancer that ensures specific exon recognition.

Here we analyzed splicing enhancer activity of purine-rich sequences from three exons in the central region of the dystrophin gene by using the *Drosophila* $dsx$ minigene as a template and found that purine-rich sequences do not necessarily exhibit splicing enhancer activity.

**MATERIALS AND METHODS**

**Selection of candidate splicing enhancer sequences.** Intron 44 in the central region of the dystrophin gene is over 180 kb and it is highly plausible that exons adjacent to intron 44 need a purine-rich exon splicing enhancer sequence to ensure correct splicing. Therefore, we searched for purine-rich sequence in adjacent exons. Since length could be one of the factors that determine splicing enhancer activity, the sequences we selected were all 30nt purine-rich segments similar in length to the 31nt sequence that showed the highest splicing enhancer activity in exon 19 (24). The trinucleotide AAG was included as frequently as possible as it is one of the proposed consensus splicing enhancer sequences (25). Three exons (exons 43, 46 and 53) were found to have 30nt long purine-rich sequences (Fig.1). The sequence in exon 43 (5'-AGCAAGAAGACAGCAGCATTGCAAAG-3' (SES 43)) was selected as a candidate. In addition, two sequences (SES 46; 5'-GGAAAAGAGCAGCAACTAAAAGAAAACG-3' and SES 53; 5'-GGAAGCTAAGGAAGAAGCTGAGCAGG-3') were selected from exons 46 and 53, respectively.

**Analysis of splicing enhancer activity.** Splicing enhancer activity was examined *in vitro* by analyzing the splicing of the heterogenous intron of the *Drosophila melanogaster doublesex (dsx)* gene pre-mRNA in HeLa cell nuclear extracts as described before (18,25). In this *in vitro* system, splicing of chimeric $dsx$ pre-mRNA is dependent on the downstream sequence. Briefly, template plasmids coding for chimeric $dsx$ pre-mRNA were constructed as shown in Fig.2. After confirming its sequence, $^{32}$P-radiolabeled pre-mRNA was prepared by *in vitro* transcription and subjected to *in vitro* splicing with HeLa cell nuclear extracts (Computer Cell Culture Center, Mons, Belgium). Reaction products were phenol extracted twice, precipitated and dissolved. The reaction products were electrophoresed on a 6% sequencing gel and then quantified (BAS2000 Bio Image Analyzer; Fuji Photo Film Co., Tokyo, Japan) as described previously (24).

Sequences that were examined for splicing enhancer activity were inserted between the BamHI and XhoI restriction enzyme recognition sites (Fig.2). These test sequences were prepared by annealing two strands of synthesized single-strand DNA that created BamHI and
FIG. 1. Schematic representation of analyzed sequences. The sequences of exons 43 (a), 46 (b) and 53 (c) are shown. Bracket indicates sequences that were examined for splicing enhancer activity. The name of each sequence is indicated below or above the bracket. Numbers in parentheses indicate splicing enhancer activity as a percentage of the control. Brackets under the sequence indicate the candidate splicing enhancer sequence in the respective exon sequence.

_XhoI_ overhanging ends. Pre-mRNA containing the 31 mer splicing enhancer sequence from exon 19 of the dystrophin gene (5'-AACTGCAAGATGCCAGCAGATCAGCTCAGGC-3') (24) was used as positive control, and the _BamHI_-digested plasmid without any test sequence was used as a negative control. The activity was measured in three experiments and expressed as percentage of that of the control activity.

_Detection of RNA binding nuclear protein._ In order to identify the protein that binds to SES 43, we synthesized 5'-FAM label 2'-O-methyl RNA that is homologous to SES 43 and used it as a probe. The probe (20 pmol) was transferred to 30 µl of reaction mixture containing 20 mM HEPES (pH 7.9 at 4°C), 200 mM KCl, 1 mM DTT, 1 mM EDTA (pH 8), 300 µg/ml BSA, 15% glycerol, 400 µg/ml E.coli tRNA, and 10 µl of HeLa cell nuclear extract. The solution was incubated at 30°C for 20 minutes. The resulting complexes were electrophoresed on 5% nondenaturing polyacrylamide gel (acrylamide/bisacrylamide ratio of 29/1) in 0.5x TBE at 4°C. Quantification of the binding complex was performed by Fluorimager 585 (Molecular Dynamics, Inc., Sunnyvale, CA). The concentration of E.coli tRNA was increased to 10 mg/ml in the binding reaction mixture in order to observe the specificity of binding.

RESULTS

_Splicing enhancer activity of three candidate sequences._ Three purine-rich sequences (SES 43, SES 46 and SES 53) selected from exons 43, 46 and 53 as candidate splicing enhancer sequences (Fig.1) were inserted into the test site of the chimeric _dsx_ minigene (Fig.
FIG. 2. Schematic representation of *Drosophila melanogaster* doublesex (*dsx*) chimeric plasmid. The chimeric plasmid consists of the entire third exon (168 nt), the third intron (114 nt), and 30 nt of the 5′ portion of the fourth exon of the *dsx* gene and the linker sequence derived from the pSP72 vector. The T7 promoter was activated to produce pre-mRNA. The exons (boxes) and intron (line) are shown. The linker sequence and the insertion site for the test sequence (closed box) are indicated. *Bam*HI and *Xho*I indicate restriction enzyme recognition sites, respectively. The annealed test oligonucleotides were cloned into the insertion site.

FIG. 3. Analysis of splicing enhancer activity of SES 43, SES 46 and SES 53. SES 43, 46 and 53 were examined for their splicing enhancer activity by inserting them to the test site of the chimeric gene (Fig. 2). Each *in vitro* splicing product was gel-electrophoresed. Both positive and negative controls are shown. The structures of the RNA products are shown schematically at the left side. The molecular size markers are fragments of 32P-labeled pBR322 obtained by *Msp*I digestion (lane M).
SPLICING ENHANCER SEQUENCE IN THE DYSTROPHIN GENE

2) and its transcripts were subjected to *in vitro* splicing. The largest amount of splicing product corresponding to mature mRNA was obtained when SES43 was inserted (Fig.3) (70.6% of the control). In contrast, only low amounts of splicing products were obtained with SES 46 and SES 53 (Fig.3) (23.6 and 31.4% of the control, respectively) (Fig.1). Thus, although SES 53 would be expected to be the strongest splicing enhancer based on its purine content and number of AAG trinucleotides (Fig.1), it showed low splicing enhancer activity. These results indicated that purine-rich sequences within dystrophin exons do not necessarily have significant splicing enhancer activity.

**Localization of splicing enhancer activity in SES43.** Since SES 43 showed significant splicing enhancer activity, we examined flanking regions to determine whether the enhancer sequence extends beyond the borders of SES 43. Another two sequences that were shifted either 12nt upstream (SES 43-1) or 6nt downstream (SES 43-2) in exon 43 relative to SES 43 (Fig.1a) were examined for splicing enhancer activity. SES 43-1 and SES 43-2 had 55.7 and 58.3% of the control activity, respectively, less than SES 43. SES 43 is the most purine-rich of the three sequenced tested (73.1%) and has three AAG trinucleotides.

**Nuclear protein that binds to splicing enhancer sequence.** Although the mechanisms by which purine-rich sequences enhance splicing remain unclear, members of the SR family of essential splicing factors have been suggested to interact directly with purine-rich enhancer sequences. In order to analyze the interaction of a nuclear protein with SES 43, the SES 43 RNA probe was examined for its ability to bind protein. Nuclear proteins that bind to the SES 43 RNA probe were identified under the conditions in which non-specific binding was abolished and their sizes were approximately 120 kDa and 40 kDa (Fig.4). These correspond to sizes of known SR proteins (7). However, the binding proteins disappeared when the tRNA concentration in the reaction mixture was increased. This indicated that nuclear proteins bind to SES 43 but the interaction is weak. We propose that this weak interaction explains the low splicing enhancer activity compared to that of the control enhancer sequence.

![FIG. 4. Nuclear proteins that bind to SES 43 RNA probe. The mixture of HeLa cell nuclear extract and SES 43 RNA probe was gel-electrophoresed. Two bands corresponding to 120 kDa and 40 kDa were visualized in low concentration of tRNA (L) whilst no band in high concentration (H).](image)

**Splicing enhancer sequence within exon 53.** Since SES 53 did not have significant splicing enhancer activity, we searched for a sequence with higher activity in exon 53. The full-sequence of exon 53 except for 16nt at either end was divided into six 30nt fragments (SES 53-2 to 7). SES 53-2, SES 53-5 and SES 53-6 showed relatively high enhancer activity.
(69.9, 79.6 and 80.2% of the control, respectively) while SES 53-3, -4 and -7 showed low activity (19.6, 42.3 and 17.5% of the control, respectively) (Fig.1c). An additional test sequence (SES 53-1) was created that corresponded to a displacement of 5 bases relative to 53-2. Remarkably, this sequence also had high activity (73.5% (Fig.1c)). These results indicated that splicing enhancer activity is located at both the 5' and 3' ends of exon 53. We were unable to find a consensus sequence in the four segments that had a relatively high enhancer activity. It was remarkable that SES 53-5 showed relatively high activity even though it overlapped SES 53 with low enhancer activity (Fig.1c). Therefore, the non-overlapping sequence TCTTAGGACAGGC in SES 53-5 seems to be important for splicing enhancer activity.

**Splicing enhancer sequence within exon 46.** SES 46 had very low splicing enhancer activity (23.6% of the control) even though it is purine-rich (76.7%) and has three AAG trinucleotides (Fig.1b). Since we reasoned that every exon in the dystrophin gene must have a splicing enhancer sequence to be properly processed, we searched for a splicing enhancer sequence in exon 46. The full-sequence of exon 46 except for 13nt or 15nt at each end was divided into four 30nt long sequences (SES 46-1, -2, -3 and -4) (Fig.1b), each of which was examined for splicing enhancer activity. Remarkably, all four sequences had relatively low activity (10.4, 12.9, 7.9 and 27.9% of the control, respectively). We conclude that exon 46 does not have a splicing enhancer sequence that can be detected using the chimeric gene system.

![Graph](https://via.placeholder.com/150)

**Content of AG nucleotides**

**FIG. 5.** Splicing enhancer activity and AG content of 16 examined sequences. Relative splicing enhancer activity (Y-axis) of each fragment was plotted against its percentage of A and G nucleotide content (X-axis). The value of exon 19 sequence is indicated as a control (triangle). There was no significant association between two parameters.

**Splicing enhancer activity and purine-rich sequence.** In order to examine the theory that purine-rich sequences can act as exon splicing enhancers, the association between the abundance of purine nucleotides in the 16 sequences examined and their splicing enhancer activities was analyzed (Fig.5). Remarkably, purine content did not correlate with splicing enhancer activity (Fig.5). Since four sequences (SES43, SES53-1, SES53-5, and SES53-6) exhibited activity of more than 70% of the control, we searched them for common features.
All four sequences had a purine content of more than 70% and at least 2 AAG trinucleotides. These results suggested that purine-rich sequence including at least 2 AAG sequence is a minimum requirement for splicing enhancer activity. However, these requirements are insufficient to confer splicing enhancer activity because two other sequences with the same characteristics (SES 46 and SES 46-4) showed low enhancer activity. We could not identify any other shared characteristics such as primary or secondary structure or location in the four splicing enhancer sequences identified here.

**DISCUSSION**

Since the dystrophin gene is characterized by the large size of its introns, exonic splicing enhancer sequence would be expected to be an essential factor for proper splicing intron. However, a comprehensive study has not been carried out on exonic splicing enhancer sequences in exons of the dystrophin gene. Here, we searched for splicing enhancer sequences in three exons in the central region containing the largest intron and found the following; (i) a purine-rich sequence is not necessarily a splicing enhancer sequence, (ii) two separate sequences in exon 53 showed similar enhancer activity, and (iii) exon 46 did not have any detectable splicing enhancer sequence. These results suggested that splicing of different exons in the dystrophin gene might be controlled by different factors.

Since purine-rich sequences have been claimed to act as splicing enhancers (26) and since we have identified purine-rich splicing enhancer sequences in two exons (exons 19 and 27) (18,24), we tested the splicing enhancer activity of the most purine-rich sequences from three exons (exon 43, 46 and 53) (Fig.1). However, only one of them (SES 43) showed significant activity (70.6% of the control) (Fig.3). Furthermore there was no association between purine content and splicing enhancer activity (Fig.5). These results indicate purine content is not the sole factor determining splicing enhancer. Although the AAG trinucleotide is proposed to a consensus sequence for splicing enhancer (25), the number of AAG triplets did not correlate with splicing activity (Fig.1). Our results suggested other factors might promote proper splicing.

Purine-rich splicing enhancers have been identified in many regulated and constitutively spliced pre-mRNAs. Most are located in exons and, in general, they activate the splicing of the upstream intron by promoting the use of a weak 3' splice site. Shapiro's splicing probability scores for the 3' splice sites of upstream introns 42, 45 and 52 were 81.4, 85.6 and 90.0, respectively (17). Since intron 42 has the weakest splice site, SES43 might promote its specific recognition. Although exon 53 is predicted to have a strong 3' splice site, it contains two enhancer sequences. Therefore, there does not appear to be a correlation between the predicted strength of the 3' splice site and the presence of detectable enhancer sequences.

Two separate regions in exon 53 showed splicing enhancer activity (Fig.1). The presence of two distinct splicing enhancer sequences in one exon has been reported in *Drosophila*, rat and bovine genes (5,12,16). They are located in exons that are alternatively spliced and found to bind different nuclear proteins. Although the incorporation of exon 53 may be regulated by nuclear protein, we are not aware of any reports of alternative splicing of exon 53. This is the first report of the identification of two distinct splicing enhancer elements within one exon of a human gene, which suggests that splicing of dystrophin transcript is also regulated by a complex system.

In this study, we employed an *in vitro* splicing system to analyze the activation of the *Drosophila dsx* minigene intron 3 splicing. This system has been employed to analyze the splicing enhancer activity of human gene transcripts from both the dystrophin gene and also the SMN gene (11,18). In addition, this system was used in studies of viral transcripts (28).
Therefore, our system is appropriate to search for splicing enhancer sequences. However, none of the sequences corresponding to segments of exon 46 showed significant splicing enhancer activity (Fig.1). This may be due to the limitation of our system, which measures the ability to activate upstream intron splicing from one intron pre-mRNA (Fig.2). Therefore, the sequences examined might exhibit splicing enhancer activity at least two introns are included in the analysis.

Exon 46 comprises 148 nt but does not appear to have significant splicing enhancer sequence. This may be due to the fact that possesses good splicing consensus sequences at 3’ and 5’ splice sites (Shapiro’s score for these splicing sites are 85.6 and 97.3, respectively). Since introns 45 and 46 comprise 20 kb and 11 kb, respectively, we considered it probable that this exon would need an additional factor to promote proper splicing, although we were unable to prove this experimentally. A splicing enhancer sequence is presumably present in the intron sequences as in other genes (2,14,19).

Recently, the list of known splicing enhancers has been expanded to include AC-rich sequences (3). In the experiments reported here, any AC-rich region with splicing enhancer activity in the tested sequence would not have been identified. Therefore, the possibility that AC-rich sequences have splicing enhancer activity in the dystrophin gene seems low. Recently, an octamer consensus, GRYYcSYR, has been proposed as an exonic splicing enhancer that is recognized by the SR protein SC35 (9). Again no such octamer consensus sequence was identified in our sequences examined here. In addition, the pentanucleotide A,C,A/G/T,G,G/C has been proposed as a consensus sequence for binding of SR 40 C (7). Every sequence bearing this consensus sequence (SES 43, SES 53-5 and the control) had significant enhancer activity. Thus, it is interesting that a 40 kDa nuclear protein bound loosely to SES 43 RNA probe (Fig.4). SR proteins have been reported to bind to splicing enhancer sequences. The ability to bind SR proteins has been correlated with the ability of enhancers to activate splicing of the upstream intron (20). Furthermore, in vitro splicing of an exon with weak splicing signals can be stimulated by incorporating a high-affinity SR protein binding site within the exon (23). We found that two proteins could bind to SES 43 RNA probe (Fig.4). Both proteins were similar in size to known SR proteins. However, the two proteins were never visualized when RNA probe and nuclear extract were mixed in the presence of a high concentration of competing tRNA. These results suggested that both proteins bind SES 43 with low affinity.

The analysis of splicing enhancer sequences leads to elucidation of the molecular mechanism of splicing in the complicated dystrophin gene and the results might be relevant for the treatment of Duchenne muscular dystrophy (13). In our previous report, we showed that an antisense oligonucleotide complementary to splicing enhancer sequence within exon 19 (88 bp) induced exon 19 skipping (15,24) and we proposed that exon skipping might be applied to correct out-of-frame mutations into in-frame mutations, as in a case of DMD having deletion of exon 20 (242 bp) (13). Very strong splicing enhancer sequences might be used to derive antisense to induce exon skipping.

ACKNOWLEDGMENTS

We thank Dr. A. Pugsley for advice and for critically reading the manuscript, and Miss N. Kageyama for her help in preparing the manuscript. This work was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan and the National Center of Neurology and Psychiatry (NCNP) of the Ministry of Health and Welfare of Japan.
SPLICING ENHANCER SEQUENCE IN THE DYSTROPHIN GENE

REFERENCES


