How small is an exon – does size matter?

C.G. Simpson, G. Clark, J. Watters & J.W.S. Brown

The removal of non-coding introns from pre-mRNA and the splicing together of translatable exon sequence is an essential requirement of gene expression and one area of post-transcriptional processing where gene expression can be regulated. The process of pre-mRNA splicing is mediated by a RNA and protein complex, called the spliceosome, that assembles on the pre-mRNA before removal of the intron takes place. Commitment to splicing requires the recognition of splice sites by factors prior to assembly of the major spliceosomal components. The accurate selection of these splice sites, in a pre-mRNA background that may contain many potential splice sites, is essential to obtain the correct splicing of exons. Two mechanisms of splice site choice, intron and exon definition, have been identified in eukaryotes that allow the accurate selection of splice sites during splicing. In exon definition, splice sites on either side of the exon are recognised by interactions between factors at the 5' splice site and the 3' end of the upstream intron. Because of the smaller size of plant introns, where only ~10% of introns are greater than 600 nt, exon definition was not expected to occur in plant splice site selection. However, we have, firstly, identified a number of splice site mutants that lead to exon skipping which is highly indicative of exon defining interactions operating in plants. Secondly, we have shown that, when two introns are found on the same transcript, one intron can enhance the splicing efficiency of the other, supporting the idea of interactions between the introns across the exon (Ann. Rep. 1996/97, 102-103). It is clear, therefore, that exon bridging interactions have an important role to play in plant splicing.

For exon defining interactions to occur, there is a minimum exon length of ~50 nt required before hindrance between the factors at the splice sites limits any interaction (Fig. 1). Despite this limitation, there are a number of examples of exons that fall below this minimal exon length, called mini-exons. In vertebrates, additional signals, in particular intron splicing enhancers, are required to promote the inclusion of a mini-exon. We have previously described the constitutive inclusion of a conserved 9 nt mini-exon in potato invertase mRNA transcripts. The elements required for the inclusion were found in the upstream intron (Ann. Rep. 1997/98, 71-73). Inclusion of this mini-exon has been studied further and detailed mutational analysis highlights the importance of a branchpoint consensus and an associated U_{11} sequence in the 3' end of the upstream intron. An

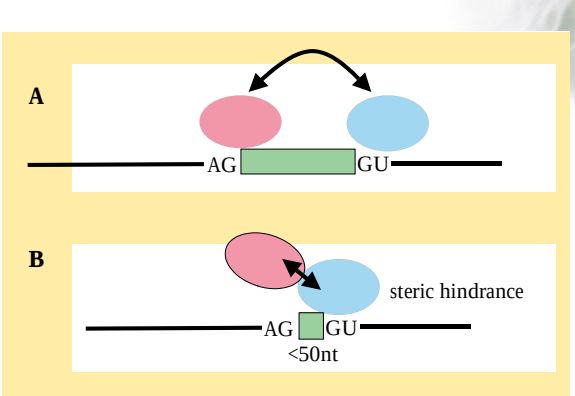

Figure 1 Steric hindrance in mini-exons
A) When an exon (box) is greater than 50nt, factors at the 3' and 5' splice sites (AG and GU respectively) are able to interact through exon definition. B) When exons are less than 50nt, factors are unable to interact and exon definition is inhibited through steric hindrance.

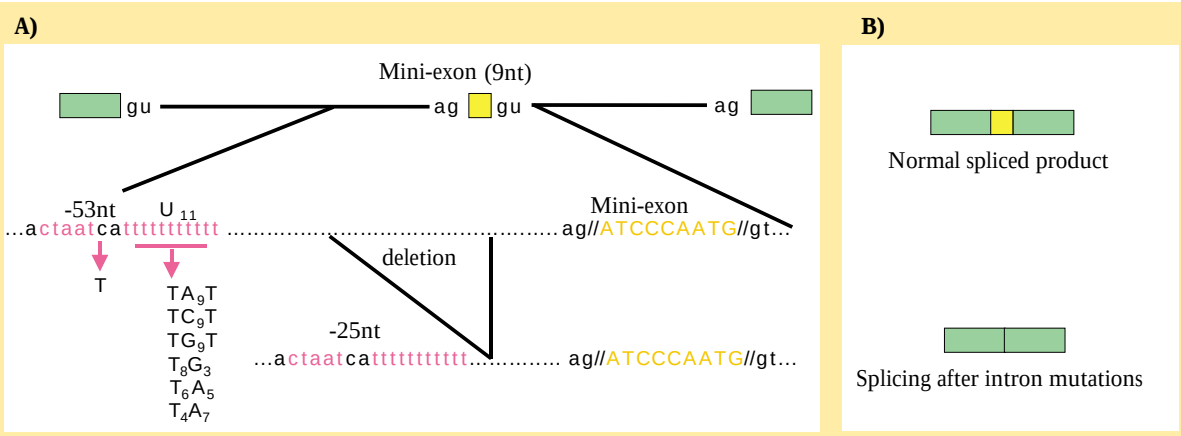

Figure 2 Detailed mutational analysis of invertase mini-exon splicing.
A) The structure of the mini-exon sequence is shown along with mutations to the branchpoint positioned –53nt from the 3' splice site and to the U_{11} element. A deletion between U_{11} and the 3' splice site brought the –53nt branchpoint and its associated U_{11} sequence to a position –25nt from the 3' splice site (boxed sequence). B) Splicing analysis of all these mutants in tobacco protoplasts led to skipping or exclusion of the 9nt mini-exon, while the normal transcript included the mini-exon.

A→U mutation to the branchpoint –53 nt upstream upstream from the 3' splice site skipped the mini-exon (Fig. 2). This shows that the upstream branchpoint nucleotide is essential for inclusion of the mini-exon. A U_{11} sequence 3 nt downstream of the putative branchpoint was mutated to a string of A, C or G nucleotides, and the number of Us were reduced to 8, 6 and 4 (Fig. 2). Virtually all mutations lead to non-inclusion of the mini-exon. This shows the essential nature of the U_{11} element for splicing of the mini-exon and a requirement for this element to be U-rich.

A striking feature of the branchpoint and U_{11} is its distance from the 3' splice site. This arrangement of branchpoint/U-rich region located at an extended distance (>25nt) from the 3' splice site is conserved among plant invertase genes. Reduction of this distance, such that the putative branchpoint was now 25 nt upstream from the 3' splice site, led to skipping of the mini-exon, indicating a need for the branchpoint to be further upstream than normal (Fig. 2). Thus, inclusion of the mini-exon appears to involve a bridging interaction between factors at the branchpoint/U-rich region and the 5' splice site flanking the mini-exon.

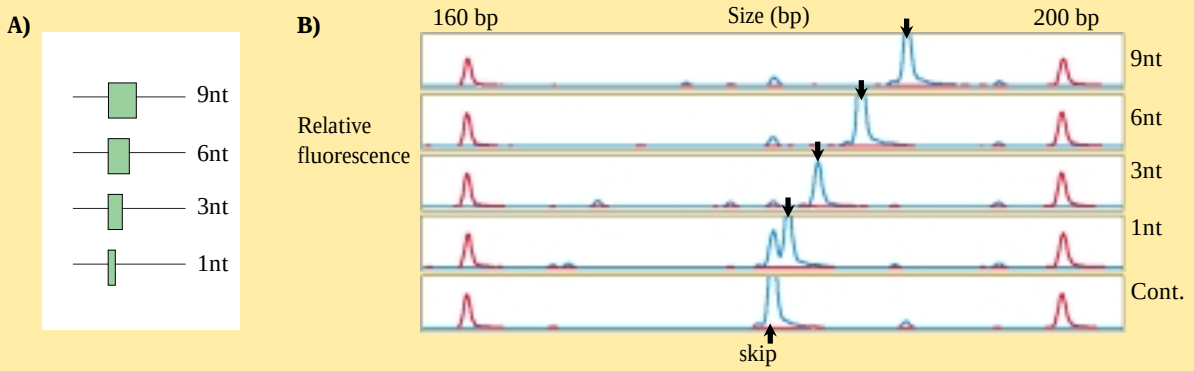

Figure 3 Splicing analysis of shortened mini-exon sequence.
A. The mini-exon sequence shown in Figure 2 was reduced in length from 9nt to 6, 3 and 1nt. The boxes represent the mini-exon.
B. RT-PCR analysis of the shortened mini-exons was performed using a fluorescent primer and the products separated on a gel. Bands found on the gel are represented here as a spectrogram of relative fluorescence against the size in base pairs (bp) of the RT-PCR product. Size marker peaks are above the 160bp and 200bp markers. The position of an RT-PCR product that represents a skipped product of 180bp is indicated. Inclusion of the differently sized mini-exons in the RT-PCR products are shown by arrows above the relevant peaks (189nt, 186nt, 183nt and 181nt for the 9, 6, 3 and 1nt exon respectively).

That the 9 nt exon is included in all invertase mRNAs suggests that the signals which we have defined represent very strong splicing signals. Therefore, it is possible that the sequence of the mini-exon is not important for its inclusion and that the signals could splice other exon sequences. To address this, the mini-exon was reduced in length from 9 nt to 6, 3 and even 1 nt in length. The 6 and 3 nt mini-exons were accurately and efficiently included while two-thirds of the final mRNA transcripts contained the 1nt mini-exon (Fig. 3). Efficient inclusion of a 1 nt exon sequence is quite exceptional and highlights the strength of the mini-exon signals. It also shows that the mini-exon sequence is irrelevant for splicing and that, in this system at least, the small size of the exon is not important. The challenge is to determine whether these signals can act as splicing enhancers to manipulate splicing and regulate expression. At present, we are looking for proteins that interact with the pre-mRNA elements involved in invertase mini-exon inclusion and are investigating other naturally-occurring plant mini-exon splicing systems.