J. Biochem. **143,** 303–310 (2008)
doi:10.1093/jb/mvm227

A Strong Exonic Splicing Enhancer in Dystrophin Exon 19 Achieve Proper Splicing Without an Upstream Polypyrimidine Tract

Yasuaki Habara[1], Mari Doshita[2], Sadako Hirozawa[2], Yuka Yokono[2], Mariko Yagi[1,*], Yasuhiro Takeshima[1] and Masafumi Matsuo[1]

[1]*Department of Pediatrics, Kobe University Graduate School of Medicine, Kobe 650-0017; and* [2]*Kobe Pharmaceutical University, Kobe 658-8558, Japan*

Received October 17, 2007; accepted November 3, 2007; published online November 26, 2007

Proper splicing is known to proceed under the control of conserved *cis*-elements located at exon–intron boundaries. Recently, it was shown that additional elements, such as exonic splicing enhancers (ESEs), are essential for the proper splicing of certain exons, in addition to the splice donor and acceptor site sequences; however, the relationship between these *cis*-elements is still unclear. In this report, we utilize dystrophin exon 19 to analyse the relationship between the ESE and its upstream acceptor site sequences. Dystrophin exon 19, which maintains adequate splicing donor and acceptor consensus sequences, encodes exonic splicing enhancer (dys-ESE19) sequences. Splice pattern analysis, using a minigene reporter expressed in HeLa cells, showed that either a strong polypyrimidine tract (PPT) or a fully active dys-ESE19 is sufficient for proper splicing. Each of these two *cis*-elements has enough activity for proper exon 19 splicing suggesting that the PPT, which is believed to be an essential *cis*-element for splicing, is dispensable when the downstream exon contains a strong ESE. This compensation was only seen in living cells but not in '*in vitro* splicing'. This suggests the possibility that the previous splicing experiments using an *in vitro* splicing system could underestimate the activity of ESEs.

Key words: acceptor site sequences, dystrophin, exonic splicing enhancer, polypyrimidine tract, splicing.

Abbreviations: AO, antisense oligonucleotide; AR, ampicillin resistant gene; BMD, Becker muscular dystrophy; BPS, branch point sequence; CV, Shapiro's splice site consensus value; DK, dystrophin Kobe; DMD, Duchenne muscular dystrophy; dys-ESE19, dystrophin exonic splicing enhancer 19; ESE, exonic splicing enhancer; PPT, polypyrimidine tract; RT–PCR, reverse transcription–PCR.

The dystrophin gene, one of the largest genes in the human genome, is characterized by the extraordinary length of its introns (*1*), and its splicing regulatory mechanism is not fully understood. In dystrophin Kobe, which harbours a 52 nt deletion in exon 19 of the dystrophin gene, exon 19 was totally skipped in mature mRNA; however, this deletion does not disturb either splicing donor or acceptor site consensus sequences (*2*). This exon skipping suggested that the deleted region contains an exonic splicing enhancer (ESE), a *cis*-acting splicing signal located within an exon that stimulates the inclusion of that exon in the mature mRNA (*3*). ESEs are thought to be recognized by protein factors such as SR proteins, which comprise a family of splicing factor containing an N-terminal RNA-binding domain and a C-terminal RS (arginine/serine-rich) domain (*4*). The exon 19 skipping observed in patients was replicated when an antisense oligonucleotide (AO), which hybridized to the region deleted in dystrophin Kobe, was introduced into cultured cell (*5*). Both *in vitro* and *in vivo* studies suggest that part of the exon sequence is essential for proper splicing; this part was named dystrophin exonic splicing enhancer 19 (dys-ESE19).

It is not clear why exon 19 requires ESE for proper splicing. Our previous study indicated that dys-ESE19 is necessary for intron 18 splicing, but not for intron 19 splicing (*6*). This suggested that the acceptor site of exon 19 is a weak splice site, as it has been suggested that ESEs are essential for the proper splicing of an exon with weak splice consensus sequences (*4*, *7*). However, the Shapiro's splice site consensus value (CV) for the acceptor site of exon 19 (*8*) is not notably low among all of the 78 acceptor sites in the dystrophin gene exons: the CV for the exon 19 acceptor site is 0.91, while the average CV for all 78 acceptor sites in dystrophin gene exons is 0.87 (*9*). Therefore, there are unknown factors or rules that determine the exon 19 splicing patterns.

In this report, we investigate the relationship between dys-ESE19 and the quality of a polypyrimidine tract (PPT) using a minigene reporter made from dystrophin exons 18, 19, 20 and the adjacent introns. Our results suggest that either a strong PPT or a fully active dys-ESE19 is sufficient for the correct splicing of dystrophin exon 19.

MATERIALS AND METHODS

Minigene Reporter Construction—To study splicing patterns, minigene expression plasmids containing three exons (exon 18, 19 and 20 of the human dystrophin gene)

*To whom correspondence should be addressed. Tel: +81-78-382-6090, Fax: +81-78-382-6099, E-mail: myagi@med.kobe-u.ac.jp

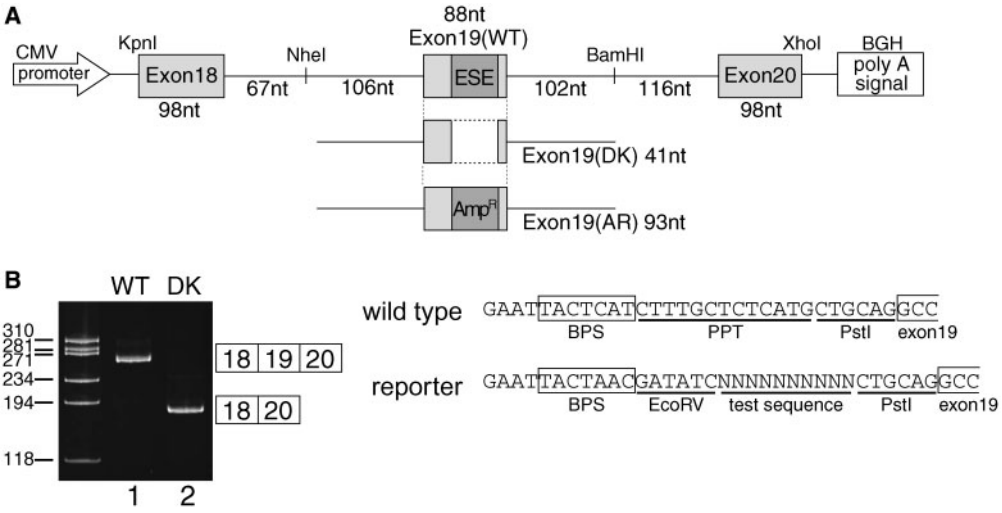

Fig. 1. **The structure of the dystrophin exon 18-19-20 minigene and examples of its splice pattern analysis.** (A) The structure of the dystrophin exon 18-19-20 minigene. Dystrophin exons 18, 19 and 20, and the adjacent introns, were amplified by PCR using primers containing a restriction enzyme cutting site, as shown in the figure. Then, three fragments (KpnI–NheI, NheI–BamHI and BamHI–XhoI) were ligated step by step using cloning vectors, and finally cloned into the pcDNA3 human expression vector with KpnI and XhoI. Two modified exon 19 constructs—DK exon 19, from which dys-ESE19 was replaced by HindIII cutting site, and AR exon 19, in which dys-ESE19 was replaced by a part of the AR gene—are shown below the WT exon 19. Open arrow and open box indicate the CMV promoter and the BGH polyadenylation signal of the pcDNA3 vector, respectively. Shaded boxes and lines indicate dystrophin exons and their flanking introns, respectively. Dark shaded boxes are dys-ESE19 (in WT exon 19) or a part of the AR gene (in AR exon 19). The sizes of the dystrophin exons and the flanking introns are also shown. The actual minigene sequences are shown in Supplementary Data Fig. S1. (B) The result of minigene reporter splice patterns. HeLa cells transfected with the minigene reporter plasmid were harvested 24 h after transfection using ISOGEN. After RNA purification and RT–PCR, samples were separated on an 8% polyacrylamide gel and stained with ethidium bromide. A 267 bp fragment (lane 1) shows that the WT exon 19 is included in the mature mRNA and a 179 bp fragment shows exon 19 skipping (lane 2). A schematic description of these RT–PCR products is shown on the right. The names of the exon 19 types are shown on the top. (C) The splice acceptor site sequence of human dystrophin exon 19 (wild type) and the sequence of the same region of the minigene reporter for PPT sequence exchange analysis (reporter). Boxes show a BPS and the first 3 nt of exon 19. Between the EcoRV and PstI sites, 11 different sequences of PPT (see Fig. 2) were inserted and tested.

and the adjacent introns were constructed (Fig. 1A) and cloned into a pcDNA3 mammalian expression vector (Invitrogen, Carlsbad, CA, USA) using KpnI and XhoI sites. The actual minigene sequences are shown in Supplementary Data Fig. S1. For minigene construction and other splicing analysis, we used standard molecular cloning methods (*10*). To make PPT- and branch point sequence (BPS)-modified reporters, an altered BPS and an EcoRV recognition site were introduced into intron 18 of the minigene using PCR-based methods. Synthesized PPT oligonucleotides containing EcoRV and PstI sites were then cloned into the same restriction site (Fig. 1C). All minigene sequences amplified by PCR were confirmed by DNA sequencing using an ABI model 310 genetic analyser (Applied Biosystems, Foster City, CA, USA).

Transfection—HeLa cells were grown in 12-well plates to ~70% confluency in DMEM (Dulbecco's modified Eagle's medium) containing 5% fetal bovine serum (Trace Biosciences, Castle Hill, Australia) at 37°C under 5% CO_2. Minigene reporter expression plasmids (0.75 μg each) were transfected into cells using Plus Reagent and Lipofectamine (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. Cells were harvested 24 h after transfection and total RNA was extracted using ISOGEN (Nippon Gene Co., Ltd., Japan).

Analysis of Splicing Products—Five micrograms of total RNA was subjected to reverse transcription (RT) using random hexamer primers or specific primers (YH353, 5′-AAGTCTCTCACTTAGC-3′) with M-MLV reverse transcriptase (Invitrogen, Carlsbad, CA, USA) in a total volume of 20 μl. PCR was performed using a forward primer corresponding to a segment of exon 18 (YH307, 5′-ATTACTCGCTCAGAAGCTGTGTTGC-3′) and a reverse primer complementary to a segment of exon 20 (YH308, 5′-AAGTCTCTCACTTAGCAACTGGCAG-3′). Amplification was carried out in a total volume of 20 μl containing 4 μl of cDNA, 2 μl of 10 × Ex *Taq* Buffer, 2 μl of 2.5 mM dNTPs, 10 pmol of each primer and 1 U of Ex *Taq* Polymerase (Takara Bio, Inc., Kyoto, Japan). PCR cycling conditions were as follows: initial denaturation at 94°C for 2 min followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 58°C for 1 min, extension at 72°C for 2 min and a final extension at 72°C for 5 min. PCR products were separated on an 8% polyacrylamide gel and stained using ethidium bromide.

In Vitro Splicing Assay—To make substrate pre-mRNA, template DNA fragments for the *in vitro* transcription were synthesized by standard PCR amplification using forward [T7(*19*), 5′-TAATACGACTCACTATAGG-3′] and reverse (YH304, 5′-CTCGAGCAGCCAGTTAAGTCTCTC AC-3′) primers with plasmid templates, which were the same plasmids utilized in the splicing assays in HeLa cells. Then, fragments were purified using a MiniElute PCR Purification Kit (QIAGEN, GmbH, Germany) and purified

name, sequences \ exon19, BPS		WT UAAC	WT UAAU	WT UCAC	WT UCAU	DK UAAC	DK UAAU	DK UCAC	DK UCAU	AR UAAC
U10	UUUUUUUUUU	+	+	+	+	+	+	+	+	+
U8G2	UUGUUUUGUU	+				+	+	+	+/−	+
U8A2	UUAUUUUAUU	+				+	+	+	+/−	+
U6G4	UUGUUGUGUG	+				−			−	+/−
U6A4	UUAUUAUAUA	+				−			−	+/−
C10	CCCCCCCCCC	+				−				−/+
U4G6	GGUGGUGUGU	+	+	+	+	−				
U4A6	AAUAAUAUAU	+	+	+	+	−				
U2G4A4	GAUGGAAUGA	+/−	+/−	−/+	−/+	−				
G6A4	GGAGGAGAGA	−				−				
A6G4	AAGAAGAGAG	−				−				

Fig. 2. **All results of splicing assay using reporter minigenes transfected into HeLa cells.** Name and sequence columns show the names and actual sequences of the PPT in the EcoRV–PstI cassette, which is shown in Fig. 1C. Exon 19 and BPS categories show the exon 19 types (WT: wild type; DK: 52 nt deletion the same as dystrophin Kobe; AR: ampcillin resistance gene fragment inserted in DK) and the last 4 nt of the branch site sequences (the first 3 nt of the branch site are same UAC for all constructs and were not shown in the figure). The splicing patterns are categorized into four types based on the results of acrylamide gel electrophoresis. Exon 19 included: +; exon 19 skipped: −; both included and skipped bands are observed and included was dominant or the densities of the two bands were almost the same: +/−; skipped band was dominant: −/+. Blank cells mean no such plasmid was made. Each cell is coloured as described below: light shaded: + or supposed to be +; dark shaded: − or supposed to be −; shaded medium colour: +/− or −/+; white: unknown.

templates were subjected to *in vitro* transcription using a mMESSAGE mMACHINE T7 Kit (Ambion, Austin, TX, USA). After DNaseI treatment, the resultant RNAs were purified on Quick Spin Columns G-50 (Roche, IN, USA) and the concentration and size of RNAs were confirmed using an ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA) and an Agilent2100 bioanalyser with an RNA nano kit (Agilent Technologies, Santa Clara, CA, USA), respectively. Standard *in vitro* splicing reactions were carried out at 37°C for 2 h in a total volume of 20 μl containing 50% (v/v) HeLa nuclear extract (HNE; CIL BIOTEC, Mons, Belgium), 1.6 mM $MgCl_2$, 0.5 mM ATP, 20 mM creatine phosphate, 25 ng pre-mRNA and 20 U of RNaseOUT (Invitrogen, Carlsbad, CA, USA). After ProteaseK treatment, RNAs were extracted by phenol–chloroform extraction and ethanol precipitation. Splice patterns were analysed using the same RT–PCR methods as described above, except that the PCR cycle number was reduced to 15. When samples showed both exon included and skipped bands, these two bands were quantified by capillary electrophoresis using an Agilent2100 bioanalyser with a DNA1000 kit. The molecular ratio of the upper and lower bands was calculated using an index of peak height number divided by nucleotide length.

RESULTS

Effects of PPT Alteration with/without ESE on the Exon 19 Splice Pattern—To investigate dys-ESE19 function during the splicing reaction, we first synthesized minigene reporters, which contain dystrophin exons 18, 19, 20 and the adjacent introns (Fig. 1A). Minigenes containing a wild-type exon 19 (WT) or a dys-ESE19 deleted exon 19, mimicking the dystrophin Kobe mutation (DK), were cloned downstream of a CMV promoter in pcDNA3 vectors and transfected into HeLa cells. One day after transfection, cells were harvested and splicing patterns were analysed by RT–PCR. The minigene with the WT exon 19 showed a normal splicing pattern (Fig. 1B, lane 1), whereas the minigene with the DK exon 19 showed an exon 19-skipped splice pattern (Fig. 1B, lane 2). This result confirmed that this minigene reporter system reproduces the *in vivo* splicing patterns seen in both healthy controls and the DK patient who has a 52 nt dys-ESE19 deletion in exon 19 (*2*). This splice pattern was also confirmed in other cell-lines (MRC5 and NIH3T3, data not shown). We concluded that this reporter system recapitulates the *in vivo* splice pattern as all of the cell-lines used showed the same splice patterns as healthy controls and the patient; we used HeLa cells for following experiments.

The splice acceptor site sequence contains three different elements (*11*, *12*): a BPS, a PPT and a 3′end consensus sequence AG. First, we analysed the relationship between the PPT and the dys-ESE19. For this purpose, we modified the PPT of exon 19 in reporter plasmids to make minigenes with different PPT sequences (Fig. 1C). In total, 11 different 10 nt artificial PPTs were inserted between the BPS and the 3′ splice site consensus sequence, and the splicing pattern with or without dys-ESE19 was analysed (Fig. 2). As the most efficient type of PPT sequence is a poly(U)-tract (*13*, *14*), a U10 test sequence was examined first. Reporters containing a U10 PPT and the WT exon 19 showed a splice pattern that included exon 19, as expected (Fig. 3A, lane 1). Interestingly, reporters with a U10 PPT and the DK exon 19 also showed a splice pattern that included exon 19 (Fig. 3A, lane 2; the band size is smaller than that in lane 1 because the DK exon 19 is shorter than the WT exon 19), indicating that dystrophin exon 19 does not require the ESE if the upstream PPT has enough strength. It also suggests the possibility that dystrophin exon 19 requires the ESE because the upstream PPT is too weak to define exon 19, even though the PPT contains enough pyrimidines (10 out of 13 nt) and its CV is sufficiently high (see 'Introduction' section). To elucidate the relationship between PPT sequence and ESE requirement, we tested a series of minigenes with different PPTs, in which pairs of uridines were replaced

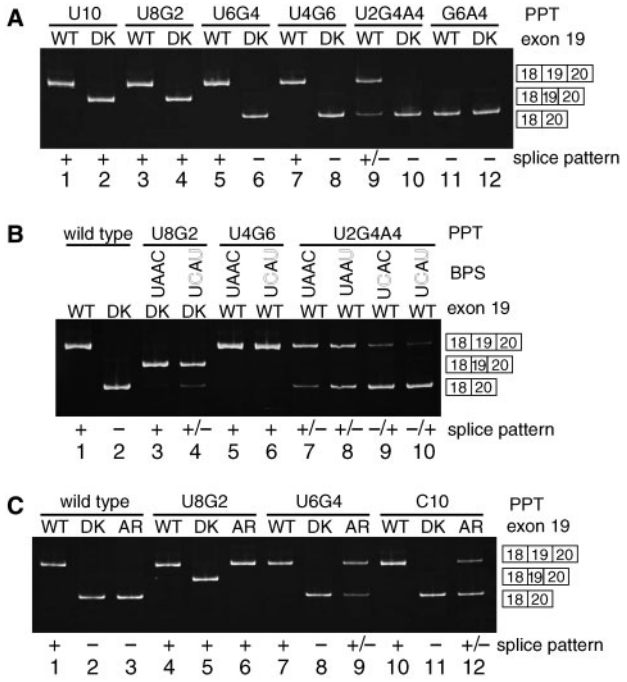

Fig. 3. **Splice pattern analysis using HeLa cells.** (A) Representative results of PPT exchange experiments. Samples were prepared in the same way as those shown in Fig. 1B. A 267 bp fragment (for example, lane 1) and a 179 bp fragment (for example, lane 6) show that WT exon 19-included and skipped splice pattern shown as Fig. 1B, respectively. A 220 bp fragment (for example, lane 2) shows that DK exon 19 is included in the mature mRNA. A schematic description of these RT–PCR products is shown on the right. The names of the tested PPTs (Fig. 2) and exon 19 types (Fig. 1A) are shown on the top. Splice patterns for exon 19 inclusion (+) and exon 19 skipping (−) are shown on the bottom. If the sample showed both included and skipped bands, when included bands were dominant they are shown as +/−, and when skipped bands were dominant they are shown as −/+. (B) Representative results of BPS exchange experiments. Samples were prepared in the same way as above. The names of the tested PPTs (Fig. 2), the last 4 nt of the BPS and the type of exon 19 are shown on the top. PPT name 'wild type' represents the same sequence as the genomic dystrophin gene, the same as the upper sequence shown in Fig. 1C. The first 3 nt (UAC) of the BPS are not shown as they are the same for all tested pre-mRNAs; the last 4 nt are shown in solid letters where they are the same nucleotides as consensus (UAAC) and outlined letters represent mutated nucleotides. A schematic description of the RT–PCR products and splice pattern for exon 19 is shown as in (A). (C) Representative results of exon length exchange experiments. Samples were prepared in the same way as above. The names of the tested PPTs, type of exon 19, schematic description of RT–PCR products and splice pattern for exon 19 are shown as in (A).

with purines or all 10 uridines were replaced with cytosines (Fig. 2). The results showed that dys-ESE19 is not necessary for proper splicing if the PPT contains eight or more uridines; however, if the PPT contains six or fewer uridines, dys-ESE19 is necessary (Fig. 3A, lanes 4 and 6). On the other hand, when the exon contains dys-ESE19, as few as four uridines are sufficient for proper exon 19 usage (Fig. 3A, lane 7), and even two uridines are enough for proper splicing of ~90% of the transcript (Fig. 3A, lane 9; Fig. 4). These results suggest that

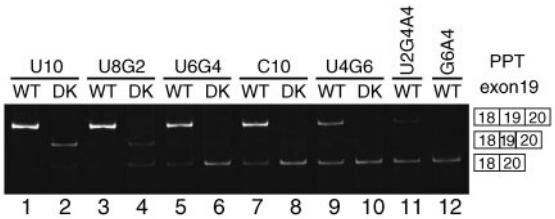

Fig. 4. ***In vitro* splicing analysis using HeLa nuclear extract.** Representative results of an *in vitro* splicing experiment are shown. After the *in vitro* splicing reaction, RNA was purified from the reaction mixture. RT–PCR and acrylamide gel electrophoresis were then performed. The names of the tested PPTs, type of exon 19 and schematic description of RT–PCR products are the same as those shown in Fig. 3.

if exon 19 does not contain the ESE, its proper splicing requires a strong PPT, and that if exon 19 contains the ESE, the upstream PPT is not necessary for proper splicing. It also suggests that the original PPT, which looks like a strong PPT, is not strong enough to promote the inclusion of exon 19 without the ESE.

Effects of BPS Mutation and Exon Length Alteration on Exon 19 Incorporation—Among the three acceptor site elements, the BPS and PPT are known for their variety. The minigene reporter used in the PPT exchange analysis illustrated in Fig. 1C contained the best BPS, UACUAAC (*12*, *15*). On the other hand, the genomic sequence of the intron 18 BPS is UACUCAU, 2 nt of which differ from the best BPS (underlined) (Fig. 1C). We analysed the effect of the alteration of these 2 nt on the splice pattern of exon 19. The fifth and seventh position of the BPS (underlined above) were mutated in minigene reporters, which were cloned into an expression vector and the resultant plasmids were introduced into HeLa cells for splicing pattern analysis using the same methods as above. For the reporter containing the PPT U8G2 and the DK exon 19, the change in the BPS slightly reduced exon 19 usage (Fig. 3B, lanes 3 and 4). For the reporter containing the U2G4A4 PPT and the WT exon 19, of the change in the BPS showed that both the fifth and seventh positions of the BPS are important for proper splicing and that when these are mutated, exon 19 usage was gradually reduced (Fig. 3B, lanes 7–10). For other cases of PPT mutation, no difference was observed due to the change in the BPS (Fig. 2; Fig. 3B, lanes 5 and 6) Considering all of these results together (Fig. 2), these2 nt changes in the BPS has an effect on the splicing efficiency of exon 19 both with and without the dys-ESE19 sequence; however, this change does not have as drastic an effect as a uridine to purine change in the PPT.

It is known that short exons are sometimes difficult to splice correctly. To confirm that the DK exon 19 (41 nt) is not too short for proper splicing compared to the WT exon 19 (88 nt), we constructed minigene reporters in which dys-ESE19 was replaced with other gene fragments that lack the ESE. For this purpose, we used a 46 nt fragment of an ampicillin resistance (AR) gene, which should not have any ESE activity because it is a bacterial gene (Fig. 1A). The fragment was selected by eye, so as not to contain too many purine residues

name, sequences \ exon19		WT	DK	AR
U10	UUUUUUUUUU	+	+/−	
U8G2	UUGUUUUGUU	+	+/−	+/−
U6G4	UUGUUGUGUG	+/−	−	−
C10	CCCCCCCCCC	+/−	−	−
U4G6	GGUGGUGUGU	+/−	−	
U2G4A4	GAUGGAAUGA	−/+	−	
G6A4	GGAGGAGAGA	−		
A6G4	AAGAAGAGAG	−		

Fig. 5. **All results of *in vitro* splicing assay using HNE.** PPTs, exon 19 types and splice pattern categorization are the same as in Fig. 2. For BPS, only UACUAAC was tested for *in vitro* splicing and so is not shown in this figure.

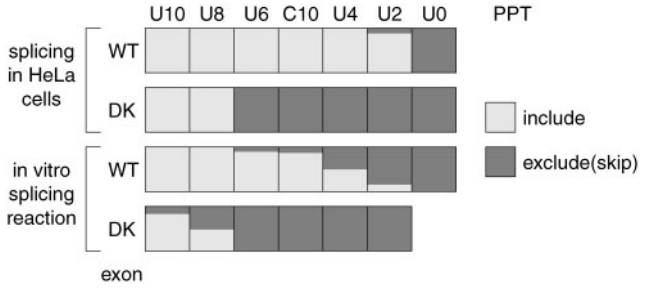

Fig. 6. **Results summary and quantification analysis of the splicing assay in HeLa cells and the *in vitro* splicing reaction with different PPTs**. Each box represents a splicing assay result and the light and dark shadowed areas in the boxes represent the molecular ratio of exon 19 inclusion/skipping. For the *in vitro* splicing assay, samples that showed both exon 19-included and exon 19-skipped splicing patterns (in Fig. 4, lanes 2, 4, 5, 7, 9 and 11) were subjected to capillary electrophoresis in an Agilent2100 and each band was quantified using peak height. For the splicing assay in HeLa cells, only the U2G4A4 sample (Fig. 3A, lane 9), showed both included and skipped splice patterns; this was amplified by a reduced number of PCR cycles (15 cycles) and analysed by same method as above. The molecular ratio of included/skipped was calculated and shown by the extent of light and dark in the shadowed area. The names of the PPTs (purine residues are omitted) are shown on the top; exon types and assay methods are shown on the left.

to avoid cloning a purine-rich ESE by accident (*16*); this fragment showed no ESE activity in an *in vitro* splicing reaction (*6*). When the exon length was expanded, AR exon 19 (93 nt) was incorporated into mRNA with a slightly higher efficiency than the DK exon 19 with a U6G4 PPT or a C10 PPT (Fig. 3C, lanes 9 and 12). Without the presence of the dys-ESE19 sequence, the U8G2 and U6G4 PPT are on the border of exon 19 inclusion/skipping determination (Fig. 3C, lanes 5 and 8). This result suggests that an exon length expansion leads to slightly more effective exon 19 recognition. The 41 nt DK exon 19 is probably slightly shorter than standard exon length required for proper exon recognition. In addition to that, AR exon 19 with a wild-type acceptor site (completely the same as the genomic dystrophin gene acceptor site) did not show exon inclusion (Fig. 3C, lane 3), confirming that dys-ESE19 is essential for the proper splicing of exon 19.

These BPS mutations and exon length alterations suggest that both the BPS and exon length have an effect on exon 19 usage in the splicing reaction; however, the effects were not drastic compared to those with/without ESE or PPT alteration.

In Vitro Splicing Assay with Same Reporter Minigene—Most previous reports that have analysed PPT strength (*13*, *14*) or the relationship between the ESE and the PPT (*17–19*) have used a biochemical method called '*in vitro* splicing' to analyse splicing patterns. We also tried the same *in vitro* splicing system using HNE. The synthesized pre-mRNAs, which had the same exon–intron structures used above, were incubated with HNE at 37°C for 2 h. After the *in vitro* splicing reaction, RNA was extracted and RT–PCR was used to analyse splicing patterns. Interestingly, the results using HeLa cells and HNE are slightly different. In HeLa cells, dys-ESE19 is not necessary for DK exon 19 inclusion when the PPT is U10 or U8G2, so there was no exon 19 skipping in these two reporters (Fig. 3A, lanes 2 and 4); on the other hand, a fraction of these two reporters showed skipping of exon 19 in *in vitro* splicing (Fig. 4, lanes 2 and 4). A similar phenomenon was observed for dys-ESE19-dependent splicing. In HeLa cells, there was no WT exon 19 skipping when the PPT was U6G4, C10 or U4G6 (Fig. 2; Fig. 3A lanes 5 and 7); however, a fraction of the mRNAs showed a WT exon 19-skipped splice pattern following *in vitro* splicing of these reporters (Fig. 4, lanes 5, 7 and 9; Fig. 5). As many samples showed both exon-inclusive and exon-skipped splice patterns, PCR products were quantified by capillary electrophoresis and the molecular ratio of exon inclusion/skipping was calculated (Fig. 6). It seems that on/off (or include/skip) regulation of exons is more strictly controlled in living cells than it is *in vitro* (Fig. 6).

Sequence Analysis of All 78 Acceptor Sites in the Dystrophin Gene—According to the experimental results, eight or more uridine residues in a 10 nt stretch has sufficient activity to be spliced correctly without an ESE. To try to predict how many exons could be spliced in an ESE-independent manner, we scanned all 78 dystrophin acceptor sites to determine whether they have a good PPT or not. If the acceptor site has eight or more uridines in a 10 nt stretch, it was categorized as a good PPT (Supplementary Data, Table S1). When we scanned sequence positions −4 to −20 of the acceptor sites (the 5′-end of the exon is at +1), 25 acceptor sites had a good PPT, and if the sequence position of the upstream end was extended to −35, 6 more sites were added to this category. Thus, approximately one-third of dystrophin exons have a good PPT and can be supposed to be spliced in an ESE-independent manner; meanwhile, the remaining two-thirds of dystrophin exons do not have a good PPT and can be supposed to contain ESEs.

We also scanned the BPSs because these also affected correct splicing efficiency (Fig. 3C). As the BPS is highly degenerate in humans (*12*), only five acceptor sites have a typical YNCTRAC (Y = C or T, R = A or G, N = any nucleotide) sequence in the −60 to −10 position in the acceptor sites (Supplementary Data, Table S1). Using a BPS finding program on an SSF website (Splicing Sequences Finder; http://www.umd.be/SSF), 56 of the 78 acceptor sites in the dystrophin gene have at least one BPS candidate sequence in the −60 to −10 position.

In other words, 22 of the 78 acceptor sites in the dystrophin gene do not have a BPS candidate sequence.

DISCUSSION

Our results show that dys-ESE19 is necessary for the correct splicing of exon 19 to compensate for the poor splicing activity of the PPT in intron 18. Meanwhile, a degenerate PPT, more than half of which contains purine residues, causes no problem as long as exon 19 has a fully active dys-ESE19. It was surprising that a PPT containing only two uridines (minigene with the U2G4A4 PPT) could be properly spliced in the majority of transcripts when exon 19 contained a dys-ESE19. We also showed that a biochemical method called '*in vitro* splicing', which is commonly used for splicing analysis, does not fully recapitulate the splicing pattern in living cells.

The first report to describe the relationship between exon sequence and PPT came out in 1991, using the *doublesex* (*dsx*) gene in *Drosophila melanogaster* (*17*). It was known that exon 4 of the *dsx* gene has a suboptimal PPT (containing 12 pyrimidine residues and seven purine residues) upstream of it and that this exon is skipped in males, but is used in females. Their report suggested that two female-specific proteins (Tra and Tra2) bind to exon 4 to encourage its inclusion in the mature mRNA in females. This report provided a prototype of the hypothesis that ESE-binding proteins bind to exons with suboptimal PPTs and stimulate exon inclusion. Several reports (*7*–*19*) followed this hypothesis, assuming that SR proteins (including Tra2) bind to ESE and stimulate binding of U2AF to the PPT for several contexts of alternatively spliced exons, in several species. Our approach to analyse the relationship between an ESE and splicing *cis*-signals, mainly PPTs, is different from previous reports in two regards. First, we analysed a human constitutively spliced exon, whereas most previous reports have analysed alternatively spliced exons among several species. It is understandable that an alternatively spliced exon would have a suboptimal PPT, because that exon should be skipped in certain conditions. However, our results showed that the constitutively spliced dystrophin exon 19 also has a suboptimal PPT, and that its splicing depends on an ESE. In addition to this, the fact that over half of the dystrophin exons could contain suboptimal PPTs suggests that a number of constitutively spliced exons, possibly over half, contain suboptimal PPTs, the proper splicing of which is dependent on an ESE. Second, we mainly used cultured cells for the splicing reporter assay, whereas a number of previous reports have used *in vitro* splicing. As far as we know, the U2G4A4 PPT that we used here is the weakest PPT, yet it showed an almost normal splicing pattern in the reporter assay that was dependent on an ESE. Why would a predominantly 'purine-tract' work well in our assay and not in others? We think this is because most previously reported splicing analysis experiments used an *in vitro* splicing system with HNE, whereas we used living cells. The on/off control of each exon is probably more strictly controlled in living cells than in nuclear extract; furthermore, ESE activity in living cells is stronger than in nuclear extract. These differences between living cells and *in vitro* splicing also suggest the possibility that when other ESE experiments show the same tendency, ESE activities could be underestimated in most splicing pattern analysis, because the experiments were performed using nuclear extract.

It is said that 'weak splice sites need ESE' (*4*); however, there is not yet a good standard method to evaluate acceptor site strength. As opposed to a splicing donor site in which the consensus sequence is relatively simple (AG|GURAGU), a splicing acceptor site consensus sequence is complicated since it contains three different elements: a BPS, a PPT and a 3′end consensus sequence AG (*11*, *12*). In addition to this, the length of the PPT and the pyrimidine/purine nucleotide ratio are various in each acceptor site, as a previous report suggested that when a PPT has a long uridine stretch, the location of the uridine stretch between the BPS and the 3′-splice site AG is variable, and that there might also be a purine stretch between these elements (*14*). YNCURAC is the consensus BPS; however, this sequence is highly degenerate and it is usually difficult to predict an actual BPS. For instance, only 5 introns have a typical YNCTRAC sequence among all 78 acceptor sites of the dystrophin gene (Supplementary Material, Table S1). Because of this flexibility, there is no good scoring method for splice acceptor sites. CV is a good indicator of splice site strength; however, it accounts for only 15 nt (−14 to +1) around the 3′ splice site (*8*) and is not sufficient to evaluate splice site strength accurately. Fairbrother *et al*. (*20*) used longer sequences (−22 to +2) than Shapiro's CV to search for ESEs in 'weak exons' and so this scoring method should be more informative than the CV; however, their calculation method also did not account for the BPS. During the actual splicing reaction in living cells, splice sites should be chosen on the basis of the balance between the strengths of each *cis*-acting signal. There are a number of reports about the role of *cis*-acting signals and how to evaluate their strength. However, the relationship between them remains unclear. There are not yet enough experiment-based conclusions to predict the contribution to splicing of each *cis*-acting signal (*21*).

The reason why we want to understand the relationship between these *cis*-acting signals and their relative contributions is for the accurate diagnosis and therapy of genetic disorders. In the case of the dystrophin gene, when a mutation (in most cases, deletion of a single exon or several exons from the genome) in the gene results in the absence of protein due to disruption of the translational reading frame of the mRNA, that patient will show a severe Duchenne muscular dystrophy (DMD) phenotype. On the other hand, if several exons are deleted but the translational reading frame is maintained correctly, the patient usually expresses a truncated protein, which is partly active, and shows a milder Becker muscular dystrophy (BMD) phenotype. Thus, for the therapy of DMD, AO-induced exon skipping, which could correct the translational reading frame from out-of-frame to in-frame in DMD patients, could possibly change those patients' phenotype from severe DMD to the milder

BMD (*1*, *22*, *23*). The mechanism of this exon skipping is thought to be through the hybridization of the AO to an ESE, thereby abolishing the exon recognition (*24*). Our results suggested that two-thirds of dystrophin exons are likely to be good therapeutic targets for AO, because two-thirds of the acceptor sites in the dystrophin gene do not contain a strong PPT and so are likely to have an ESE in the downstream exon. Indeed, exon 19, which contains an ESE and does not have a strong PPT upstream of it, is a good therapeutic target exon for AO treatment. This exon was skipped by AO in *in vitro* splicing experiments (*6*) and in primary cultures of cells from a DMD patient with an exon 20 deletion in the dystrophin gene (*5*). The expression of newly synthesized dystrophin protein induced by the AO was also confirmed in cultured cells (*5*). This AO treatment was finally applied to the patient and in-frame mRNA and dystrophin protein expression was confirmed (*25*). These results suggest that two-thirds of dystrophin exons are good candidate therapeutic targets; on the other hand, it is likely not a good idea to try exon skipping for the remaining one-third of exons.

In summary, we show that either a strong PPT or a strong ESE is sufficient for proper exon recognition during the splicing reaction. It was surprising that a predominantly 'purine tract' does not disturb splicing when the downstream exon has an ESE, suggesting that when a downstream exon contains a strong ESE, the PPT is almost dispensable. This high activity of an ESE is seen only in living cells, not in '*in vitro* splicing' reactions, suggesting the possibility that previous experiments may have underestimated ESE activity. Analysis of all of the dystrophin acceptor sites suggests that two-thirds of the dystrophin exons are potentially spliced in an ESE-dependent manner and the remaining one-third are possibly ESE-independent exons. Here we elucidated the relationship between the dys-ESE19 and the upstream PPT, only one relationship of many between two *cis*-acting signals that regulate splicing patterns. Splicing events are the result of a fine 'balance of power' between a number of regulatory elements (*21*); thus, clarifying each relation could be a way to achieve an understanding of the splicing machinery and it enables the development of therapies for genetic disorders by splicing-associated methods.

Supplementary Data are available at JB Online.

We thank Dr K. Inoue and Dr K. Saiki for helpful advice and discussion. This work was supported by grant from the Ministry of Education, Culture, Sports, Science and Technology. Y.H. was supported by a grant from the Japan Foundation of Neuroscience and Mental Health.

REFERENCES

1. Matsuo, M. (1996) Duchenne/Becker muscular dystrophy: from molecular diagnosis to gene therapy. *Brain Dev.* **18**, 167–172
2. Matsuo, M., Masumura, T., Nishio, H., Nakajima, T., Kitoh, Y., Takumi, T., Koga, J., and Nakamura, H. (1991) Exon skipping during splicing of dystrophin mRNA precursor due to an intraexon deletion in the dystrophin gene of Duchenne muscular dystrophy Kobe. *J. Clin. Invest.* **87**, 2127–2131
3. Blencowe, B.J. (2000) Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.* **25**, 106–110
4. Graveley, B.R. (2000) Sorting out the complexity of SR protein functions. *RNA* **6**, 1197–1211
5. Takeshima, Y., Wada, H., Yagi, M., Ishikawa, Y., Minami, R., Nakamura, H., and Matsuo, M. (2001) Oligonucleotides against a splicing enhancer sequence led to dystrophin production in muscle cells from a Duchenne muscular dystrophy patient. *Brain Dev.* **23**, 788–790
6. Takeshima, Y., Nishio, H., Sakamoto, H., Nakamura, H., and Matsuo, M. (1995) Modulation of *in vitro* splicing of the upstream intron by modifying an intra-exon sequence which is deleted from the dystrophin gene in dystrophin Kobe. *J. Clin. Invest.* **95**, 515–520
7. Tian, M. and Maniatis, T. (1994) A splicing enhancer exhibits both constitutive and regulated activities. *Genes Dev.* **8**, 1703–1712
8. Shapiro, M.B. and Senapathy, P. (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15**, 7155–7174
9. Sironi, M., Pozzoli, U., Cagliani, R., Comi, G.P., Bardoni, A., and Bresolin, N. (2001) Analysis of splicing parameters in the dystrophin gene: relevance for physiological and pathogenetic splicing mechanisms. *Hum. Genet.* **109**, 73–84
10. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA
11. Kramer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65**, 367–409
12. Burge, C.B., Tuschl, T., and Sharp, P.A. (1999) Splicing of precursors to mRNAs by the spliceosomes. in *The RNA World* (Gesteland, R.F., Chek, T.R., and Atkins, J.F., eds.) 2nd edn. pp. 525–560, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA
13. Roscigno, R.F., Weiner, M., and Garcia-Blanco, M.A. (1993) A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *J. Biol. Chem.* **268**, 11222–11229
14. Coolidge, C.J., Seely, R.J., and Patton, J.G. (1997) Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Res.* **25**, 888–896
15. Zhuang, Y.A., Goldstein, A.M., and Weiner, A.M. (1989) UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci. USA* **86**, 2752–2756
16. Tanaka, K., Watakabe, A., and Shimura, Y. (1994) Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol. Cell. Biol.* **14**, 1347–1354
17. Hoshijima, K., Inoue, K., Higuchi, I., Sakamoto, H., and Shimura, Y. (1991) Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in *Drosophila*. *Science* **252**, 833–836
18. Wang, Z., Hoffmann, H.M., and Grabowski, P.J. (1995) Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* **1**, 21–35
19. Graveley, B.R., Hertel, K.J., and Maniatis, T. (2001) The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* **7**, 806–818
20. Fairbrother, W.G., Yeh, R.-F., Sharp, P.A., and Burge, C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science* **297**, 1007–1013
21. Buratti, E., Baralle, M., and Baralle, F.E. (2006) Defective splicing, disease and therapy: searching for master checkpoints in exon definition. *Nucleic Acids Res.* **34**, 3494–3510

22. Aartsma-Rus, A., Bremmer-Bout, M., Janson, A.A.M., den Dunnen, J.T., van Ommen, G.-J.B., and van Deutekom, J.C.T. (2002) Targeted exon skipping as a potential gene correction therapy for Duchenne muscular dystrophy. *Neuromusc. Disord.* **12**, S71–S77
23. McClorey, G., Fletcher, S., and Wilton, S. (2005) Splicing intervention for Duchenne muscular dystrophy. *Curr. Opin. Pharmacol.* **5**, 529–534
24. Aartsma-Rus, A., De Winter, C.L., Janson, A.A.M., Kaman, W.E., Van Ommen, G.-J.B., Den Dunnen, J.T., and Van Deutekom, J.C.T. (2005) Functional analysis of 114 exon-internal AONs for targeted DMD exon skipping: indication for steric hindrance of SR protein binding sites. *Oligonucleotides* **15**, 284–297
25. Takeshima, Y., Yagi, M., Wada, H., Ishibashi, K., Nishiyama, A., Kakumoto, M., Sakaeda, T., Saura, R., Okumura, K., and Matsuo, M. (2006) Intravenous infusion of an antisense oligonucleotide results in exon skipping in muscle dystrophin mRNA of Duchenne muscular dystrophy. *Pediatr. Res.* **59**, 690–694