Multisite and bidirectional exonic splicing enhancer in *CD44* alternative exon *v3*

Elena Vela, Josep M. Hilari, Xavier Roca, et al.

References	This article cites 35 articles, 17 of which can be accessed free at: http://rnajournal.cshlp.org/content/13/12/2312.full.html#ref-list-1
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**

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 ISAMAT[1]

[1]Fundación Echevarne, Barcelona 08037, Spain
[2]Department 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 has been applied to genes that have multiple variable exons that appear to be included in or excluded from the processed mRNA in a complex manner to produce multiple mRNA isoforms from the same gene (Breitbart et al. 1987).

The CD44 family of cell-surface glycoproteins mediates the response of cells to their cellular microenvironment by regulating growth, survival, differentiation, and motility. All CD44 proteins are encoded by a single and highly conserved gene containing 20 exons, 12 of which undergo alternative splicing. *CD44* exons are distributed in the gene in four regions: two constant regions consisting of exons 1–5 and exons 15–17, which are subject to general constitutive splicing; a region composed of exons 18 and 19, which show an alternate use of a short or long cytoplasmic tail, respectively; and a central region that spans exons 6a–14, also known as variable exons *v1–v10*. 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

---

conditions, including cell adhesion and migration, hematopoiesis, lymphocyte activation, assembly of extracellular matrix, and metastatic spread of tumors (Herrlich et al. 1993; Gunthert et al. 1995; Ruiz et al. 1995; Lesley et al. 1997). We have previously used human breast ductal carcinoma to show that the CD44 isoform expression pattern of the breast carcinoma and its lymph-node metastases is equivalent, presenting a stepwise inclusion of variable exons (Roca et al. 1998). Variable exon inclusion in this model starts with the 3′-most variable exons (i.e., *v10* and *v9*) in low molecular weight isoforms and gradually moves toward 5′ variable exons (i.e., *v2*) with a subsequent increase in isoform molecular weight. In contrast, normal cells express low molecular weight CD44 isoforms. This alternative splicing pathway, that reflects a 3′-to-5′ directional and stepwise inclusion of variable exons, is distinguishable from an alternative splicing pathway that involves exclusively variable exon *v3*, observable in both normal and carcinoma breast tissue. Exon *v3* is subject to the 3′-to-5′ exon inclusion preference but is also present in a low molecular weight isoform corresponding to CD44 *v3* alone. These pathways are not mutually exclusive and can be easily integrated into a single hypothetical model of *CD44* complex alternative splicing (Roca et al. 1998).

Here we report the role of *CD44 v3* exon sequences in *v3* splicing using deletion mutants and extensive site-directed mutagenesis. We found a *cis*-acting splicing enhancer of novel structure and organization within exon *v3*, and we provide proof of function and mode of operation both in the context of *CD44* and when located in an exogenous gene. The significance of these findings is discussed in light of the operational nature of multisite and bidirectional splicing enhancers described for other gene loci.

RESULTS

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). Confir-

mation 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
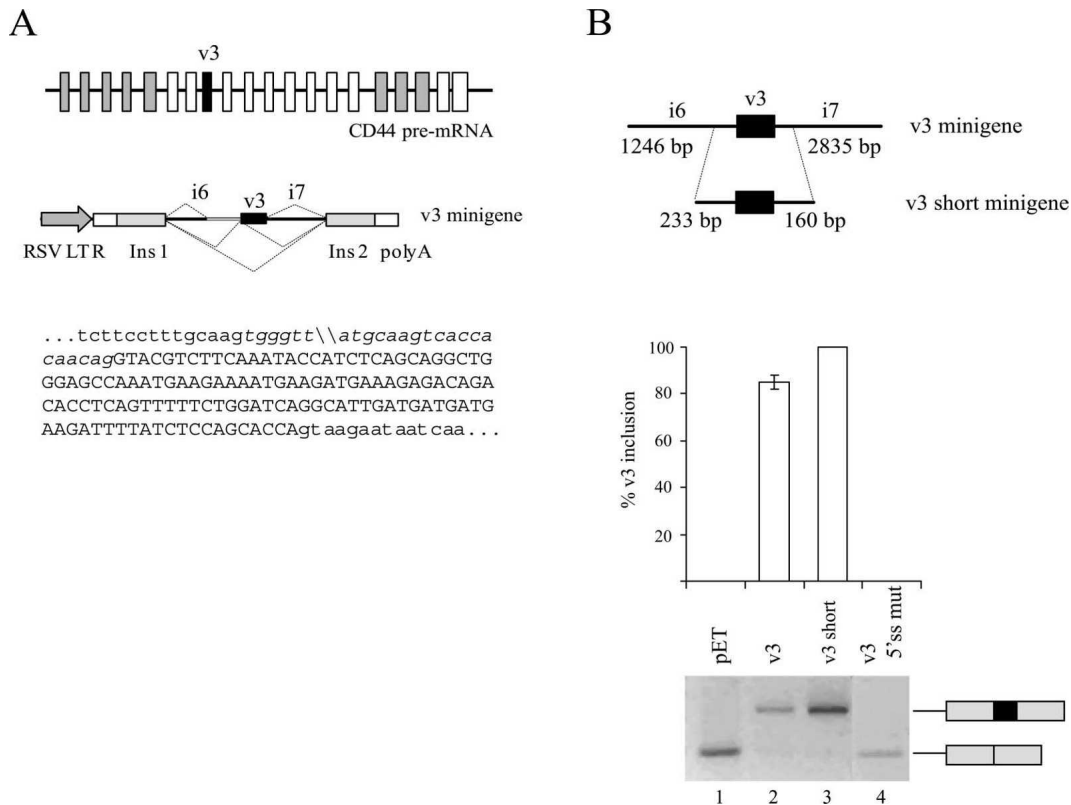
FIGURE 1. *CD44*. (*A, top*) Genomic structure of the human *CD44* gene (gray boxes, constitutive exons; white boxes, alternative exons; black box, variable exon *v3*; black line, introns). (*Middle*) Partial structure of *v3* minigene. The reporter *v3* minigene containing a contiguous genomic fragment 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).

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* sequence contains seven GAR motifs which may contribute during splicing regulation as reported in other exons (Lavigueur et al. 1993; Watakabe et al. 1993; Xu et al. 1993; Bourgeois et al. 1999). In order to test their level of contribution in the context of the X nonamer described above, the Del B3 construct was selected for transfection experiments as it contains only one X nonamer with a single internal and three flanking GAR motifs. These GAR motifs have been mutated from GAA to cAA and from GAG to cAG (Fig. 4A) and tested by transient transfection. These experiments (Fig. 4B) show that flanking GAR motifs have no effect on splicing (Fig. 4B, lanes 2–5), and mutation of the internal motif carries a marginal drop in exon inclusion (Fig. 4B, lanes 6 and 7), suggesting that *v3* inclusion requires additional *cis*-acting

elements, contrary to the situation reported for other ESEs (Tanaka et al. 1994; Ramchatesingh et al. 1995).

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* minigene 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
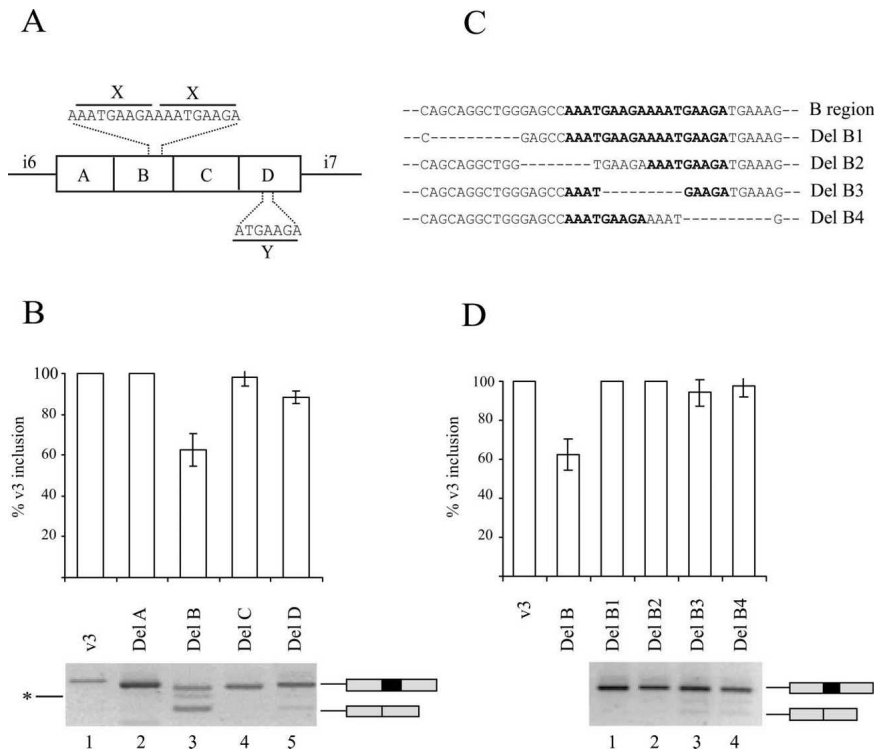
A

C

```
--CAGCAGGCTGGGAGCCAAATGAAGAAAATGAAGATGAAAG--   B region
--C----------GAGCCAAATGAAGAAAATGAAGATGAAAG--   Del B1
--CAGCAGGCTGG--------TGAAGAAAATGAAGATGAAAG--   Del B2
--CAGCAGGCTGGGAGCCAAAT---------GAAGATGAAAG--   Del B3
--CAGCAGGCTGGGAGCCAAATGAAGAAAAT----------G--   Del B4
```

B

D

FIGURE 2. Deletion mutants. (*A*) Large deletion mutants. *CD44* exon *v3* was arbitrarily divided into four regions: A, nucleotides 4–25; B, nucleotides 21–61; C, nucleotides 57–93; D, nucleotides 94–121. Deletions of each region were created in the context of pET *v3* with short introns by PCR and subcloning procedures. Locations of *v3* splicing enhancer motifs (X and Y) are indicated. (*B*) RT-PCR amplification of splicing products after transient transfection of large deletion mutants in the MCF-7 cell line: lane *1*, *v3*; lane *2*, Del A; lane *3*, Del B; lane *4*, Del C; lane *5*, Del D. Extra band (*) corresponds to (−49) pseudo-exon previously described (Vela et al. 2006). (*C*) Small deletion mutants. *CD44 v3* B region was subdivided into four regions: B1, nucleotides 22–33; B2, nucleotides 32–42; B3, nucleotides 41–51; B4, nucleotides 50–60. Deletions of each region were created in the *v3* construct by PCR and subcloning. In bold, repeated X element is indicated. (*D*) RT-PCR amplification of splicing products after transient transfection of small deletion mutants in the MCF-7 cell line: lane *1*, Del B1; lane *2*, Del B2; lane *3*, Del B3; lane *4*, Del B4.

(*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 AAATG GGTA was mutagenized to AAATGgaagA (*BRCA1* NL X) and to tgATGgaagA (*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, YX, XmX, XmY, XYm, Yym, and XmYm (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,
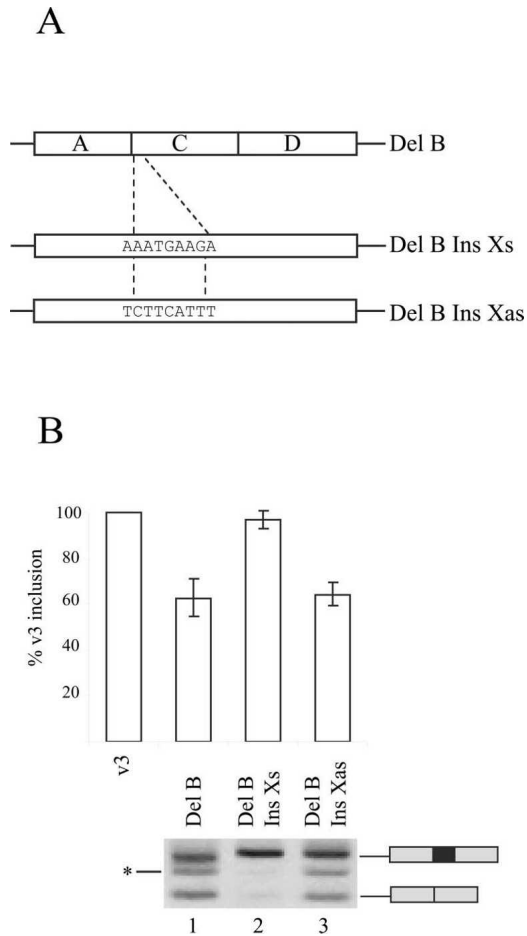
A

B

FIGURE 3. Insertion mutants. (*A*) Functional proof in *CD44* constructs was approached by insertion of double-stranded oligonucleotides into the Del B mutant and RT-PCR tests of their splicing capacity. Del B Ins Xs mutant was generated by insertion of the AAATGAAGA oligonucleotide. Del B Ins Xas mutant was generated by insertion of the TCTTCATTT oligonucleotide. (*B*) RT-PCR amplification of splicing products after transient transfection of insertion mutants in the MCF-7 cell line: lane *1*, Del B; lane *2*, Del B Ins Xs; lane *3*, Del B Ins Xas. Extra bands (*) in lanes *1* and *3* correspond to the (−49) pseudo-exon (Vela et al. 2006).

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
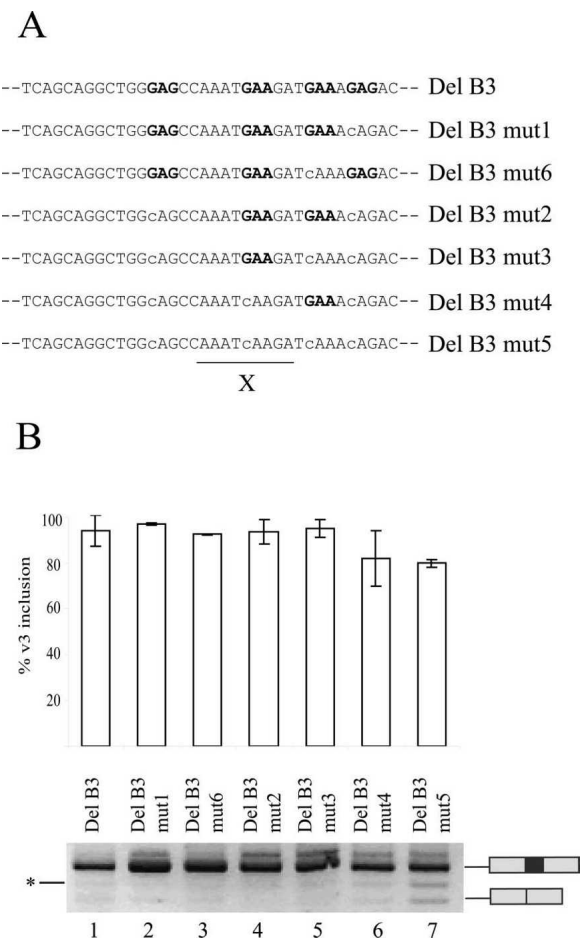
A

B

FIGURE 4. GAR mutants. (*A*) GAR motifs located in the context of X nonamer were mutated by PCR-directed mutagenesis from the Del B3 construct. These motifs have been mutated from GAA to cAA and from GAG to cAG. Active GAR motifs are indicated in bold. Mutated positions are indicated in lower case letters. Sequence corresponding to X nonamer is indicated. (*B*) RT-PCR amplification of splicing products after transient transfection of GAR mutants in the MCF-7 cell line: lane *1*, Del B3 construct; lane *2*, mutant 1; lane *3*, mutant 6; lane *4*, mutant 2; lane *5*, mutant 3; lane *6*, mutant 4; lane *7*, mutant 5. Extra band (*) corresponds to the (−49) pseudo-exon (Vela et al. 2006).

A

ATGCTGAGTTTGTG BRCA1 wt
ATGCT**t**AGTTTGTG BRCA1 NL

Exon 17 — Exon 18 — Exon 19   BRCA1 minigene

AGGAAAATGGGTAGTTA
AGGA**AAATG**aag**A**GTTA X
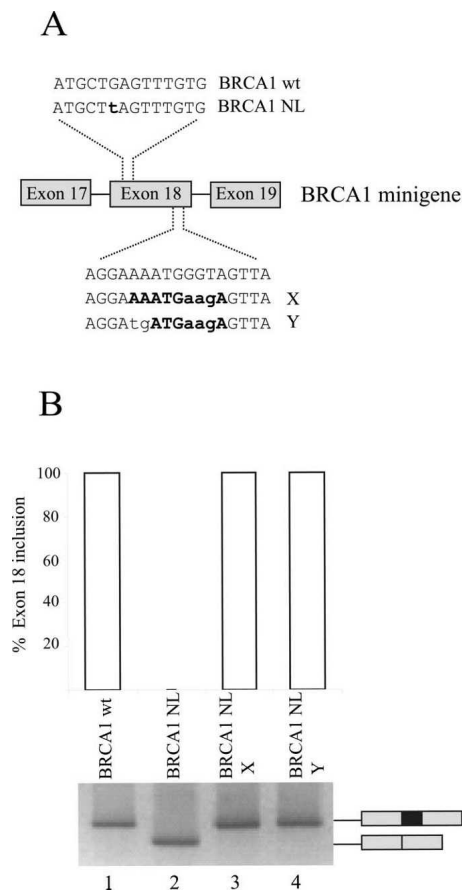AGGAtg**ATG**aag**A**GTTA Y

B

FIGURE 5. Proof-of-function constructs. (*A*) Functional proof in the heterologous context of *BRCA1* minigene system. X and Y motifs were created in *BRCA1* NL minigenes by PCR-directed mutagenesis. (*B*) RT-PCR amplification of splicing products after transient transfection of *BRCA1* minigenes into the MCF-7 cell line: lane *1*, *BRCA1* wt; lane *2*, *BRCA1* NL; lane *3*, *BRCA1* NL X; lane *4*, *BRCA1* NL Y.

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 XXYm-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/GUAAGU (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 UUUUUU CCCUNCAG/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 (XXY) 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.
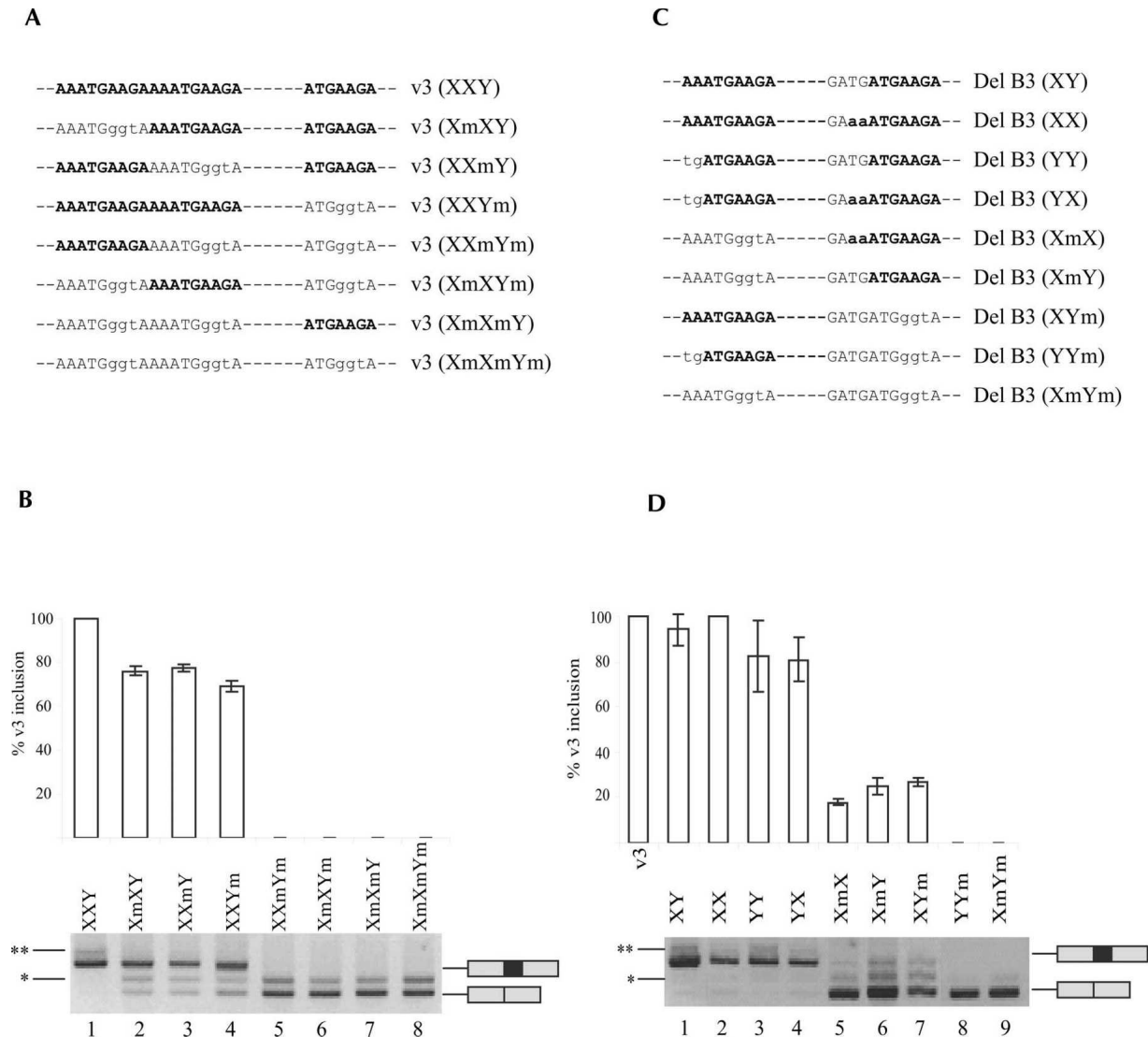
A

```
--AAATGAAGAAAATGAAGA------ATGAAGA--  v3 (XXY)
--AAATGggtAAAATGAAGA------ATGAAGA--  v3 (XmXY)
--AAATGAAGAAAATGggtA------ATGAAGA--  v3 (XXmY)
--AAATGAAGAAAATGAAGA------ATGggtA--  v3 (XXYm)
--AAATGAAGAAAATGggtA------ATGggtA--  v3 (XXmYm)
--AAATGggtAAAATGAAGA------ATGggtA--  v3 (XmXYm)
--AAATGggtAAAATGggtA------ATGAAGA--  v3 (XmXmY)
--AAATGggtAAAATGggtA------ATGggtA--  v3 (XmXmYm)
```

C

```
--AAATGAAGA-----GATGATGAAGA--  Del B3 (XY)
--AAATGAAGA-----GAaaATGAAGA--  Del B3 (XX)
--tgATGAAGA-----GATGATGAAGA--  Del B3 (YY)
--tgATGAAGA-----GAaaATGAAGA--  Del B3 (YX)
--AAATGggtA-----GAaaATGAAGA--  Del B3 (XmX)
--AAATGggtA-----GATGATGAAGA--  Del B3 (XmY)
--AAATGAAGA-----GATGATGggtA--  Del B3 (XYm)
--tgATGAAGA-----GATGATGggtA--  Del B3 (YYm)
--AAATGggtA-----GATGATGggtA--  Del B3 (XmYm)
```

B

D

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*, XXmYm; lane *6*, XmXYm; lane *7*, XmXmY; lane *8*, XmXmYm. (*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).

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 an 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
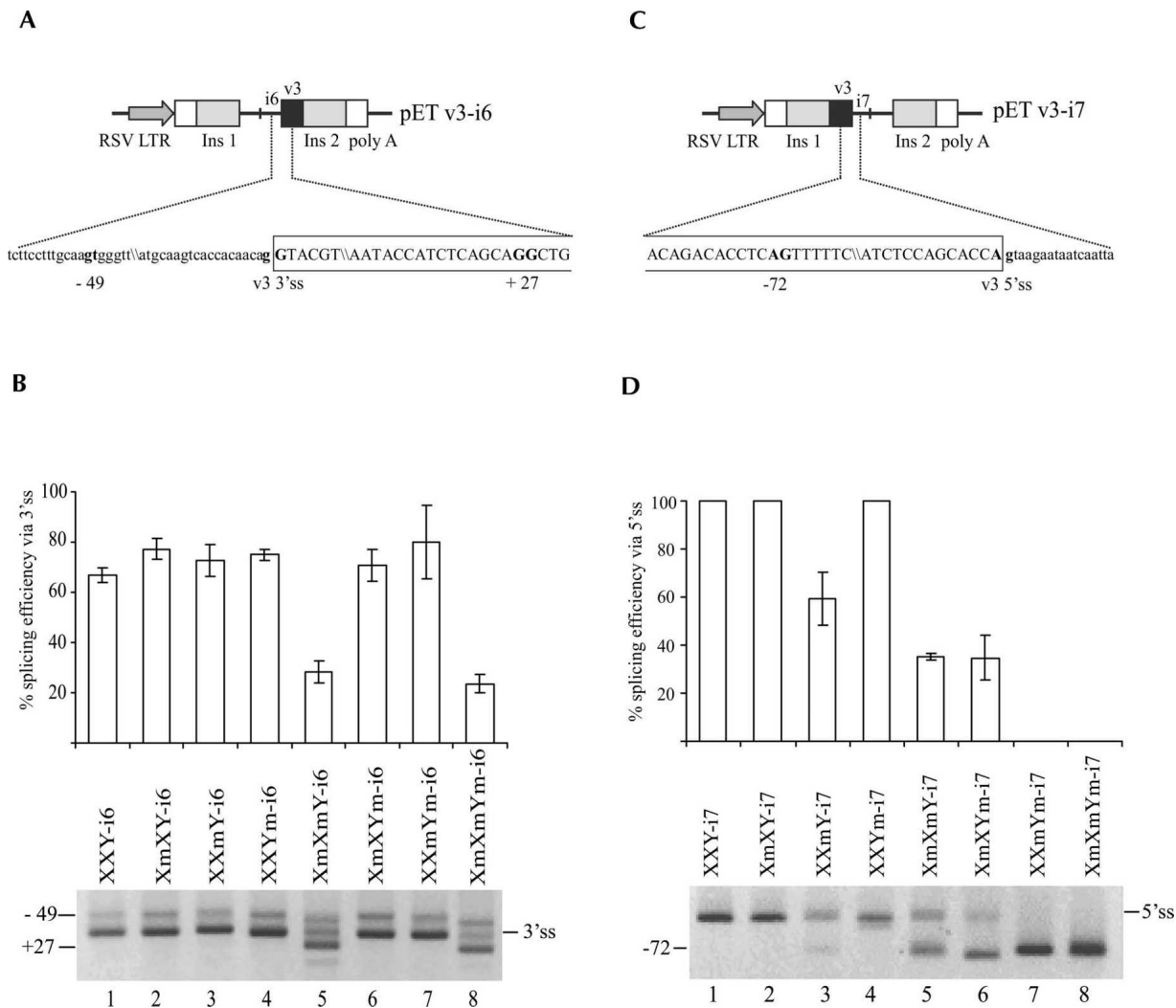
FIGURE 7. Single-intron minigenes. (*A*) *v3*-i6 minigenes were generated by PCR subcloning from *v3* constructs, deleting the *v3* 5′ss and intron 7. Exon sequences are indicated in upper case. Intron sequences are indicated in lower case. 3′ss are indicated in bold. Positions of cryptic 3′ss are numbered 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*, XXmY-i6; lane *4*, XXYm-i6; lane *5*, XmXmY-i6; lane *6*, XmXYm-i6; lane *7*, XXmYm-i6; lane *8*, XmXmYm-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*, XXmY-i7; lane *4*, XXYm-i7; lane *5*, XmXmY-i7; lane *6*, XmXYm-i7; lane *7*, XXmYm-i7; lane *8*, XmXmYm-i7.

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
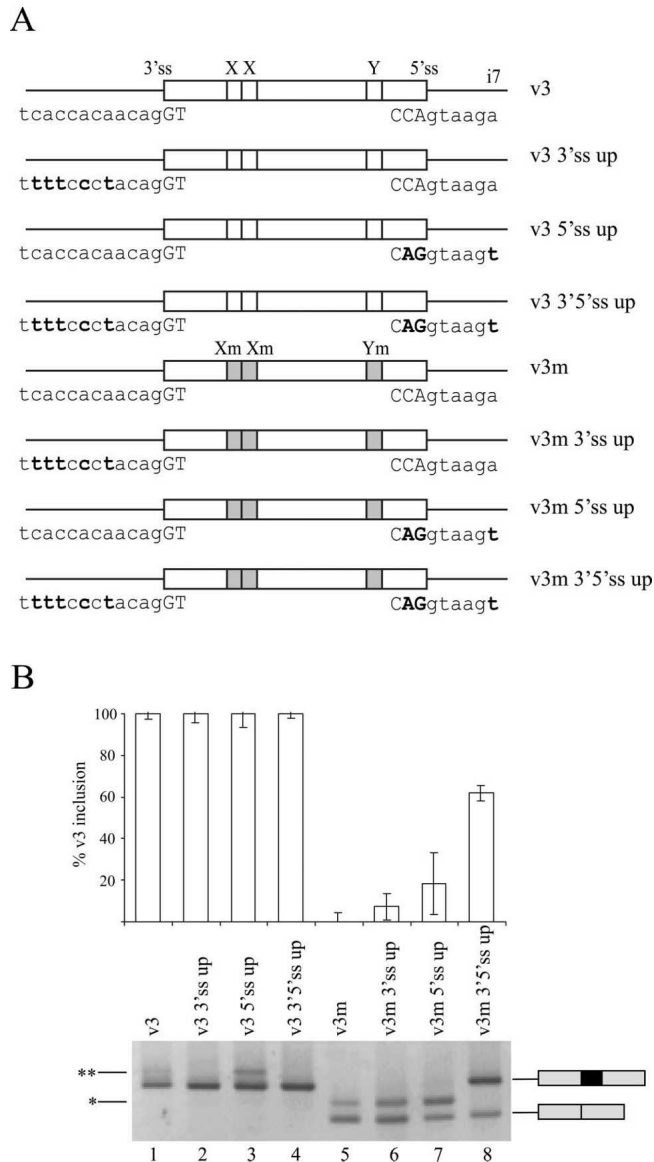
Vela et al.

A

B

FIGURE 8. Characterization of elements involved in *CD44* exon *v3* splicing. (*A*) Localization of *v3* splicing elements. Active XX and Y motifs are indicated as small white boxes. Mutant XX and Y motifs are indicated as small gray boxes. *v3* splice-site sequences are indicated. Upper case: exonic nucleotides. Lower case: intronic nucleotides. *v3* splice sites were mutated to match consensus sequences (3′ss and/or 5′ss up) by PCR-directed mutagenesis from *v3* and *v3*m constructs. Mutated positions are indicated in bold. (*B*) RT-PCR amplification of splicing products after transient transfection of splice-site mutants in the MCF-7 cell line: lane *1*, *v3*; lane *2*, *v3* 3′ss up; lane *3*, *v3* 5′ss up; lane *4*, *v3* 3′5′ss up; lane *5*, *v3*m; lane *6*, *v3*m 3′ss up; lane *7*, *v3*m 5′ss up; lane *8*, *v3*m 3′5′ss up. Extra bands (*) correspond to the (−49) pseudo-exon (Vela et al. 2006).

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 (ATGAAGAA, TGAAGAAA, 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 GAAAAT. In addition, the Y heptamer overlaps with the GATGAA 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 (*dsx*RE) 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 *dsx*RE 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 (*fru*RE) from *Drosophila* consists of three 13-nt repeat elements. In this case, only one repeat element within *fru*RE is necessary and sufficient to activate the female-specific splice site (Lam et al. 2003). *fru*RE and *dsx*RE 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 *E1A* 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 insuline, 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′-GCGTATGTTTTCTAGACAGCCTT GAATAAT-3′) with I7R5bX (5′-AATGAATCATCATAGACCAG GTTGT-3′) and I6F3X (5′-CTTTCAGCTTTCTAGAAGAGAATA TTCAGTT-3′) with I7R1X (5′-ATATTTTCCCCAACTTCTAGA AACATTCTA-3′). The first set, I6FX and I7R5bX, amplifies the full length of introns 6 and 7. The second set, I6F3X and I7R1X, 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 XbaI 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′-CCATGGTACCAGGCTGGGAGCCAAATGAA-3′) and R5′-v3Acc (5′-AAGAGGTACCTGTTGTTGTGACTTGCAT TA-3′), deleting from nucleotide 4 to nucleotide 25. Del B mutant was generated with primers F5′v3Xho (5′-AGATGCTCGAGACA GACACCGCAGTTTTTC-3′) and R5′v3Xho (5′-GCCTCTCGA GATGGTATTTGAAGACGTAC-3′), deleting from nucleotide 21 to nucleotide 61. Del C mutant was generated with primers F3′v3Nsi (5′-GATCATGCATTGATGATGATGAAGATTTT-3′) and R3′v3Nsi (5′-GTCTCATAGCATCTTCATTTTCTTCATTT GG-3′), deleting from nucleotide 57 to nucleotide 93. Del D mutant was generated with primers F3′v3Sph (5′-TTATCG CATGCACCAGTAAGAATAATCAAT-3′) and R3′v3Sph (5′-AT CATGCATGCCTGATCCAGAAAAACTGAG-3′), deleting from

nucleotide 94 to nucleotide 121. PCR fragments were inserted into pUC18 vectors, sequenced, and subcloned into pET vectors.

Small deletion mutants were generated by inserting PCR fragments into the Del B construct. Del B1 mutant was generated with primers DelB1F (5′-TCGAGCCAAATGAAGAAAATGA-3′) and pETR (5′-ATGGGTACCGAGCTCCAC-3′). Del B2 mutant was generated with primers DelB2F (5′-TCAGCAGGCTGGGAA GAAAATGAAGATGAAAGA-3′) and pETR. Del B3 mutant was generated with primers pETF (5′-CCTCGAGGTCGACGGT AT-3′) and DelB3R (5′-TCTTTCATCTTCATTTGGCTCCCAGC CTG-3′). Del B4 mutant was generated with primers pETF and Del B4R (5′-TCATTTTCTTCATTTGGCTC-3′).

Insertion mutants were generated by insertion of double-stranded oligonucleotides into the Del B mutant. Del B Ins Xs mutant was generated by insertion of the AAATGAAGA oligonucleotide. Del B Ins Xas mutant was generated by insertion of the TCTTCATTT oligonucleotide.

Point mutants were generated by PCR with the Site Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions.

Single-intron constructs were synthesized using PCR from pET *v3* short and *v3* XmXmYm constructs. For *v3*-i7 constructs, a PCR primer was used that initiates *v3* at an EcoRV restriction site followed by nucleotide 10. For *v3*-i6 constructs, a PCR primer was used that terminates *v3* at nucleotide 111 followed by an EcoRV restriction site.

BRCA1 minigene constructs

BRCA1 minigene NL (Liu et al. 2001) consisting of exons 17, 18, and 19 (kindly provided by Dr. A. Krainer, Cold Spring Harbor Laboratory) was excised from its original pIND vector and subcloned into the BamHI/XbaI sites of the pET vector for exon trap experiments.

All constructs were confirmed by sequencing with BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems).

RT-PCR analysis

RT-PCR was performed from 2 μg of total RNA extracted with Ultraspec RNA solution (Biotecx Laboratories). The first-strand reaction was performed with the First-Strand Synthesis Kit (Pharmacia), with Trap#1 primer (5′-GATCCACGATGC-3′) for *v3* minigene constructs and Random primer for *BRCA1* minigene constructs, followed by PCR using oligonucleotide primers N5′trap in *Insuline* exon 1 (5′-CCTGCTCATCCTCTGGGAGC-3′) and N3′Trap in *Insuline* exon 2 (5′-AGGTCTGAAGGT CACGGGCC-3′).

PCR conditions were 94°C for 1 sec, 62°C for 5 sec, and 72°C for 10 sec, for 40 cycles. Under these conditions, ratios between the amplification of exon inclusion and exon skipping products did not change when different numbers of cycles were performed. PCR products were separated by electrophoresis in a 2.5% agarose gel, visualized with ethidium bromide stain, and quantified as net intensity of each band with the Kodak-EDAS 120 system. Exon inclusion was defined as inclusion/(inclusion + skipping). Results were confirmed in at least three independent transient transfection assays. The variability between independent assay results is expressed as standard deviation bars in all graphics.

REFERENCES

Bajorath, J. 2000. Molecular organization, structural features, and ligand binding characteristics of CD44, a highly variable cell surface glycoprotein with multiple functions. *Proteins* **39:** 103–111.

Bell, M.V., Cowper, A.E., Lefranc, M.P., Bell, J.I., and Screaton, G.R. 1998. Influence of intron length on alternative splicing of CD44. *Mol. Cell. Biol.* **18:** 5930–5941.

Bourgeois, C.F., Popielarz, M., Hildwein, G., and Stevenin, J. 1999. Identification of a bidirectional splicing enhancer: Differential involvement of SR proteins in 5′ or 3′ splice-site activation. *Mol. Cell. Biol.* **19:** 7347–7356.

Breitbart, R.E., Andreadis, A., and Nadal-Ginard, B. 1987. Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* **56:** 467–495.

Cartegni, L., Chew, S.L., and Krainer, A.R. 2002. Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nat. Rev. Genet.* **3:** 285–298.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* **31:** 3568–3571. doi: 10.1093/nar/gkg616.

Cooper, T.A. 2005. Use of minigene systems to dissect alternative splicing elements. *Methods* **37:** 331–340.

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.

Fairbrother, W.G., Yeo, G.W., Yeh, R., Goldstein, P., Mawson, M., Sharp, P.A., and Burge, C.B. 2004. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. *Nucleic Acids Res.* **32:** W187–W190. doi: 10.1093/nar/gkh393.

Goren, A., Ram, O., Amit, M., Keren, H., Lev-Maor, G., Vig, I., Pupko, T., and Ast, G. 2006. Comparative analysis identifies exonic splicing regulatory sequences—The complex definition of enhancers and silencers. *Mol. Cell* **22:** 769–781.

Graveley, B.R. 2000. Sorting out the complexity of SR protein functions. *RNA* **6:** 1197–1211.

Graveley, B.R. 2001. Alternative splicing: Increasing diversity in the proteomic world. *Trends Genet.* **17:** 100–107.

Gunthert, U., Stauder, R., Mayer, B., Terpe, H.J., Finke, L., and Friedrichs, K. 1995. Are CD44 variant isoforms involved in human tumour progression? *Cancer Surv.* **24:** 19–42.

Herrlich, P., Zoller, M., Pals, S.T., and Ponta, H. 1993. CD44 splice variants: Metastases meet lymphocytes. *Immunol. Today* **14:** 395–399.

Hertel, K.J. and Maniatis, T. 1998. The function of multisite splicing enhancers. *Mol. Cell* **1:** 449–455.

Lam, B.J., Bakshi, A., Ekinci, F.Y., Webb, J., Graveley, B.R., and Hertel, K.J. 2003. Enhancer-dependent 5′-splice-site control of fruitless pre-mRNA splicing. *J. Biol. Chem.* **278:** 22740–22747.

Lavigueur, A., La Branche, H., Kornblihtt, A.R., and Chabot, B. 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes & Dev.* **7:** 2405–2417.

Lesley, J., Hyman, R., English, N., Catterall, J.B., and Turner, G.A. 1997. CD44 in inflammation and metastasis. *Glycoconj. J.* **14:** 611–622.

Liu, H.X., Cartegni, L., Zhang, M.Q., and Krainer, A.R. 2001. A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nat. Genet.* **27:** 55–58.

Ponta, H., Sherman, L., and Herrlich, P.A. 2003. CD44: From adhesion molecules to signaling reglators. *Nat. Rev. Mol. Cell Biol.* **4:** 33–45.

Ramchatesingh, J., Zahler, A.M., Neugebauer, K.M., Roth, M.B., and Cooper, T.A. 1995. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.* **15:** 4898–4907.

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

Reed, R. and Palandjian, L. 1997. Spliceosome assembly. In *Eukaryotic mRNA processing* (ed. A.R. Krainer), pp. 103–129. Oxford University Press, Oxford, UK.

Roca, X., Mate, J.L., Ariza, A., Munoz-Marmol, A.M., Von Uexkull-Guldeband, C., Pellicer, I., Navas-Palacios, J.J., and Isamat, M. 1998. CD44 isoform expression follows two alternative splicing pathways in breast tissue. *Am. J. Pathol.* **153:** 183–190.

Ruiz, P., Schwarzler, C., and Gunthert, U. 1995. CD44 isoforms during differentiation and development. *Bioessays* **17:** 17–24.

Screaton, G.R., Bell, M.V., Jackson, D.G., Cornelis, F.B., Gerth, U., and Bell, J.I. 1992. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proc. Natl. Acad. Sci.* **89:** 12160–12164.

Sheth, N., Roca, X., Hastings, M.L., Roeder, T., Krainer, A.R., and Sachidanandam, R. 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* **34:** 3955–3967. doi: 10.1093/nar/gkl556.

Stamenkovic, I., Amiot, M., Pesando, J.M., and Seed, B. 1989. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell* **56:** 1057–1062.

Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T.A., and Soreq, H. 2005. Function of alternative splicing. *Gene* **344:** 1–20.

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.

Vela, E., Roca, X., and Isamat, M. 2006. Identification of novel splice variants of the human CD44 gene. *Biochem. Biophys. Res. Commun.* **343:** 167–170.

Watakabe, A., Tanaka, K., and Shimura, Y. 1993. The role of exon sequences in splice site selection. *Genes & Dev.* **7:** 407–418.

Xu, R., Teng, J., and Cooper, T.A. 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. *Mol. Cell. Biol.* **13:** 3660–3674.

Zhang, X.H. and Chasin, L.A. 2004. Computational definition of sequence motifs governing constitutive exon splicing. *Genes & Dev.* **18:** 1241–1250.

Zhang, X.H., Kangsamaksin, T., Chao, M.S., Banerjee, J.K., and Chasin, L.A. 2005. Exon inclusion is dependent on predictable exonic splicing enhancers. *Mol. Cell. Biol.* **25:** 7323–7332.

Zhu, J., Shendure, J., Mitra, R.D., and Church, G.M. 2003. Single molecule profiling of alternative pre-mRNA splicing. *Science* **301:** 836–838.