MOLECULAR AND CELLULAR BIOLOGY, Dec. 1997, p. 7099–7107 Vol. 17, No. 12
0270-7306/97/$04.00+0

U1 Small Nuclear RNA-Promoted Exon Selection Requires a Minimal Distance between the Position of U1 Binding and the 3′ Splice Site across the Exon

DONG-YOUN HWANG AND JUSTUS B. COHEN*

Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261

Received 25 July 1997/Accepted 4 September 1997

Both experimental work and surveys of the lengths of internal exons in nature have suggested that vertebrate internal exons require a minimum size of approximately 50 nucleotides for efficient inclusion in mature mRNA. This phenomenon has been ascribed to steric interference between complexes involved in recognition of the splicing signals at the two ends of short internal exons. To determine whether U1 small nuclear ribonucleoprotein, a multicomponent splicing factor that is involved in the first recognition of splice sites, contributes to the lower size limit of vertebrate internal exons, we have taken advantage of our previous observation that U1 small nuclear RNAs (snRNAs) which bind upstream or downstream of the 5′ splice site (5′SS) stimulate splicing of the upstream intron. By varying the position of U1 binding relative to the 3′SS, we show that U1-dependent splicing of the upstream intron becomes inefficient when U1 is positioned 48 nucleotides or less downstream of the 3′SS, suggesting a minimal distance between U1 and the 3′SS of approximately 50 nucleotides. This distance corresponds well to the suggested minimum size of internal exons. The results of experiments in which the 3′SS region of the reporter was duplicated suggest an optimal distance of greater than 72 nucleotides. We have also found that inclusion of a 24-nucleotide miniexon is promoted by the binding of U1 to the downstream intron but not by binding to the 5′SS. Our results are discussed in the context of models to explain constitutive splicing of small exons in nature.

Pre-mRNA splicing requires at least four sequence signals, the 5′ splice site (5′SS) signal, 3′SS signal, branchpoint sequence (BPS), and polypyrimidine tract (PPT). Recognition of these signals by small nuclear RNAs (snRNAs) and proteins results in the formation of a splicing complex, the spliceosome, which directs accurate and efficient intron removal in a two-step reaction (for reviews, see references 19, 38, and 44). In mammalian genes, where the splicing signals are only loosely conserved and potential cryptic splice sites are abundant, it remains to be fully understood how multiple exons and introns are recognized and committed to splicing in an orderly and accurate manner (see reference 4 for a review).

According to the exon definition model, SSs in vertebrates are recognized in pairs across exons (1, 46), which are fairly short and homogeneous in size compared to vertebrate introns. Among the strongest arguments for this model, mutations in either SS signal at the boundaries of internal exons frequently cause exon skipping in vivo and in vitro (30, 31, 35, 37, 40, 55, 59). Thus, both SS signals are ignored when only one is mutated, but the distal signals across the upstream and downstream introns remain active. Coordinate recognition of exon boundaries is also strongly suggested by the observation that mutations which create SSs inside introns can activate nearby cryptic sites of opposite polarity to form a new exon (40, 58).

Exon definition is believed to involve exon-bridging interactions that support communication between factors at the two ends (1, 46). For example, mutations that improve a weak 5′SS to provide a better recognition sequence for U1 snRNA stimulate the 3′SS across the exon in vitro (18, 31) and reduce the accessibility of the upstream 3′SS to antisense oligonucleotides (12), indicating that events at the 5′SS influence the activity of the 3′SS at the other end of the exon. Likewise, we have previously shown that splicing of an upstream intron is stimulated by the binding of U1 snRNA at or near the 5′SS downstream (24). The idea that these effects are mediated by exon-bridging interactions is supported by the observation that base pairing of U1 snRNA with the 5′SS increases binding of the essential splicing factor U2AF65 to the PPT across the exon (23); U2AF65 promotes association of U2 small nuclear ribonucleoprotein (snRNP) with the branchpoint region (49). Furthermore, the recognition of exons with weak splicing signals can be stimulated by exon-internal sequences (referred to as exonic splicing enhancers) which appear to be recruitment sites for complexes that may interact simultaneously with U1 snRNP at the 5′SS and U2AF65 at the 3′SS (51, 68). It has previously been shown that exonic splicing enhancers promote the binding of U2AF65 to the upstream 3′SS region (61) and stimulate the removal of the upstream intron (20, 32, 39, 62, 65), just as a downstream 5′SS does (23, 31, 55). By using purified recombinant proteins, it has been also demonstrated that U2AF65 and U1 snRNP can be connected through a series of protein-protein interactions which involve SR proteins (63), and SR proteins have been detected both in the E complex, the first specific prespliceosome complex (51), and in enhancer-binding complexes that support E complex formation (32, 43, 51, 53, 57). SR proteins, named for their characteristic domains that are rich in arginine-serine repeats (for reviews, see references 16 and 34), can bind RNA with a degree of sequence specificity (22, 54, 69) and can interact with other proteins, such as the U1 snRNP-specific 70,000-molecular-weight protein (U1 70K protein), the small subunit of U2AF (known as U2AF35), and other SR proteins (26, 63). Accordingly, SR proteins are strong candidates to support exon definition as bridging factors between SSs across exons, whether or not they are recruited to the exon by an internal

* Corresponding author. E-mail: jbc+@pitt.edu.

enhancer (1). Indeed, SR proteins appear to play an early role in SS recognition, as suggested by their ability to influence alternative SS choice (14, 17, 27, 66).

Simultaneous recognition of exon termini, as proposed by the exon definition model, is problematic when the size of the exon is insufficient to accommodate the footprints of complexes that assemble at the two termini, especially when these complexes must be properly situated to interact (1, 52). It has previously been shown that the last 17 or 18 nucleotides (nt) of different exons are protected from RNase digestion by splicing complexes which are assembled on the 5′SS in vitro (5, 28). Likewise, the results of previous experiments with uncapped human β-globin pre-mRNA have suggested that a region starting at least 15 nt upstream of the 5′SS is protected from the activity of an endogenous 5′-to-3′ exonuclease (48). Furthermore, U1 snRNP and other factors that bind to pseudo-5′SSs 15 to 29 nt upstream of the *Drosophila* P element third intron in somatic extracts have been shown to inhibit the binding of U1 snRNP to the correct 5′SS, which may account for somatic inhibition of splicing of this intron (50). In more general systems, mutual interference between nearby U1 binding sites has also previously been demonstrated (14, 42). In addition, previous splicing inhibition experiments with antisense oligonucleotides have demonstrated that at least 25 nt of exon sequence from both the 5′SS and the 3′SS are engaged in important interactions in vitro (12), suggesting that at least 50 nt of exon sequence are occupied by complexes formed on splice sites that flank internal exons. Indeed, it was found for a constitutive exon that internal deletion caused skipping, which was reversed when the length of the exon was increased to 51 nt (10). Consistent with these different observations, a previous survey of exon lengths in nature has shown that no more than 4% of vertebrate internal exons are shorter than 50 nt (21) and it has been suggested that recognition and splicing of these short exons involve specialized approaches (1).

U1 snRNA normally interacts with 5′SSs through base pairing. Results from other laboratories (11, 23, 31, 46) as well as our own work (24) indicate that this interaction stimulates the 3′SS upstream, at the other end of the exon. We have previously shown, however, that this activity of U1 is not conditional on base pairing at the 5′SS, since U1 snRNAs with base pairing specificity for sequences upstream or downstream of the 5′SS, shift U1 RNAs, also stimulate the upstream 3′SS (24). Shift U1 RNAs have mutations at the 5′ end of U1 snRNA in the sequence that normally pairs with 5′SSs (9). These mutations specifically direct shift U1 RNAs to complementary sequences of choice elsewhere along the pre-mRNA, as demonstrated for a number of shift U1 RNAs by mutation of the intended binding site, which abolished activity, and complementary mutation of the shift U1 RNA, which restored activity (9, 24). In our system, shift U1 RNAs promote splicing to a normally silent 3′SS upstream at the expense of splicing to a constitutive 3′SS downstream (24) (see Fig. 1).

Using shift U1 RNAs to vary the position of U1 binding, we examined the possibility that the normal requirement for U1 snRNP binding to the 5′SS of internal exons dictates a minimum exon size and that shorter exons can be spliced when U1 is moved outside. We show here that activation of the upstream 3′SS by U1 was dramatically reduced when U1 approached the 3′SS. Increasing the distance between the 3′SS and the position of U1 binding, without changing the distance between U1 and the 5′SS, restored 3′SS activity, which demonstrates that the distance between U1 and the 3′SS, not the 5′SS, is critical. The minimum distance for efficient 3′SS activation suggested by our results is approximately 50 nt, which is similar to the lower size limit of vertebrate internal exons, as derived from deletion experiments and statistical analyses (1, 10, 21). The results of experiments with competing 3′SSs further indicate an optimal distance between U1 and the 3′SS of 72 nt or more. These results are consistent with increasing constraints on the interaction between U1 snRNP and factors that assemble on the 3′SS when the U1 binding site is moved closer to the 3′SS; at 50 nt or less, the two sites may no longer be accessible simultaneously. We also show that recognition of a 24-nt miniexon was promoted by base pairing of U1 with sequences in the downstream intron but not by base pairing of U1 with the 5′SS region. This suggests that U1, rather than some other factor which interacts with the 5′SS, is at least partly responsible for the lower size limit of internal exons. Our results indicate that small introns in nature must rely on special mechanisms to circumvent the size constraints imposed by the normal binding of U1 to the 5′SS.

MATERIALS AND METHODS

Plasmids. Reporter plasmids pGT+ and pIDX:CTC (previously called phGH.FSX:CTC), control vector pDNase ^ E5, the two U6 plasmids (see Fig. 5), and the generation of shift U1 plasmids were described before (9, 24, 25). The previous names for some of the U1 plasmids used here were U1-α(-23) for U55, U1-α(-6) for U72, U1-αGT2 for U85, and U1-α+14 for U99 (9, 25).

p2(3′ss), which contains a tandem duplication of the 3′SS region (see Fig. 3A), was constructed from three parts. A 976-bp fragment containing the 5′ half of the reporter gene from an *Xba*I site just before the human growth hormone (hGH) start codon to a *Bgl*I site at the start of IDX was obtained from plasmid phGH.FSX:A5 (8); the 3′ overhang left by *Bgl*I digestion was removed with T4 DNA polymerase. A second fragment was generated by PCR with a forward primer [Yn(+); see below] that initiates with the first nucleotide of the presumed BPS upstream of IDX and a reverse primer (Ras/hGH2860) complementary to a region inside intron C2; pGT+ was used as the template. The PCR fragment was treated with Klenow DNA polymerase and digested with *Bst*EII, which cuts in the reverse primer sequence. The two fragments were ligated to a vector piece obtained by digestion of phGH (an hGH construct without IDX [8]) with *Xba*I and *Bst*EII. p2(3′ss)+21 was generated by a similar approach; the 1,009-bp *Xba*I-*Bgl*I fragment from p2(3′ss) was used, and a PCR fragment was generated with a different forward primer, IDX+21, which starts in the *Bgl*I site near the 5′ end of IDX and specifies a 21-nt insertion (see below).

pME:CTC was also created by the insertion of two fragments into phGH. A 994-nt fragment containing the 5′ half of the reporter gene was generated by digestion of phGH.FSX:wt (24) with *Bgl*I, removal of the *Bgl*I protruding end with T4 DNA polymerase, and digestion with *Hin*dIII. The second fragment was generated by PCR with pIDX:CTC as the template and with forward primer ME and reverse primer Ras/hGH2860; forward primer ME starts 10 nt before the 3′ end of IDX and has an *Xho*I site fused upstream. The PCR fragment was treated with Klenow DNA polymerase and digested with *Bst*EII. The two fragments were then ligated to a *Hin*dIII-*Bst*EII vector fragment from phGH. The sequences of all PCR-derived regions were verified by DNA sequencing.

S1 probes. End-labeled probes with tails of unrelated DNA were prepared as described previously from pUC recombinants with the specific portion of the probe in the polylinker region (24). To generate probe 1 of Fig. 2, a 331-bp *Nhe*I-*Pst*I fragment from pGT+, covering IDX and part of the two flanking introns, was placed between the *Xba*I and *Pst*I sites of pUC118. For probes 1 of Fig. 3 and 4, the corresponding *Nhe*I-*Pst*I fragments from p2(3′ss) and p2(3′ss)+21 were introduced into pUC119. Probe 1 of Fig. 2 was produced by PCR with the pUC recombinant, a pUC-specific forward primer (118-1004), and a 5′-end-labeled reverse primer [Ras2676(−)] that recognizes IDX. Probes 1 of Fig. 3 and 4 were generated in the same way, except that primer 118-1004 was replaced by Seq, a different pUC-specific primer.

Assays. 293 cells were transfected by the calcium phosphate procedure with 4 to 8 μg of pDNase ^ E5, 5 to 10 μg of hGH reporter plasmid, 10 to 20 μg of U1 plasmid, and 10 to 20 μg of U6 plasmid where indicated. Metabolic labeling, immunoprecipitation, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, cytoplasmic RNA preparation, and S1 nuclease protection were performed as described previously (8, 24).

Oligonucleotide sequences. The sequences of the primers used are as follows: Yn(+), CACTGACTCCTCTCCCTTGAC (intron C2; forward); IDX+21, CTCTGGCACTGGACTACAGCACTTCACCTCTAGCTCCAGCTCCGG (IDX; forward; insertion underlined); ME, CTCGAGCCCTCGC(C/G)(A/C)(G/T)G (IDX; forward); Ras/hGH2860, CCGAGGTCACCGCTCCGGCCTGGCTCAG (intron C2; reverse); Ras2676(−), GGGGGTCCCAGAGGGTCC (IDX; reverse); 118-1004, CACAGGAAACAGCTATGACCATG (pUC); and Seq, GTTTTCCCAGTCACGAC (pUC).

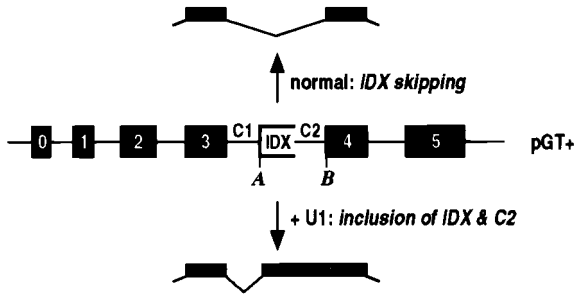

FIG. 1. Reporter gene in construct pGT+ (center) and splicing in the absence (normal) or presence (+U1) of foreign U1 snRNAs. In the center, hGH exons are represented by black boxes, IDX is represented by an open-ended empty box, and introns are represented by connecting lines; the open end of the IDX box symbolizes the absence of a functional 5′SS. The 3′SSs used in the absence (*B*) or presence (*A*) of foreign U1 RNAs are indicated. In the diagrams at the top and bottom, the regions included in spliced mRNA are represented by black bars. The difference between the two diagrams is the exclusion or inclusion of IDX and intron C2 (24, 25), referred to as IDX skipping and inclusion of IDX and C2, respectively.

RESULTS

Experimental system. We have previously reported that shift U1 RNAs that bind to a region between a silent upstream 3′SS (*A*; Fig. 1) and a constitutive downstream 3′SS (*B*; Fig. 1) activate the upstream site, which occurs at the expense of splicing to the downstream site unless a functional 5′SS is present between the two 3′SSs (24). The reporter constructs in these experiments expressed the LGH gene, which was interrupted between exons 3 and 4 by an insertion from the human H-*ras* gene containing a splicing-defective copy of the H-*ras* exon IDX (Fig. 1). In construct pGT+, the 5′SS of IDX was inactive due to mutations that abolished base pairing with endogenous U1 and U6 snRNAs (25). Without base pairing between U1 and the 5′SS, the 3′SS across IDX was also silent and IDX was excluded from mature mRNA (Fig. 1, normal). We showed that the 3′SS could be activated by coexpression of U1 RNAs with compensatory base changes to the 5′SS mutations as well as by shift U1 RNAs with 9-nt complementarities to regions upstream or downstream of the mutated 5′SS (9, 24). In these situations (Fig. 1, +U1), splicing occurred between exon 3 and IDX (intron C1), not between exons 3 and 4 (intron C). However, splicing between IDX and exon 4 (intron C2) was not observed because the 5′SS was defective in other interactions (24). Thus, the entire region from the start of IDX to the end of exon 4 was included in mature mRNA (inclusion of IDX and C2), creating a situation of competing 3′SSs, with exclusive use of the downstream site (*B*) under normal conditions and predominant use of the upstream site (*A*) in the presence of foreign U1 RNA binding between *A* and *B*. We used this system to determine whether U1-dependent activation of the upstream site requires a minimum distance between this site and the position of U1 binding downstream.

Activation of an upstream 3′SS as a function of the position of U1 binding. Shift U1 RNAs which could bind to sequences inside IDX or downstream were generated. The binding positions of these shift U1 RNAs relative to the upstream 3′SS *A* and relative to the defective 5′SS at the end of IDX are diagrammed in Fig. 2A; the nomenclature notes the distance between the binding site for each shift U1 RNA and the upstream 3′SS. The actual binding sequences for the different shift U1 RNAs in construct pGT+ are indicated in Fig. 2B, which shows the sequence of IDX and flanking H-*ras* sequences. IDX is a natural exon, but in pGT+, the sequence

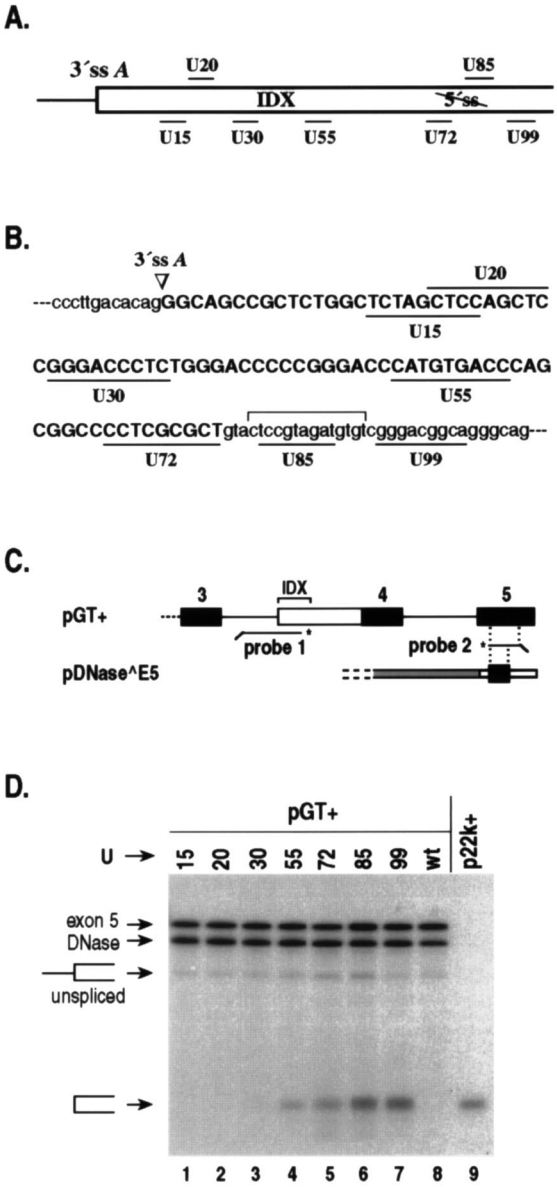

FIG. 2. Activation of a 3′SS by U1 snRNAs targeted to different positions downstream. (A) Positions of the recognition sites for different U1 RNAs in pGT+. The name of each U1 RNA indicates the distance in nucleotides between the first position of its recognition site and the upstream 3′ splice junction (3′SS *A*). The position of the silent 5′SS at the end of IDX is indicated. (B) Sequences of IDX (uppercase letters) and flanking regions (lowercase letters) in pGT+, showing the binding sites (under- or overlined) for various shift U1 RNAs. The entire sequence is from the human H-*ras* gene, with the exception of the bracketed region, which has mutations to inactivate the 5′SS (for details, see reference 25). (C) S1 nuclease protection probes previously described (24). Both probes had specific (horizontal line) and nonspecific (diagonal tail) portions, and both were end labeled (*). In the representation of pDNase ^ E5, the DNase coding sequence is the shaded area and the black box is the hGH exon 5 insertion downstream. In the pGT+ illustration, IDX and intron C2 are shown as a single open box to symbolize that in our experiments these elements are either both excluded from spliced RNA or included together. (D) S1 protection analysis of cytoplasmic RNAs from transiently transfected 293 cells. Protected fragments are identified on the left. The two upper bands (exon 5 and DNase) were derived from probe 2, and the two lower bands were derived from probe 1. Probe 2 was omitted in lane 9. wt, wild type.

just downstream of IDX has mutations that inactivate the 5′SS (Fig. 2B) (25).

Plasmids expressing different shift U1 RNAs were cotransfected with pGT+ into 293 cells together with a vector for standardization, pDNase^E5; the standardization vector expresses DNase protein from a cDNA with a portion of hGH exon 5 in its 3′ untranslated region (Fig. 2C) (24). Cytoplasmic RNAs were isolated 2 days later and analyzed by S1 nuclease protection assay with the two end-labeled probes illustrated in Fig. 2C. Probe 2 detected all hGH and DNase transcripts as two bands of distinct sizes (exon 5 and DNase, respectively) (Fig. 2D) (24). Probe 1 detected activation of the 3′SS of IDX (smallest band) (Fig. 2D) as well as unspliced pGT+ RNA (Fig. 2D); the small product was identified by comparison with lane 9, where RNA from cells transfected with plasmid p22k+, which expresses an IDX-containing hGH cDNA that lacks introns C1 and C2 (9), was analyzed.

The results in Fig. 2D show that the 3′SS of IDX was activated by U1 snRNAs directed to binding sites 55 nt or more downstream (lanes 4 through 7), with the efficiency of activation increasing at greater distances to an apparent plateau at 85 nt (lanes 6 and 7). However, no activation was seen with U15, U20, or U30, which have recognition sites starting 30 nt or less downstream of the 3′SS (Fig. 2D, lanes 1 through 3). These results suggest both a lower limit to the distance between the U1 binding site and the 3′SS upstream and an optimal distance. However, they do not exclude the alternative possibility that the observed differences are simply due to inherent differences among the U1 RNAs used. In addition, they do not demonstrate that the reference point for the apparent position effect is the upstream 3′SS, instead of some downstream signal that stimulates U1 activity by proximity, for example, the (mutated) 5′SS signal at the end of IDX.

Distance to the U1 binding site determines upstream 3′SS activity and choice. To address the uncertainties mentioned above, two new constructs, p2(3′ss) and p2(3′ss)+21, were derived from pGT+. Both had a 33-nt tandem duplication that spanned the junction of intron C1 and IDX (3′SS*A*), including the presumed BPS, PPT, and 3′-AG of intron C1, and the first 8 nt of IDX; this duplication created alternative 3′SSs for intron C1 (Fig. 3A). In addition, p2(3′ss)+21 had a 21-nt insertion of random sequence inside IDX, immediately upstream of the recognition site for U15 (Fig. 3A; see Materials and Methods for the sequence of the insertion). The two constructs were tested as described above by cotransfection with U1 and DNase vectors and S1 protection analysis. The binding sites for the four U1 RNAs included in this experiment are indicated in Fig. 3A, along with their distances to the tandem 3′SSs upstream.

As discussed below, the results in Fig. 3B demonstrate that both 3′SSs can be used and that the distal site (relative to the position of U1 binding) is preferred over the proximal site when U1 is close, but not when U1 is moved farther away. In addition, they demonstrate that these position effects relate to an upstream reference point, presumably the 3′SS region, not to a downstream reference point, and that the activities of different U1 RNAs differ as a function of binding position, not as a consequence of inherent differences between them, for example, in binding affinities or expression levels.

With p2(3′ss) (Fig. 3B, lanes 6 through 10), the proximal 3′SS was activated by U85 (lane 9) but was not activated by the three U1 RNAs that bind closer, at distances of 55 nt or less (lanes 6 through 8). Instead, the last three U1 RNAs activated the distal 3′SS, located 48 to 88 nt upstream, although with different efficiencies; in particular, U15 (Fig. 3B, lane 6) was very inefficient compared to U30 and U55 (lanes 7 and 8),

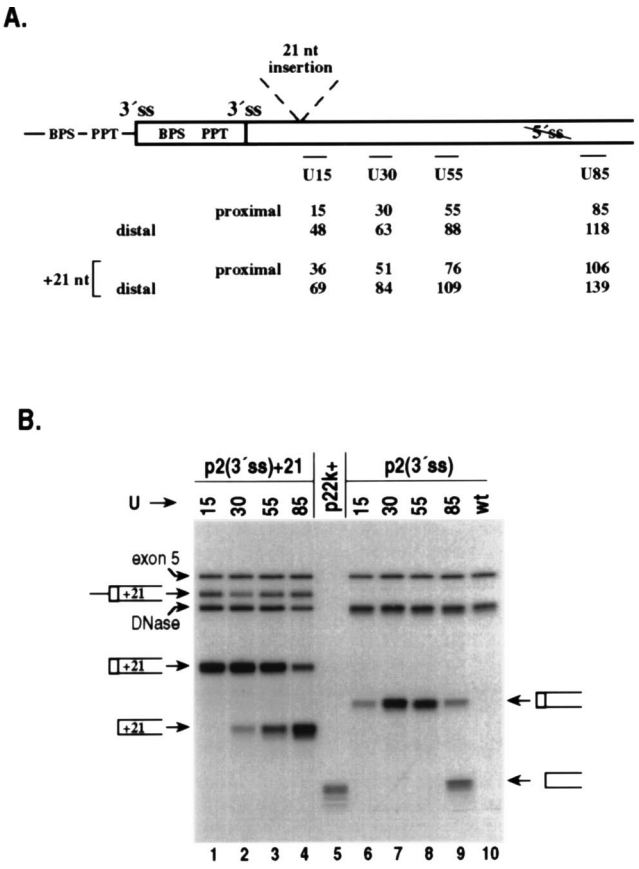

FIG. 3. Activation of alternative 3′SSs by U1 RNA binding downstream and effects of an insertion between 3′SSs and U1 binding sites. (A) U1 binding sites are illustrated below a representation of IDX in the two constructs with duplicated 3′SS regions, p2(3′ss) and p2(3′ss)+21; the second construct also had a 21-nt insertion in IDX. The distances between U1 binding sites and upstream (distal) and downstream (proximal) 3′ splice junctions are indicated below each construct. (B) S1 protection analysis of cytoplasmic RNAs from transiently transfected 293 cells. The types of probes used were the same as those used in Fig. 2, but probe 1 had the 3′SS duplication with (lanes 1 through 4) or without (lanes 5 through 10) the 21-nt insertion in IDX; probe 2 was omitted in lane 5. Products are identified on the sides. The product protected by unspliced reporter RNA is indicated only for lanes 1 through 4 (second band from the top); the corresponding product in lanes 6 through 10 appears as a shadow just below the DNase band.

which is reminiscent of previous results where proximity to the 3′SS appeared to reduce activation. However, instead of reaching a plateau with increasing distance to U1, the activity of the distal 3′SS peaked with U30 (Fig. 3B, lane 7) and decreased thereafter (lanes 8 and 9), ultimately in favor of the proximal site (lane 9).

In the second reporter construct, p2(3′ss)+21, the distances between the positions of U1 binding and the two 3′SSs are 21 nt greater than those in the first construct, p2(3′ss), but there is no change in the positions of U1 binding relative to possible downstream points of reference. The extra 21 nt clearly facilitated activation of the distal 3′SS by U15, now 69 nt away (Fig. 3B, lane 1), indicating that U15 showed little or no activity before because its binding site was too close to either of the upstream 3′SSs (Fig. 2D, lane 1; Fig. 3B, lane 6). The extra 21 nt also allowed activation of the proximal 3′SS by U1 RNAs that could not demonstrably do so before, i.e., U30 and U55 (Fig. 3B; compare lanes 2 and 3 to lanes 7 and 8), which were now 51 and 76 nt, respectively, removed from the proximal site.

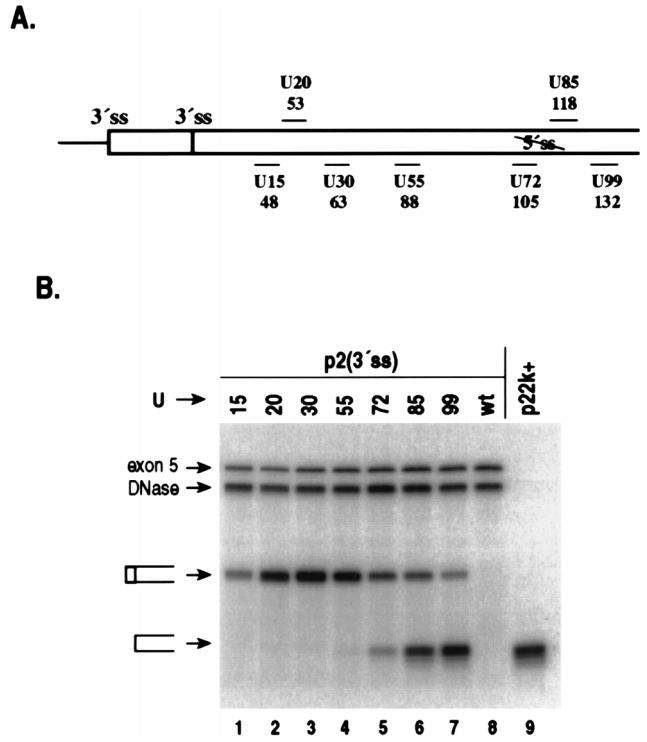

FIG. 4. Activation of alternative 3′SSs by a range of U1 RNAs. (A) The positions of U1 binding sites in the IDX region of reporter plasmid p2(3′ss) are illustrated, and the distances to the distal 3′SS upstream are shown below each U1 RNA. (B) S1 nuclease protection results. The experimental conditions were the same as those described in the legends to Fig. 2 and 3. Probe 1 was the same as that used for lanes 5 through 10 of Fig. 3, and probe 2 was omitted in lane 9.

To further define the spacing requirements between the 3′SS and the U1 binding position, additional shift U1 RNAs were tested with p2(3′ss). These included U20 and U72 (Fig. 4A), which bind in suggested transition areas (Fig. 3B, lanes 6 through 9). The results are shown in Fig. 4B. Taking into account slight differences in the exon 5 signals in Fig. 4B, lanes 1 through 3, a significant drop in the activity of the distal 3′SS is apparent between U20 and U15, pointing to a transition at a distance of around 50 nt (the distances to the distal site are shown in Fig. 4A). The activity of the distal site peaked at 63 nt (Fig. 4B, lane 3), and this peak was followed by a progressive decrease in activity at 88 nt and beyond (lanes 4 through 7), where competition with the proximal site became increasingly evident. In lane 4 of Fig. 4B, a very slight signal for the proximal site can be observed but is absent in lane 3, and there is a slight decrease in the signal for the distal site. In Fig. 3B, competition was evident in lane 2, but not in lane 1. This suggests that the proximal site becomes competitive when U1 binds approximately 50 nt downstream (55 [Fig. 4B, lane 4] and 51 [Fig. 3B, lane 2] nt). Thus, there are two notable transitions, when U1 binds 50 nt downstream of the distal site (where this site gained activity rapidly [Fig. 4B, lanes 1 and 2]) and when U1 binds 50 nt downstream of the proximal site (where signs of competition between the two sites were first observed). This indicates that 50 nt represents a boundary between distances that do and do not substantially restrict activation of upstream 3′SSs by U1.

U1 snRNA promotes splicing of a miniexon from inside the downstream intron, but not when it is targeted to the 5′SS. To test directly the implication that exons which are smaller than 50 nt cannot usually rely on base pairing between U1 and the

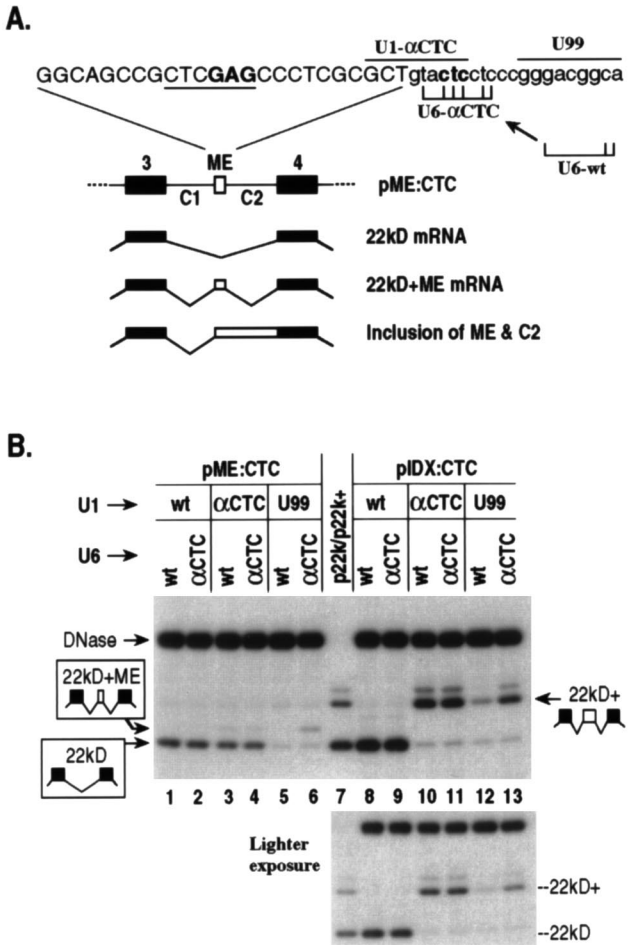

FIG. 5. U1- and U6-promoted inclusion of a miniexon. (A) The miniexon construct, pME:CTC, has the central portion of IDX replaced by GAG (bold uppercase letters), which reduces the exon (ME [for miniexon]) to 24 nt (uppercase letters) and creates an *Xho*I site (underlined). The sequence at the 5′ end of intron C2 is also shown (lowercase letters), with differences from the wild-type sequence in boldface type (positions 4 through 6; agu→ctc). The recognition sites for U1-αCTC and U99 are overlined, and potential base pairing interactions between U6-αCTC and the start of intron C2 are represented by the forked shape with six teeth; the forked shape with three teeth represents wild-type (wt) U6, which has three fewer positions of complementarity to the same sequence of intron C2. Alternative splicing patterns are illustrated. (B) Analysis of proteins produced by the miniexon construct (pME:CTC) and the corresponding full-length IDX construct (pIDX:CTC) (previously called phGH.FSX:CTC [24]) by immunoprecipitation of medium from transiently transfected 293 cells. A plasmid expressing secreted DNase protein was cotransfected as an internal control, except in lane 7, where cDNAs for 22kD (normal hGH without IDX or ME) and 22kD+ (hGH with IDX) were transfected together (8). Two days after transfection, cells were incubated with [^{35}S]methionine-cysteine, the medium was immunoprecipitated with combined hGH- and DNase-specific antibodies, and precipitates were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (8). Products are identified on the sides of the gel. A lighter exposure of the right side of the gel (lanes 7 through 13) is included below the gel to correct for a general difference in expression between the two reporter constructs (compare the 22kD signals in lanes 1 and 8).

5′SS to promote their inclusion in spliced RNA, two hGH-based constructs, one with full-length (81-nt) IDX (pIDX:CTC) and the other with a 24-nt miniexon derived from IDX by internal deletion (pME:CTC), were compared (Fig. 5A). These constructs had positions 4 through 6 of wild-type intron C2 changed from AGT to CTC to prevent the pairing of endogenous U1 with the 5′SS sequence (24). Expression of the two vectors after transfection was analyzed by a well-charac-

terized protein expression assay used previously to distinguish the splicing patterns illustrated in Fig. 5A. The first pattern, exclusion of the internal exon, produces normal hGH protein, referred to here as 22kD. The second pattern is inclusion of the internal exon, which results in the production of a larger protein, termed 22kD+ in the case of full-length IDX and 22kD+ ME for the miniexon. The third pattern is that described above for pGT+ in the presence of shift U1 RNAs, splicing of intron C1 and inclusion of the miniexon or IDX together with intron C2. This occurs when the upstream 3′SS is activated by U1 without activation of the mutated 5′SS (Fig. 1). This pattern produces no detectable protein because translation terminates in intron C2. Increased splicing of this type therefore results in a decrease in the total amount of hGH proteins produced (24). The protein assay was used here because protection of miniexon S1 probes was very inefficient as a result of the size of the miniexon.

Figure 5B shows that each vector alone produced only 22kD, confirming that the CTC mutation prevents inclusion of the internal exon, whether it is IDX or the miniexon (lanes 1 and 8) (lane 7 shows the positions of the smaller [22kD] protein and the larger [22kD+] protein produced by cotransfection of two cDNA vectors). U1-αCTC is a variant U1 snRNA with complete complementarity to the mutated 5′SS signal itself (positions −3 to +6 relative to the splice junction; Fig. 5A). As observed previously with U99 (24), coexpression of U1-αCTC with the IDX construct reduced IDX skipping in favor of IDX inclusion (Fig. 5B, lane 10). In contrast, U1-αCTC expression had essentially no effect on the miniexon construct (Fig. 5B, lane 3), indicating that U1-αCTC is unable to promote inclusion of the miniexon or splicing of intron C1 alone. To determine whether this is due to unfavorable positioning of U1-αCTC, U99 was tested. The results show that U99, which binds in the C2 region downstream of IDX and the miniexon, reduced the amount of 22kD for both reporter constructs (Fig. 5B, lanes 5 and 12) compared to that for wild-type U1 (lanes 1 and 8). We have shown previously for the IDX construct that the reduction in the amount of 22kD is caused by increased splicing of intron C1 and that a corresponding increase in the amount of 22kD+ is not seen with U99 or other shift U1 RNAs because intron C2 remains mostly unspliced (24) (see below). Likewise, the reduced 22kD signal generated by the miniexon construct (Fig. 5B, lane 5) indicates that U99 stimulates the splicing of intron C1 across the miniexon as well; in this case, no splicing of intron C2 is evident. Thus, although U1-αCTC did not reduce the amount of 22kD produced by the miniexon construct (Fig. 5B, lane 3) and therefore apparently did not stimulate intron C1 splicing, U99 did, which is consistent with a requirement for a minimal distance between U1 and the 3′SS. The distance between U99 and the 3′SS across the miniexon (35 nt) is at the lower end of the minimum range suggested above, which may explain why the U99-mediated decrease in 22kD production is less for the miniexon construct than for the IDX construct (approximately twofold by phosphorimager measurement [not shown]), but this difference is perhaps smaller than expected.

Exon inclusion was evident in Fig. 5B, lane 12 but was not evident in lane 5, suggesting that splicing of intron C2 may be more difficult for the miniexon construct than for the IDX construct. This raises the possibility that the smallness of the miniexon limits the accessibility of the 5′SS to factors other than U1, i.e., factors that are necessary for splicing intron C2. On the other hand, the IDX construct was expressed throughout at a higher level than was the miniexon construct (Fig. 5B; compare lanes 1 and 8), and the difference in exon inclusion between lanes 5 and 12 becomes minor when this is taken into account. A shorter exposure is included in Fig. 5B to illustrate this suggestion. Thus, IDX inclusion may in fact be only slightly more efficient than is miniexon inclusion in the presence of U99, with the difference attributable to the twofold difference in 3′SS activity in the two situations. This argues that the defect in intron C2 splicing is not necessarily aggravated by the size of the miniexon.

To test this suggestion directly, we determined whether the mutated 5′SS sequence of the miniexon is accessible to a U6 snRNA with compensatory base changes, U6-αCTC (Fig. 5A) (wild-type U6 has three positions of complementarity to the mutated 5′SS sequence; U6-αCTC has six). We showed previously for pIDX:CTC that U6-αCTC increases the splicing of intron C2 in the presence of U99, which results in increased IDX inclusion (25); this was confirmed here (Fig. 5B; compare lanes 12 and 13). As indicated by the appearance of the 22-kDa + ME protein (Fig. 5B, lane 6), coexpression of the miniexon vector and U99 with U6-αCTC also promoted exon inclusion (compare to lane 5). Therefore, the 5′SS region of the miniexon is accessible to both U6 snRNA and other factors that are necessary for the recognition of the 5′SS and splicing of intron C2. When they are standardized to the 22-kDa bands (Fig. 5B, lanes 1 and 8), the exon inclusion bands (lanes 6 and 13) differ by no more than twofold in intensity, suggesting that intron C2 splicing is equally efficient in these two situations. Accordingly, the short distance between the 3′SS and the 5′SS of the miniexon does not appear to hinder recognition and processing of the 5′SS by factors other than U1.

It is noteworthy that intron C2 was spliced with equal efficiencies in the absence and presence of U6-αCTC when U1 bound to the mutated 5′SS of IDX (Fig. 5B, lanes 10 and 11), but not when U1 bound further downstream (lanes 12 and 13). We have proposed previously that U1 facilitates base pairing of U6 snRNA with pre-mRNA sequences at or near the U1 binding site, and we have provided evidence that endogenous U6 snRNA can associate with the mutated 5′SS of pIDX:CTC, although it does so inefficiently because of limited base complementarity (25). The results in Fig. 5B, lanes 10 through 13, are consistent with the interpretation that U1-αCTC, but not U99, directs endogenous U6 snRNA to the 5′SS efficiently (compare lanes 10 and 12) and that U6-αCTC, because of its increased complementarity to the 5′SS, is less dependent on guidance by U1. It is also possible that U99 misdirects endogenous U6 and, to a smaller degree, U6-αCTC to positions inside the intron, but this has not been shown. Either way, our results argue that the main 5′SS binding factor which limits exon size is U1, not U6 or some other factor, which suggests that the decision to splice or skip internal exons on the basis of size is made at an early stage.

DISCUSSION

SS signals in vertebrates are believed to be recognized initially in pairs across exons, but this is problematic when the two signals of a pair are too close together to accommodate the various factors involved in their recognition (1, 2, 10, 52). Indeed, a lower size limit for efficient splicing of vertebrate internal exons has been suggested by experimental and statistical results (1, 10, 21), but it is not certain that this limit reflects constraints at the stage of exon definition. Furthermore, although most vertebrate exons are larger than the suggested limit, there are multiple exceptions, but it is not generally understood how these exceptions circumvent the usual size constraints. To get to these issues, it is useful to define the conditions that normally prevent the efficient inclusion of small exons.

U1 snRNP is a large splicing factor that almost certainly plays an important role in exon definition (11, 23, 31). We previously found that the normal base pairing interaction of U1 with the 5′SS can be functionally replaced by base pairing away from the 5′SS, which stimulates the upstream 3′SS, as does the normal base pairing interaction, without changing the position of the 5′SS (9, 24). By using this observation, we asked whether the normal interaction of U1 with the 5′SS and the dependence of 3′SS activity on this interaction dictate a minimal distance between 5′ and 3′SSs across internal exons. The results show that stimulation of an upstream 3′SS by U1 targeted downstream requires a minimum distance between the 3′SS and U1, and this distance is very close to the lower size limit for efficient splicing of vertebrate internal exons. This suggests that the size constraint on vertebrate internal exons is due at least in part to steric hindrance between U1 snRNP at the 5′SS and factors that are responsible for recognition of the 3′SS.

The 5′SS is recognized by different factors at different stages of spliceosome assembly (for reviews, see references 4, 7, 29, 44, 47, and 60). U1 snRNP binds at a very early stage and is a component of the E (early) or commitment complex. E complex also contains SR proteins (51), and it has previously been suggested that one of these, SF2/ASF, binds the 5′SS directly (69) and interacts from this position with the U1 70K protein (26); it has not been demonstrated, however, that this actually occurs during normal spliceosome assembly. At subsequent stages, U6 snRNA and U5 snRNA interact with the 5′SS through base pairing (reviewed in reference 33), and protein factors such as p220, a U5 snRNP component, are found in close proximity to the 5′SS (45, 60, 64). It is not a priori obvious, therefore, at which stage and by which factors the proposed size constraint on internal exons is generated. Previous observations are consistent with the suggestion that U1 snRNP, perhaps together with associated factors, is at least partly responsible. For example, it has previously been shown that an exon sequence of 15 to 25 nt is protected from RNase digestion and oligonucleotide-mediated inhibition of splicing by a complex that is formed on the 5′SS (5, 12, 28, 48), and it has been demonstrated that U1 snRNP is a component of this 5′SS complex (5, 28). Therefore, U1 snRNP and associated factors may occupy up to half of the previously suggested minimal distance between the 5′SS and 3′SS, and it can be easily imagined that the other half is necessary to accommodate exon-bridging proteins together with factors that recognize the 3′SS. Our results support this possibility, suggesting that the decision between inclusion and exclusion of short internal exons occurs at an early stage.

Here we have shown that U1 can activate an upstream 3′SS from a position of approximately 50 nt or more downstream but cannot when the distance is smaller. With duplicated 3′SSs, the activity of the distal site increased rapidly when U1 was moved downstream through the 50-nt boundary, but activity started to decrease when the 50-nt boundary for the proximal site was reached (Fig. 4). From this point on, competition between the two sites was increasingly evident. To exclude the possibility that this pattern was a function of the position of U1 relative to a downstream point of reference, a sequence insertion just upstream of U1 binding sites was tested (Fig. 3). Although the insertion did not change the position of U1 relative to downstream sequences, the pattern of 3′SS activity shifted as if each U1 had been moved further downstream, showing that the distance between U1 and the upstream 3′SSs, not the proximity of U1 and any downstream element, was critical. This experiment also demonstrated that each U1, including U15, was capable of activating an upstream 3′SS as long as its binding site was more than 50 nt away. We conclude, therefore, that activation of an upstream 3′SS by U1 requires a minimal spacing of approximately 50 nt, which is very close to the previously suggested lower size limit for efficient inclusion of vertebrate internal exons (1, 10, 21). This is consistent with the idea that the lower size limit of internal exons is determined at least in part by the spatial requirements for interaction between U1 snRNP at the 5′SS and factors that assemble upstream at the 3′SS.

Our experiment with the 24-nt miniexon expanded these observations by showing that the miniexon could be included in spliced RNA on the condition that U1 did not bind the 5′SS but bound further downstream (Fig. 5). In this situation, enhanced binding of U6 snRNP to the 5′SS did not inhibit activation of the 3′SS 24 nt upstream, as suggested by the equally low levels of exon-skipping product (22-kDa protein) in the presence and absence of U6-αCTC; inhibition would have increased exon skipping. Conversely, the proximity of the 3′SS did not reduce the accessibility of the 5′SS to U6 or other factors that are required for processing of the 5′SS, since the U6-αCTC-dependent increases in exon inclusion for the miniexon and full-length IDX were similar. Therefore, it appears that U1, not some other factor(s) that normally interacts with the 5′SS, limits the proximity of 5′- and 3′SSs across internal exons. This limitation could be due to direct steric incompatibility between the presence of U1 at the 5′SS and factors that bind at the 3′SS or to problems in fitting the usual exon definition bridge over a short, perhaps inflexible, exon. Our results suggest that the proximity between SSs across exons is less restricted once the focus shifts from exon definition to the removal of flanking introns, perhaps because SS complexes at this later stage take less exon space, because exon flexibility and accommodation of exon-bridging complexes are no longer necessary, or both.

Activation of the upstream 3′SS by U99 was only twofold less efficient with the miniexon than with full-length IDX, which is a small difference, given that the U1-3′SS spacing for the miniexon was well below 50 nt; in our earlier experiments, activation was already significantly reduced at 48 nt (Fig. 3) and was essentially undetectable at 30 nt (Fig. 2). Since activation of the correct 3′SS was not measured directly in the miniexon experiment, it is possible that U99 activated aberrant 3′SSs further upstream, which would decrease the amount of the 22-kDa protein just like activation of the correct 3′ss. However, the production in the presence of U6-αCTC of the correct miniexon protein, 22 kDa + ME, at a level only twofold below that of the 22-kDa+ (full-length IDX) protein argues against this possibility; our experience with these proteins suggests that the 22 kDa + ME protein is not significantly more stable than is the 22-kDa+ protein. Instead, the results suggest that factors other than linear distance influence the ability of U1 to activate upstream 3′SSs. This is reasonable, since the real distance between U1 and the 3′SS is influenced by RNA secondary structure, which is sequence dependent. It can be imagined, for example, that the *Xho*I site at the center of the miniexon facilitates bending or folding of the exon so that U99 and the upstream 3′SS can be brought sufficiently close together to establish a connection between them without major friction; a different sequence may keep the RNA in a more extended shape, thereby increasing the distance that must be bridged. In this view, short exons are problematic for exon definition because they cannot be looped out, but this is of course only one of several possibilities that can be experimentally tested.

Although most vertebrate internal exons fall within the range of 50 to 300 nt, an estimated 4% are shorter, raising the

question of how these exons overcome the usual spacing constraints (1). The simplest way would be to select and process the termini of small exons sequentially, i.e., to circumvent traditional exon definition. An example of this approach has previously been described in a study of a constitutive, 7-nt troponin I microexon (52). In this case, the downstream SS of the microexon is recognized first as the end of a larger exon, including the upstream intron and exon, and the upstream SS appears to be recognized later. However, this may be a viable approach only if at least one of the flanking introns is sufficiently short to be included in the initial composite exon, which is not always the case; sequential splicing of larger introns has also been demonstrated with truncated substrates in vitro (18, 41), but it is not clear if this reflects the natural situation. It has previously been noted that small exons tend to be surrounded by strong SSs (1, 10), and skipping of an internally deleted exon was reversed by improving the flanking SSs (11). The flanking introns in this experiment were short, and it is possible that strong SSs with short flanking introns eliminate the need for exon definition altogether (56), thereby allowing sequential recognition of SSs across exons. However, strong SSs with short flanking introns are not universal features of small vertebrate exons (67), and it is likely that other mechanisms exist to either facilitate sequential recognition and processing of SSs across small exons or to somehow avoid steric problems during traditional exon definition. These mechanisms probably rely on special intron sequences.

Different intronic elements have previously been shown to enhance the inclusion of small exons (3, 6, 36, 67), but it is not clear how these elements work. In our system, various sequences targeted by U1 snRNAs, both exonic and intronic, behave as splicing enhancers that activate the upstream 3′SS. Conventional exonic splicing enhancers and normal 5′SSs also stimulate use of the upstream 3′SS, and the suggested mechanism is the same, establishment of communication with the 3′SS through bridging factors (4, 23, 44, 51, 61). It appears, therefore, that exonic and intronic splicing enhancers can act alike, suggesting that certain natural intronic splicing enhancers likewise function by providing anchoring sites for downstream factors that stimulate the 3′SS upstream. It is likely, although not certain, that conventional exonic splicing enhancers in internal exons communicate simultaneously with the upstream and downstream SSs (51), thereby supporting traditional exon definition, and intronic enhancers could do the same, for example, by first folding over the exon. For small exons, it is conceivable that such an arrangement directs the binding of U2AF and SR proteins to the two SSs but denies access to the region of a large particle such as U1 snRNP; mechanisms that involve steric hindrance to prevent the binding of U1 to the 5′SS have previously been described (13–15, 50). Alternatively, it is possible that intronic enhancers, including our U1 binding sites, uncouple the recognition of SSs across exons. In the troponin I situation referred to above, recognition of the upstream 3′SS is delayed until the downstream 5′SS has been recognized and processed (52). If enhancers selectively promote recognition of the 3′SS, they could accomplish the reverse, i.e., rapid occupation of the 3′SS to deny U1 access to the 5′SS until the upstream intron is spliced. Either way, our work suggests that intronic enhancers which promote the splicing of small exons must deal with the normal binding of U1 to the 5′SS by accelerating recognition of the 5′SS to imitate the troponin I situation or by shielding the 5′SS from U1, temporarily or permanently.

In summary, our results are consistent with previous suggestions that simultaneous recognition of the two ends of exons is problematic for short exons because of steric constraints. They suggest that U1 snRNP is at least partly responsible for this limitation and that any mechanism that can substitute functionally for the usual base pairing between U1 and the 5′SS should facilitate the splicing of small exons. Whether intronic enhancers of miniexon splicing act in this manner or by facilitating sequential recognition and processing of SSs is a question for future studies.

ACKNOWLEDGMENTS

We thank Christine Milcarek and Gretchen Edwalds-Gilbert for constructive discussions and Paula Grabowski for helpful comments on an early version of the manuscript.

This work was supported in part by grant CN-74 from the American Cancer Society.

REFERENCES

1. **Berget, S. M.** 1995. Exon recognition in vertebrate splicing. J. Biol. Chem. **270:**2411–2414.
2. **Black, D. L.** 1991. Does steric interference between splice sites block the splicing of a short c-*src* neuron-specific exon in non-neuronal cells? Genes Dev. **5:**389–402.
3. **Black, D. L.** 1992. Activation of c-*src* neuron-specific splicing by an unusual RNA element in vivo and in vitro. Cell **69:**795–807.
4. **Black, D. L.** 1995. Finding splice sites within a wilderness of RNA. RNA **1:**763–771.
5. **Black, D. L., B. Chabot, and J. A. Steitz.** 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell **42:**737–750.
6. **Carlo, T., D. A. Sterner, and S. M. Berget.** 1996. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. RNA **2:**342–353.
7. **Chabot, B.** 1996. Directing alternative splicing: cast and scenarios. Trends Genet. **12:**472–478.
8. **Cohen, J. B., S. D. Broz, and A. D. Levinson.** 1993. U1 small nuclear RNAs with altered specificity can be stably expressed in mammalian cells and promote permanent changes in pre-mRNA splicing. Mol. Cell. Biol. **13:**2666–2676.
9. **Cohen, J. B., J. E. Snow, S. D. Spencer, and A. D. Levinson.** 1994. Suppression of mammalian 5′ splice-site defects by U1 small nuclear RNAs from a distance. Proc. Natl. Acad. Sci. USA **91:**10470–10474.
10. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNAs with short internal exons. Mol. Cell. Biol. **11:**6075–6083.
11. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. **12:**2108–2114.
12. **Dominski, Z., and R. Kole.** 1994. Identification and characterization by antisense oligonucleotides of exon and intron sequences required for splicing. Mol. Cell. Biol. **14:**7445–7454.
13. **Eng, F. J., and J. R. Warner.** 1991. Structural basis for the regulation of splicing of a yeast messenger RNA. Cell **65:**797–804.
14. **Eperon, I. C., D. C. Ireland, R. A. Smith, A. Mayeda, and A. R. Krainer.** 1993. Pathways for selection of 5′ splice sites by U1 snRNPs and SF2/ASF. EMBO J. **12:**3607–3617.
15. **Eperon, L. P., I. R. Graham, A. D. Griffiths, and I. C. Eperon.** 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell **54:**393–401.
16. **Fu, X. D.** 1995. The superfamily of arginine/serine-rich splicing factors. RNA **1:**663–680.
17. **Ge, H., and J. L. Manley.** 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. Cell **62:**25–34.
18. **Grabowski, P. J., F. U. Nasim, H. C. Kuo, and R. Burch.** 1991. Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. **11:**5919–5928.
19. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7:**559–599.
20. **Hampson, R. K., L. La Follette, and F. M. Rottman.** 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. Mol. Cell. Biol. **9:**1604–1610.
21. **Hawkins, J. D.** 1988. A survey on intron and exon lengths. Nucleic Acids Res. **16:**9893–9908.
22. **Heinrichs, V., and B. S. Baker.** 1995. The *Drosophila* SR protein RBP1 contributes to the regulation of *doublesex* alternative splicing by recognizing RBP1 RNA target sequences. EMBO J. **14:**3987–4000.
23. **Hoffman, B. E., and P. J. Grabowski.** 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. Genes Dev. **6:**2554–2568.
24. **Hwang, D. Y., and J. B. Cohen.** 1996. Base pairing at the 5′ splice site with U1 small nuclear RNA promotes splicing of the upstream intron but may be

dispensable for splicing of the downstream intron. Mol. Cell. Biol. **16:**3012–3022.
25. **Hwang, D. Y., and J. B. Cohen.** 1996. U1 snRNA promotes the selection of nearby 5′ splice sites by U6 snRNA in mammalian cells. Genes Dev. **10:**338–350.
26. **Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Lührmann, M. A. Garcia-Blanco, and J. L. Manley.** 1994. Protein-protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. Nature **368:**119–124.
27. **Krainer, A. R., G. C. Conway, and D. Kozak.** 1990. The essential pre-mRNA splicing factor SF2 influences 5′ splice site selection by activating proximal sites. Cell **62:**35–42.
28. **Krämer, A.** 1987. Analysis of RNase-A-resistant regions of adenovirus 2 major late precursor-mRNA in splicing extracts reveals an ordered interaction of nuclear components with the substrate RNA. J. Mol. Biol. **196:**559–573.
29. **Krämer, A.** 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. Annu. Rev. Biochem. **65:**367–409.
30. **Krawczak, M., J. Reiss, and D. N. Cooper.** 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum. Genet. **90:**41–54.
31. **Kuo, H. C., F. H. Nasim, and P. J. Grabowski.** 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science **251:**1045–1050.
32. **Lavigueur, A., H. La Branche, A. R. Kornblihtt, and B. Chabot.** 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.
33. **Madhani, H. D., and C. Guthrie.** 1994. Dynamic RNA-RNA interactions in the spliceosome. Annu. Rev. Genet. **28:**1–26.
34. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. Genes Dev. **10:**1569–1579.
35. **Marvit, J., A. G. DiLella, K. Brayton, F. D. Ledley, K. J. H. Robson, and S. L. C. Woo.** 1987. GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylketonuria. Nucleic Acids Res. **15:**5613–5628.
36. **Min, H., R. C. Chan, and D. L. Black.** 1995. The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. Genes Dev. **9:**2659–2671.
37. **Mitchell, P. J., G. Urlaub, and L. Chasin.** 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in Chinese hamster ovary cells. Mol. Cell. Biol. **6:**1926–1935.
38. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to mRNA by the spliceosome, p. 303–357. *In* R. F. Gesteland and J. F. Atkins (ed.), The RNA world. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
39. **Nagoshi, R. N., and B. S. Baker.** 1990. Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns. Genes Dev. **4:**89–97.
40. **Nakai, K., and H. Sakamoto.** 1994. Construction of a novel database containing aberrant splicing mutations of mammalian genes. Gene **141:**171–177.
41. **Nasim, F. H., P. A. Spears, H. M. Hoffmann, H. C. Kuo, and P. J. Grabowski.** 1990. A sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5′ splice site in preprotachykinin pre-mRNA. Genes Dev. **4:**1172–1184.
42. **Nelson, K. K., and M. R. Green.** 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. Genes Dev. **2:**319–329.
43. **Ramchatesingh, J., A. M. Zahler, K. M. Neugebauer, M. B. Roth, and T. A. Cooper.** 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.
44. **Reed, R.** 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. Curr. Opin. Genet. Dev. **6:**215–220.
45. **Reyes, J. L., P. Kois, B. B. Konforti, and M. M. Konarska.** 1996. The canonical GU dinucleotide at the 5′ splice site is recognized by p220 of the U5 snRNP within the spliceosome. RNA **2:**213–225.
46. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10:**84–94.
47. **Rosbash, M., and B. Séraphin.** 1991. Who's on first? The U1 snRNP-5′ splice site interaction and splicing. Trends Biochem. Sci. **16:**187–190.
48. **Ruskin, B., and M. R. Green.** 1985. Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. Cell **43:**131–142.
49. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell **52:**207–219.
50. **Siebel, C. W., L. D. Fresco, and D. C. Rio.** 1992. The mechanism of somatic inhibition of *Drosophila* P-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5′ splice site control U1 snRNP binding. Genes Dev. **6:**1386–1401.
51. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol. **14:**7670–7682.
52. **Sterner, D. A., and S. M. Berget.** 1993. In vivo recognition of a vertebrate mini-exon as an exon-intron-exon unit. Mol. Cell. Biol. **13:**2677–2687.
53. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. **7:**2598–2608.
54. **Tacke, R., and J. L. Manley.** 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. EMBO J. **14:**3540–3551.
55. **Talerico, M., and S. M. Berget.** 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. **10:**6299–6305.
56. **Talerico, M., and S. M. Berget.** 1994. Intron definition in splicing of small *Drosophila* introns. Mol. Cell. Biol. **14:**3434–3445.
57. **Tian, M., and T. Maniatis.** 1993. A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. Cell **74:**105–114.
58. **Treisman, R., S. H. Orkin, and T. Maniatis.** 1983. Specific transcription and RNA splicing defects in five cloned β-thalassaemia genes. Nature **302:**591–596.
59. **Treisman, R., N. J. Proudfoot, M. Shander, and T. Maniatis.** 1982. A single-base change at a splice site in a β^0-thalassemic gene causes abnormal RNA splicing. Cell **29:**903–911.
60. **Umen, J. G., and C. Guthrie.** 1995. The second catalytic step of pre-mRNA splicing. RNA **1:**869–885.
61. **Wang, Z., H. M. Hoffmann, and P. J. Grabowski.** 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. RNA **1:**21–35.
62. **Watakabe, A., K. Tanaka, and Y. Shimura.** 1993. The role of exon sequences in splice site selection. Genes Dev. **7:**407–418.
63. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75:**1061–1070.
64. **Wyatt, J. R., E. J. Sontheimer, and J. A. Steitz.** 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. Genes. Dev. **6:**2542–2553.
65. **Xu, R., J. Teng, and T. A. Cooper.** 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. Mol. Cell. Biol. **13:**3660–3674.
66. **Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth.** 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. Science **260:**219–222.
67. **Zhang, L., M. Ashiya, T. G. Sherman, and P. J. Grabowski.** 1996. Essential nucleotides direct neuron-specific splicing of γ2 pre-mRNA. RNA **2:**682–698.
68. **Zuo, P., and T. Maniatis.** 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes Dev. **10:**1356–1368.
69. **Zuo, P., and J. L. Manley.** 1994. The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5′ splice sites. Proc. Natl. Acad. Sci. USA **91:**3363–3367.