Multisite and bidirectional exonic splicing enhancer in CD44 alternative exon v3

Elena Vela, Josep M. Hilari, Xavier Roca, Ana M. Muñoz-Mármo, Aurelio Ariza and Marcos Isamat

RNA 2007 13: 2312-2323; originally published online Oct 16, 2007; Access the most recent version at doi:10.1261/rna.732807

References
This article cites 35 articles, 17 of which can be accessed free at: http://www.rnajournal.org/cgi/content/full/13/12/2312#References

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here

Notes

To subscribe to RNA go to: http://www.rnajournal.org/subscriptions/

© 2007 RNA Society
Multisite and bidirectional exonic splicing enhancer in CD44 alternative exon v3

ELENA VELA,1,2,3 JOSEP M. HILARI,2,3 XAVIER ROCA,1,2,4 ANA M. MUÑOZ-MÁRMOL,2 AURELIO ARIZA,2 and MARCOS ISAMAT1

1Fundación Echevarne, Barcelona 08037, Spain
2Department of Pathology, Hospital Universitari Germans Trias i Pujol, Autonomous University of Barcelona, Barcelona 08916, Spain

ABSTRACT

The human CD44 gene encodes multiple isoforms of a transmembrane protein that differ in their extracellular domains as a result of alternative splicing of its variable exons. Expression of CD44 is tightly regulated according to the type and physiological status of a cell, with expression of high molecular weight isoforms by inclusion of variable exons and low molecular weight isoforms containing few or no variable exons. Human CD44 variable exon 3 (v3) can follow a specific alternative splicing route different from that affecting other variable exons. Here we map and functionally describe the splicing enhancer element within CD44 exon v3 which regulates its inclusion in the final mRNA. The v3 splicing enhancer is a multisite bipartite element consisting of a tandem nonamer, the XX motif, and an heptamer, the Y motif, located centrally in the exon. Each of the three sites of this multisite enhancer partially retains its splicing enhancing capacity independently from each other in CD44 and shows full enhancing function in gene contexts different from CD44. We further demonstrate that these motifs act cooperatively as at least two motifs are needed to maintain exon inclusion. Their action is differential with respect to the splice-site target abutting v3. The first X motif acts on the 3′ splice site, the second X motif acts on both splice sites (as a bidirectional exonic splicing enhancer), and the Y motif acts on the 5′ splice site. We also show that the multisite v3 splicing enhancer is functional irrespective of flanking intron length and spatial organization within v3.

Keywords: CD44 gene; alternative splicing; exonic splicing enhancer; v3 inclusion

INTRODUCTION

Alternative pre-mRNA splicing is a very common event in higher eukaryotes which enables the production of multiple mRNA isoforms from a single gene. Alternative splicing provides a powerful means to expand the proteome diversity with a prevalence of up to 60% of all human genes (Stamm et al. 2005). Several genes encode transcripts that are alternatively spliced to produce up to tens of thousands of different mRNAs (Graveley 2001). The CD44 gene is becoming a model for the study of loci that undergo complex alternative splicing to originate a wide diversity of proteins. The category of complex alternative splicing applies to this central region, which is responsible for the incorporation of the variable domains to shape, predominantly, the extracellular, membrane-proximal stem structure of the protein (Stamenkovic et al. 1989; Screaton et al. 1992; Bajorath 2000; Ponta et al. 2003).

The expression of certain CD44 isoforms has been implicated in important physiological and pathological...
Alternative splicing of CD44 v3 is independent from flanking intron length or sequence

Previous reports have described the capacity of CD44 exon v3 to be included in or excluded from the final CD44 mRNA in a manner specific for v3 and independent from the general 3’-to-5’ directional variable exon inclusion trend (Roca et al. 1998; Zhu et al. 2003). In order to map the splicing regulatory elements within v3, we performed a complete dissection of this exon by exon trap and mutagenesis experiments (Fig. 1A).

Human CD44 v3 exon spans 126 base pairs (bp) and is flanked by introns 6 and 7, which are 1246 and 2835 bp in length, respectively. Experimental constrains prompted the use of a shortened version of such introns in exon trap experiments concerning v3 (Fig. 1B). The influence of intron length on splicing efficiency has been studied and reported to be inversely correlated to exon inclusion in some instances during CD44 splicing (Bell et al. 1998). Confirmation that intron length does not affect v3 inclusion was provided by means of transient transfection of a shortened version of both flanking introns whenever the authentic splice sites are left intact (Fig. 1B, lane 3). When the 5’ splice site (5’ss) was mutated from GTAAG to GTAGG, v3 was excluded from the final mRNA (Fig. 1B, lane 4). Full inclusion of exon v3 in the shorter intron construct implies that introns 6 and 7 contain no elements necessary for v3 splicing and that alternative splicing control must therefore reside within exon v3 itself (Fig. 1B, cf. lanes 2 and 3).

CD44 v3 inclusion is regulated by an internal splicing enhancer

In an attempt to locate the elements regulating the alternative splicing of v3, four regions (A, B, C, and D) spanning the whole of v3 were arbitrarily defined and deleted independently to construct exon trap deletion mutants Del A (nucleotides 4–25), Del B (nucleotides 21–61), Del C (nucleotides 57–93), and Del D (nucleotides 94–121) as shown in Figure 2A. Subsequent transient transfection of these constructs into the MCF-7 cell line and RT-PCR analysis of the spliced products (Fig. 2B) revealed a 30% drop in inclusion efficiency for the Del B construct (Fig. 2B, lane 3) and a minor decrease in inclusion in the case of Del D (Fig. 2B, lane 5). These findings indicate the presence of exonic splicing enhancers in these regions. Inclusion of the previously described –49) pseudo-exon appeared in the Del B construct. This pseudo-exon is generated from a cryptic 3’ splice site (3’ss) in intron 6, 49 nucleotides (nt) upstream from the classical 3’ss of exon v3 (Vela et al. 2006) as seen in Figure 1A.

Region B was further divided into four subregions, namely, B1 (nucleotides 22–33), B2 (nucleotides 32–42), B3 (nucleotides 41–51), and B4 (nucleotides 50–61), and four independent deletion mutants spanning each of these subregions were constructed (Fig. 2C). Exon trap analysis of these constructs (Fig. 2D) did not show the expected decrease in v3 inclusion when compared to the original wt v3 construct. A detailed analysis of the sequences deleted in small deletion mutants revealed that a 9-bp sequence motif (AAATGAAGA), named X, was present in all four constructs, and that whenever this sequence was interrupted during the generation of the deletion, the motif was reconstructed after rejoining the sequences downstream (see Fig. 2C). In order to test the hypothesis that this 9-nt motif could indeed be responsible for alternative splicing control of v3, the sequence was inserted into the Del B construct (Fig. 3A) and analyzed by transient transfection into MCF-7 and RT-PCR (Fig. 3B). Insertion of the 9-nt oligo into Del B (Fig. 3B, construct Del B Ins Xs) restored the phenotype to wt v3 inclusion levels (Fig. 3B, lane 2), while insertion of the antisense version of this sequence into the same site (Fig. 3B, construct Del B Ins Xas) did not change v3 inclusion levels (Fig. 3B, cf. lanes 1 and 3). This
finding reveals that the X nonamer motif is playing a role in v3 inclusion in these constructs.

Positive exonic cis-acting elements, known as exonic splicing enhancers (ESEs) are often found in a purine-rich context and may be functionally associated to the latter (Cartegni et al. 2002). The v3 minigene contains a contiguous genomic fragment of CD44 v3 and flanking introns was cloned in the polylinker of the exon trap vector (pET). Promoter (RSV LTR), insuline (Ins 1 and Ins 2) exons, CD44 intron 6 containing the 49-bp pseudo-exon (gray line), v3 (black box), intron 7, and possible splicing routes (thin lines) are indicated. (Bottom) CD44 exon v3 sequence: upper case, exon; lower case, intron boundaries; lower-case italics, 49-bp pseudo-exon. (B, top) Intron length in v3 minigenes. v3 minigene contains full-length introns 6 and 7. In the v3 short minigene, intron 6 was shortened from 1246 to 233 bp and intron 7 from 2835 to 160 bp. (Bottom) Transient transfection of the minigenes in the MCF-7 cell line: lane 1, vector without insert; lane 2, v3 with full-length flanking introns 6 and 7; lane 3, v3 with shorter versions of introns 6 and 7; lane 4, v3 with full-length introns containing a point mutation in the 5’splice site (GTAAG to GTAGG).

**CD44 v3 splicing enhancer motifs are functional in a different gene context**

Region B contains two copies of the X motif as a tandem repeat, and a 7-nt version of this motif (ATGAAGA), named Y, is present as a single copy in the D region (Fig. 2A) that also affects, albeit to a lesser extent, the inclusion of v3 (Fig. 2B, lane 5). Functional proof of the inherent splicing enhancing capacity of the X nonamer and Y heptamer was obtained by their independent insertion in the context of an heterologous exon in the BRCA1 mini-gene system. The BRCA1 minigenes consist of exons 17, 18, and 19 of the human BRCA1 gene, where the introns flanking exon 18 have been shortened (Liu et al. 2001). In the wild-type version (BRCA1 wt), exon 18 is efficiently included. However, a G/T transversion at position 6 of exon 18 disrupts a critical ESE, leading to exon 18 skipping.

---

**FIGURE 1.** CD44. (A, top) Genomic structure of the human CD44 gene (gray boxes, constitutive exons; white boxes, alternative exons; black box, variable exon v3; black line, introns). (Middle) Partial structure of v3 minigene. The reporter v3 minigene containing a contiguous genomic fragment of CD44 v3 and flanking introns was cloned in the polylinker of the exon trap vector (pET). Promoter (RSV LTR), insuline (Ins 1 and Ins 2) exons, CD44 intron 6 containing the 49-bp pseudo-exon (gray line), v3 (black box), intron 7, and possible splicing routes (thin lines) are indicated. (Bottom) CD44 exon v3 sequence: upper case, exon; lower case, intron boundaries; lower-case italics, 49-bp pseudo-exon. (B, top) Intron length in v3 minigenes. v3 minigene contains full-length introns 6 and 7. In the v3 short minigene, intron 6 was shortened from 1246 to 233 bp and intron 7 from 2835 to 160 bp. (Bottom) Transient transfection of the minigenes in the MCF-7 cell line: lane 1, vector without insert; lane 2, v3 with full-length flanking introns 6 and 7; lane 3, v3 with shorter versions of introns 6 and 7; lane 4, v3 with full-length introns containing a point mutation in the 5’splice site (GTAAG to GTAGG).
(BRCA1 NL). Both the X nonamer and the Y heptamer were independently recreated in the BRCA1 NL minigene near the 3' end of exon 18, where the sequence AAATGG GTA was mutagenized to AAATGaagA (BRCA1 NL X) and to tggATGaagA (BRCA1 NL Y) (Fig. 5A). RT-PCR analysis (Fig. 5B) showed that both elements rescue exon 18 inclusion (lanes 3 and 4), and demonstrated that both the X nonamer and the Y heptamer are bona fide ESEs that could function in an exonic context different from the CD44 gene. From that moment the AAATGGGTA sequence was named X mutant (Xm) and ATGGGTA sequence was named Y mutant (Ym), and they were used as negative-splicing control ESEs in subsequent experiments.

**CD44 v3 splicing enhancer operates as multisite and bidirectional element**

If an ESE is important for exon v3 inclusion, then the X nonamer and Y heptamer sequences within v3 should be required for maximal activity. In order to ascertain the level of cooperativity between the tandem repeat of X and Y, each of the three motifs within v3 was mutated in the context of the v3 minigene. The X nonamer was mutated from AAATGAAGA to AAATGggtA and the Y heptamer from ATGAAGA to ATGggtA to generate Xm and Ym (Fig. 6A) as suggested by their lack of splicing activity in the BRCA1 NL minigene experiments. These mutants were transfected into MCF-7 cells and exon inclusion analysis (Fig. 6B) showed that mutation of one of the three motifs present lowered exon inclusion levels to the same extent (Fig. 6B, lanes 2–4). Subsequent mutation of any two of the motifs caused exon skipping (Fig. 6B, lanes 5–7), as did the triple mutant (Fig. 6B, lane 8). These results suggest that the splicing enhancing strength is comparable in all three motifs, and a minimum of two of them is required to enable splicing, which indicates some degree of cooperativity between these sites.

To determine the level of cooperativity and dependence on relative positioning of the X and Y motifs, a series of mutants containing a combination of only two motifs from X, Y, Xm, and Ym was used in exon inclusion analysis. These constructs were based on the Del B3 minigene and included XX, YY, XY, XmX, XmY, YXm, and YmXm (Fig. 6C). These results (Fig. 6D) demonstrate that at least two motifs are needed to maintain exon inclusion during splicing (Fig. 6D, lanes 1–4), and mutation of one of them, independently of position, leads to a dramatic increase in exon skipping (Fig. 6D, lanes 5–8). Mutation of both sites eliminates detectable v3 inclusion (Fig. 6D, lane 9). Furthermore, this result suggests that whenever there is only one functional motif, X and Y have slightly different properties, as X retains some function in both positions and Y only exerts some promotion on splicing in its original position relative to X (Fig. 6D, cf. lanes 5 and 8). This position-dependent effect is commonplace during splicing regulation (Goren et al. 2006). In addition, cooperativity between motifs is gene context–dependent as either motif by itself promotes efficient splicing in the BRCA1 experiments (Fig. 5B) but not in the CD44 v3 context. The differences on splicing efficiency may be explained by the synergy of any single motif with other motifs already present in the BRCA1 exon 18 or to exon context effects,
such as differences in relative position of the motif within the exon, exon size, or identity of the splice sites. It is not possible, on the basis of the data provided herein, to ascribe the cause of such differences to one, several, or all of the effects listed above.

The first step in the splicing of an exon involves the recognition of both 5’ss and 3’ss (Reed and Palandjian 1997). Although the recognition of one splice site often facilitates the use of the other through a process known as exon definition, the two splice sites can also be recognized independently and are both potential sites of regulation (Reed 1996; Graveley 2000). To investigate further which of the splice sites of exon v3 was regulated by the X and Y motifs, single-intron minigene constructs were made in which exon v3 was truncated either immediately upstream of the 5’ss or immediately downstream of the 3’ss, yielding minigene constructs v3-i6 and v3-i7, respectively.

In v3-i6 minigenes, the 5’ss and intron 7 were deleted in order to characterize the regulation of the 3’ss by v3 ESE (Fig. 7A). In this context, all combinations of XXY mutants were created and splicing efficiency of intron 6, via v3 3’ss, was analyzed by RT-PCR (Fig. 7B). In these constructs, two cryptic 3’ss are functional: the previously described (−49) 3’ss (Vela et al. 2006) and the (+27) 3’ss. Analysis of splicing efficiency considered the use of the authentic v3 3’ss versus the cryptic 3’ss. Mutation of one of the motifs maintained splicing in all constructs (Fig. 7B, lanes 2–4), but mutation of two motifs demonstrated that splicing via
3'ss was only maintained in XmXYm-i6 and XXmYm-i6 (Fig. 7B, lanes 6 and 7) and dropped to 30% in XmXmY-i6 and triple mutant constructs (Fig. 7B, lanes 5 and 8). These results demonstrate that only the X motifs maintained splicing via the 3'ss and the Y motif failed in the recognition of the 3'ss, so that intron 6 is inefficiently spliced.

In v3-i7 minigenes, the 3'ss and intron 6 were deleted in order to analyze the regulation of the 5'ss by v3 ESEs (Fig. 7C). All combinations of XXY mutants were generated and splicing efficiency of intron 7, via v3 5'ss, was analyzed by RT-PCR (Fig. 7D). In these constructs a cryptic 5'ss was also functional: (−72) 5'ss. Analysis of splicing efficiency considered the use of the authentic v3 5'ss versus the cryptic 5'ss. In constructs with two active motifs, splicing via 5'ss was maintained in XmXY-i7 and XXmYm-i7 (Fig. 7D, lanes 2 and 4) and was reduced to 60% in XxmY-i7 (Fig. 7D, lane 3). Mutation of two motifs caused a drop to 40% of splicing efficiency via 5'ss in XmXmY-i7 and XmXYm-i7 (Fig. 7D, lanes 5 and 6) and complete exon exclusion in XXmYm-i7 and triple-mutant constructs (Fig. 7D, lanes 7 and 8). These results demonstrate that the second X and Y motifs maintained splicing via 5'ss, but the first X does not participate in the splice-site recognition and thus intron 7 cannot be spliced.

Taken together, the combined results from the one-intron experiments suggest that there is a cooperative action between the motifs for the recognition of the splice sites. This cooperativity contributes to the splicing activity when two active motifs are present in the construct. However, when only one motif remains active, it manifests directional action in splice-site recognition. In this way, the first X nonamer acts on the recognition of the 3'ss, the second X nonamer acts in the recognition of both splice sites (bidirectional action), and the Y heptamer acts on the recognition of the 5'ss.

Background splicing via 3'ss in v3-i6 minigenes suggests that other factor(s) are influencing 3'ss usage, in this construct specifically, as compared with v3-i7 experiments, or other results presented herein where no splicing is achieved in the absence of ESE. Likewise, v3-i6 splicing analyses do not reach expected 100% efficiency as noted in the rest of our results. This may respond to the presence of the active cryptic (−49) 3'ss, which competes in this test. Functional demonstration of the use of this cryptic splice site under physiological conditions is available elsewhere (Vela et al. 2006).

The consensus sequence of mammalian 5'ss is CAG/UUAAGU (Sheth et al. 2006). The 5'ss of CD44 exon v3 differs from the consensus at positions −1, −2, and +6. The consensus sequence of mammalian 3'ss is UUUUUUC/CCCUNCA/G. The 3'ss of CD44 exon v3 differs from the consensus at positions −5, −8, −9, and −10. To test the influence of 5'ss and 3'ss strength on v3 splicing, we constructed minigenes in which the splice sites of v3 were mutated to match the consensus sequence of mammalian splice sites (Fig. 8A). These constructs, with active enhancer (XXX) or triple mutant enhancer (XmXmYm), were transiently transfected in MCF-7 and analyzed by RT-PCR (Fig. 8B). While the splicing enhancer was active, consensus splice sites did not increase the splicing efficiency, which was already maximal (Fig. 8B, lanes 2–4). In the case of mutant ESE, the consensus splice sites marginally increased v3 inclusion, and only when both splice sites were improved to match the consensus did exon inclusion reach significant levels (Fig. 8B, lanes 6–8). These observations indicate that changing the v3 splice sites to perfectly match the consensus is not sufficient for constitutive inclusion of v3 and that the presence of the ESE described herein is required to achieve efficient exon inclusion.

All minigene constructs described in this work were also tested in the monkey kidney COS-7 cell line, with equivalent results (data not shown). This implies that CD44 exon v3 regulation by the XXY ESE is not restricted to human cells but that it is also operational in other animal models and tissue types.
DISCUSSION

We had previously described that splicing of human CD44 exon v3 is regulated independently from other variable exons in a breast tissue model. Here we exhaustively dissected the sequence elements that dictate the splicing behavior of exon v3, by minigene transfection and RT-PCR analysis. This strategy has allowed us to define exon sequences whose mutation decreases the abundance of spliced transcripts containing exon v3, in particular a tandem nonamer (AAATGAAGA, named the X motif) and a heptamer (ATGAAGA, named the Y motif). These motifs act cooperatively as their presence together within v3 regulates exon inclusion. Their action is differential with respect to the splice-site target abutting v3. The first X motif acts on the 3'ss, the second X motif acts on both splice sites (as a bidirectional ESE), and the Y motif acts on the 5'ss. Both of these sequence motifs retain their splicing enhancing activity in a heterologous gene context in transient transfection experiments, defining them as bona fide ESEs. We also show that the multisite v3 splicing enhancer is functional, irrespective of flanking intron length and its spatial organization within v3. As CD44 v3-cassette inclusion is detected in both normal breast tissue and in breast...

FIGURE 6. Point mutants. (A) XX and Y point mutants were generated by PCR-directed mutagenesis from the v3 minigene. Active X and Y motifs are indicated in bold. Mutated positions are indicated in lower case. (B) RT-PCR amplification of splicing products after transient transfection of XX and Y point mutants in the MCF-7 cell line: lane 1, XXY; lane 2, XmXY; lane 3, XXmY; lane 4, XXym; lane 5, XXym; lane 6, XmXYm; lane 7, XmXYm; lane 8, XmXym. (C) X and Y point mutants were generated by PCR-directed mutagenesis from the Del B3 construct. Active X and Y motifs are indicated in bold. Mutated positions are indicated in lower case. (D) RT-PCR amplification of splicing products after transient transfection of X and Y point mutants in the MCF-7 cell line: lane 1, XY; lane 2, XX; lane 3, YY; lane 4, YX; lane 5, XmX; lane 6, XmY; lane 7, XYm; lane 8, YYm; lane 9, XmYm. Extra bands (*) correspond to the (−49) pseudo-exon (Vela et al. 2006).
carcinoma (Roca et al. 1998), its regulation through the ESE described herein may be biologically relevant in both normal and transformed cells.

Expression of minigene pre-mRNAs by transient transfection provides a rapid assay for loss-of-function/gain-of-function analyses and for cis- and trans-acting factors that affect splicing regulation (Cooper 2005). A different and complementary approach to identify ESEs in vitro involves biocomputing methods based on the over-representation of these sequences in exons. We have evaluated the level of theoretical confidence for the elements functionally identified in exon v3 using available biocomputing resources. Fairbrother’s relative enhancer and silencer classification by unanimous enrichment (RESCUE) biocomputing method identified 238 hexamer sequences that occurred more frequently in exons with weak splice sites (Fairbrother et al. 2002). Submission of wild-type CD44 v3 sequences to this Web server (Fairbrother et al. 2004) predicted three ESE

**FIGURE 7.** Single-intron minigenes. (A) v3-i6 minigenes were generated by PCR subcloning from v3 constructs, deleting the v3 5’ss and intron 7. Exon sequences are indicated in upper case. Intron sequences are indicated in lower case. 3’ss are indicated in bold. Positions of cryptic 3’ss are numbered according to authentic 3’ss. (B) RT-PCR amplification of splicing products after transient transfection of v3-i6 constructs in the MCF-7 cell line. XX and Y mutants were generated by PCR-directed mutagenesis. RT-PCR products generated by the use of the different 3’ss were quantified as net intensity of each band. Splicing efficiency of intron 6 via v3 3’ss, was defined as (use of authentic v3 3’ss)/(use of authentic v3 3’ss + (-49) 3’ss + (+27) 3’ss): lane 1, XXY-i6; lane 2, XmXY-i6; lane 3, XXYm-i6; lane 4, XXXm-i6; lane 5, XnmYm-i6; lane 6, XmXYm-i6; lane 7, XXmYm-i6; lane 8, XnmXmYm-i6. (C) v3-i7 minigenes were generated by PCR subcloning from v3 constructs, deleting intron 6 and the v3 3’ss. Exon sequences are indicated in upper case. Intron sequences are indicated in lower case. 5’ss are indicated in bold. Position of the cryptic 5’ss is numbered according to authentic 5’ss. (D) RT-PCR amplification of splicing products after transient transfection of v3-i7 constructs in the MCF-7 cell line. XX and Y mutants were generated by PCR-directed mutagenesis. RT-PCR products generated by the use of different 5’ss were quantified as net intensity of each band. Splicing efficiency of intron 7 via 5’ss, was defined as (use of authentic v3 5’ss)/(use of authentic v3 5’ss + (-72) 5’ss): lane 1, XXY-i7; lane 2, XmXY-i7; lane 3, XXYmY-i7; lane 4, XXXmY-i7; lane 5, XnmYmY-i7; lane 6, XmXYm-i7; lane 7, XXmYm-i7; lane 8, XnmXmYm-i7.
hexamer matches in the X motif (AATGAA, ATGAAG and TGAAGA) and two ESE hexamers in the Y motif (ATGAAG and TGAAGA). Consistent with this prediction, submission of mutated v3 sequences (XmXmYm) provided no matches for hexamer enhancers. RNA octamers described by Zhang and Chasin (2004) are found more frequently in exons than in pseudo-exons or intronless genes. Octamers with a higher prevalence in exons are designated as putative exonic splicing enhancers (PESEs), and those with lower prevalence in exons are designated as putative exonic splicing silencers (PESSs). Analysis of CD44 exon v3 sequences using this algorithm (Zhang et al. 2005) reveals several overlapping PESEs in wild-type v3 sequences—five in the XX region (ATGAAGA, TGAAGAA, GAA GAAAA, ATGAAGAT, and GAAGATGA) and two in the Y motif (GATGAAGA and ATGAAGAT)—but no PESSs. Interestingly, when mutated v3 sequences are tested (XmXmYm), the PESEs theoretically identified in wild-type v3 become PESSs. This observation supports the presence of the two motifs which we have functionally characterized as parts of a multisite ESE for v3 in this report. Furthermore, the motifs described during this work are represented in the collection of 285 exonic splicing regulatory sequences (ESRs) recently reported by the Ast group (Goren et al. 2006). The XX tandem nonamer contains several hexamers predicted by Goren et al. (2006) as ESR: AATGAA, GAAGAA, AAGAAA, and GAAAAAT. In addition, the Y heptamer overlaps with the GATGAAGA and GAA GAT hexamers. Some of the hexamers contained in that collection have been validated in vivo to show enhancer or silencer capacity depending on the ESR environment (Goren et al. 2006). Our assays of the effect on splicing of the XX and Y motifs have shown only enhancing capacity. However, whether they can also function as silencers in other positions, in the same context or in different gene contexts, has not been demonstrated.

The homologs of human CD44 exon v3 in other organisms maintain a high degree of evolutionary conservation. The enhancer motifs X and Y described herein are conserved in other primates like Macaca mulatta (GenBank NC_007871) and Pan troglodytes (GenBank NC_006478). In other mammals like Mus musculus (GenBank NT_039207), Rattus norvegicus (GenBank NM_012924), and Canis familiaris (GenBank L28931), the first X motif differs in one unique nucleotide position and the second X motif and Y motif remain 100% conserved. Based on its behavior in MCF-7 and COS-7 cell lines, where two of the motifs are enough to address v3 inclusion, we would expect v3 in these organisms to promote its inclusion via the same ESE. The presence of two contiguous identical nonamers (XX) in CD44 exon v3 resembles other known multisite ESEs. The doublesex repeat element (dsexRE) from Drosophila contains six nearly identical copies of a 13-nt repeat sequence, each of which is recognized by regulatory proteins. Each repeat constitutes an independent and equivalent binding site for Tra and Tra2, and cooperative binding occurs within but not between individual repeats (Hertel and Maniatis 1998). Multiple repeats of the dsexRE increase splicing efficiency by increasing the probability of splice-site activation rather than by formation of multiple
interactions between the enhancer complex and components bound to the weak 3’ss of the dsx pre-mRNA. Likewise, the fruitless repeat element (fruRE) from Drosophila consists of three 13-nt repeat elements. In this case, only one repeat element within fruRE is necessary and sufficient to activate the female-specific splice site (Lam et al. 2003). fruRE and dsxRE function similarly in activating 5’ss and 3’ss, respectively, and these ESE complexes are capable of recruiting targets that contain multiple spliceosomal components required for the initial recognition of 5’ss and 3’ss. The precise mechanistic nature of the CD44 exon v3 ESE awaits further characterization, but an approximation to its mode of operation is provided by its preference in targeting specific splice sites.

Most ESEs have been shown to promote recognition of either the upstream 3’ss or the downstream 5’ss. Only a few ESEs have been shown to simultaneously promote the recognition of both upstream and downstream splice sites, thus working as a bidirectional splicing enhancer. Bourgeois et al. provided the first characterization of a purine-rich bidirectional splicing enhancer (Pu1 and Pu2), in the adenovirus EIA pre-mRNA (Bourgeois et al. 1999). Pu1 is involved in splicing activation through the 5’ss, whereas both Pu1 and Pu2 are required for an optimal activation of the 3’ss. In addition to the high level of sequence identity between Pu2 and the v3 X motif, this model may set a precedent for the v3 ESE function, where the first X is involved in 3’ss recognition and the second in both 5’ss and 3’ss activation.

In order to identify putative ESEs responsive to human SR proteins SF2/ASF, SC35, SRp40, and SRp55, CD44 v3 sequences were analyzed with the ESEfinder web server (Cartegni et al. 2003). This approach failed to match SR-responsive elements in the XX and Y regions, suggesting that the effect of these motifs is not mediated by any of these four proteins. Submission of CD44 v3 sequences to ESRsearch web tool (Goren et al. 2006) had equivalent results: none of the motifs contained in v3 ESE matched any known trans-factors’ binding site. Our attempts to identify specific trans-factors acting on the isolated motifs of the v3 ESE, by RNA-affinity precipitation, have systematically failed to reveal specific RNA–protein binding (data not shown). This failure calls for a change in the trans-factor identification strategy that should consider the novel v3 composite ESE as a whole functional splicing enhancer unit, in line with our functional cis-acting characterization that defines a complex operational multisite unit during the splicing process.

The sequence and functional information generated in the course of this work sets the basis for trans-factor identification and contributes to the construction of a full mechanistic model for CD44 alternative splicing. We report a novel structure and organization within exon v3 in the form of a functional multisite bipartite bidirectional ESE.

### MATERIALS AND METHODS

#### Cell culture and transfections

In order to analyze the regulation of CD44 v3 splicing in a breast tissue model, the MCF-7 cell line was selected as an in vitro model of human breast carcinoma cells due to its high efficiency of transfection, in our experience, during transient transfection assays as compared to other cell lines. The COS-7 cell line (African green monkey kidney) was selected as a model unrelated to human breast tissue.

Cell lines were purchased from the American Type Culture Collection (ATCC). MCF-7 cells were cultured in Minimum Essential Medium-Eagle supplemented with 10% fetal calf serum, 2 mM l-glutamine, 0.1 mM nonessential amino acids, 1 mM sodium pyruvate, 0.01 mg/mL bovine insulin, 100 units/mL penicillin and gentamicin. COS-7 cells were grown in Dulbecco’s modified Eagle medium supplemented with 10% fetal calf serum, 2 mM l-glutamine, 100 units/mL penicillin and gentamicin. Cell lines were tested periodically for the absence of mycoplasma contamination following a standard PCR method. MCF-7 and COS-7 cells were transiently transfected using 200 ng of DNA and the LipofectAMINE PLUS Reagent (Gibco) according to the manufacturer’s instructions. Cells were harvested 24 h after transfection to extract total RNA.

#### CD44 v3 minigene constructs

A fragment of the human genomic CD44 sequence (GeneBank NT_009237), containing exon v3 and flanking introns 6 and 7, was amplified by PCR using two sets of primers containing XbaI restriction sites, I6FX (5’-GGGTATGTTTTTCTAGACGCTTT GATAAT-3’) with I7R5bX (5’-AATGGAATCATGATGACGCG GTTGTT-3’) and I6F3X (5’-CTTTTCAGCTTTTCTAGAGGAGATA TTCAGTT-3’) with I7RX (5’-ATATTTTTCCCAACTCTAGA AACATTCTA-3’). The first set, I6FX and I7RX, amplifies the full length of introns 6 and 7. The second set, I6FX3 and I7RX1, amplifies from nucleotide 1012 in intron 6 to nucleotide 160 in intron 7.

The genomic inserts were cloned into a pUC18 vector with SureClone Ligation Kit (Pharmacia) and subcloned into the Xbal site within the multiple cloning site of the mammalian expression exon trap (pET) vector (MobITec), resulting in the minigene constructs pET v3 and pET v3 short. Large deletion mutants were generated by combining PCR fragments from the above constructs, giving rise to deletions in the v3 exon sequence. Del A mutant was generated with primers F5’-v3Acc (5’-CCATGGTCACGCCGTCGAGCCAAATGGA-3’) and R5’-v3Acc (5’-AAGGATCCTGTGTTGTTGACCTTCAT TA-3’), deleting from nucleotide 4 to nucleotide 25. Del B mutant was generated with primers F5’-v3Xho (5’-AGATGATCCGAGACA GACACCCGATTTTCC-3’) and R5’-v3Xho (5’-GCCTCTCGA GATGTATTTGGAGCGATGC-3’), deleting from nucleotide 21 to nucleotide 61. Del C mutant was generated with primers F3’-v3Nsi (5’-GATCAGATCTGAGAGAAGAGATTTT-3’) and R3’-v3Nsi (5’-GTCTCATGACTGATCCTTCTCTCTTTG-3’), deleting from nucleotide 57 to nucleotide 93. Del D mutant was generated with primers F3’-v3Sph (5’-TTATCG CATGCCACGAGATAATACGTT-3’) and R3’-v3Sph (5’-AT CATGATGCCCTGATCAGAGGAAACTGAG-3’), deleting from...
reaction was performed with the First-Strand Synthesis Kit Ultraspec RNA solution (Biotecx Laboratories). The first-strand RT-PCR was performed from 2

RT-PCR analysis

RT-PCR was performed from 2

2322 RNA, Vol. 13, No. 12

REFERENCES


