The EMBO Journal vol.12 no.3 pp.1249 – 1255, 1993

Conversion of a *trans*-spliced *C.elegans* gene into a conventional gene by introduction of a splice donor site

Richard Conrad, Ruey Fen Liou[1] and Thomas Blumenthal[2]

Department of Biology, Indiana University, Bloomington, IN 47405, USA
[1]Present address: PO Box 23-106, Taiwan 10098, ROC
[2]Corresponding author

Communicated by J.E.Sulston

In *Caenorhabditis elegans*, pre-mRNAs that are *trans*-spliced are distinguished by the presence of an 'outron', intron-like RNA at the 5′ end followed by a splice acceptor. We report that *trans*-splicing of the *rol-6* gene can be completely suppressed simply by introducing a donor site into its 173 nt outron, at a site 50 nt upstream of the *trans*-splice site, thereby converting *rol-6* into a conventional gene with a spliced intron near its 5′ end. When the consensus donor site was inserted at sites further upstream it was less effective in replacing *trans*-splicing with *cis*-splicing. Surprisingly, the length of the intron was not the important variable, since lengthening of the 50 nt intron to 250 nt did not restore *trans*-splicing. Apparently the context into which the splice site was introduced determined the efficiency of its use. These results support the conclusion that the sole signal for *trans*-splicing is the presence of an outron. Clearly, *cis*- and *trans*-splice acceptor sites are interchangeable, allowing the possibility of competition between the two types of splicing.

Key words: mRNA processing/nuclear mRNA splicing/nematode/alternative splicing

Introduction

The phenomenon of intermolecular RNA splicing, called *trans*-splicing, was discovered in trypanosomes (Murphy *et al.*, 1986; Sutton and Boothroyd, 1986) and has since been reported in nematodes (Krause and Hirsh, 1987), trematodes (Rajkovic *et al.*, 1990) and *Euglena* (Keller *et al.*, 1992). In all of these cases a small untranslated leader, referred to as a spliced leader or SL, is spliced onto many recipient messages (reviewed in Nilsen, 1989; Agabian, 1990; Donelson and Zeng, 1990; Green, 1991; Huang and Hirsh, 1992). The donor of the SL is a small (~100 nt) RNA called the spliced leader RNA (SL RNA), which is packaged into a small nuclear ribonucleoprotein particle, the SL snRNP. This particle is closely related to the U-snRNPs involved in the assembly of the spliceosome, which is involved in the intramolecular or *cis*-, splicing process (Thomas *et al.*, 1988; Van Doren and Hirsh, 1988). The mechanism of *trans*-splicing is similar to *cis*-splicing: a branched intermediate analogous to the lariat intermediate of *cis*-splicing has been found both in trypanosomes (Murphy *et al.*, 1986; Sutton and Boothroyd, 1986) and in nematodes (Thomas *et al.*, 1988; Bektesh and Hirsh, 1988). There are also profound differences between the *trans*-splicing processes in trypanosomes and nematodes. In trypanosomes there are no introns in the transcripts and all messages are *trans*-spliced. In nematodes, introns are common and not all transcripts are *trans*-spliced, therefore two decisions must be made that are not made in trypanosomes: whether or not to *trans*-splice near the 5′ end of the pre-mRNA and whether to *trans*- or *cis*-splice at each acceptor site.

Many nematode genes are of the conventional variety in that transcription initiates at the 5′ end of the first exon. However, initiation of transcription of some nematode genes occurs within a region of the pre-mRNA that resembles an intron. We have termed this 5′ intron an 'outron' and shown that the presence of an outron, which is spliced out, is sufficient to specify *trans*-splicing (Conrad *et al.*, 1991). *Caenorhabditis elegans* has the added complication of a second SL RNA, SL2 RNA (the original is now known as SL1 RNA). Each SL RNA contributes its leader to the pre-mRNAs of a specific set of genes (Huang and Hirsh, 1989). *Cis*- and *trans*-splicing in *C.elegans* share many similarities. The SL RNA is bound by the Sm proteins, which are also bound to the U snRNPs (Thomas *et al.*, 1988; Van Doren and Hirsh, 1988). At least three of these U-snRNPs, U2, U4 and U6, have been shown to be required for both *cis*- and *trans*- splicing in an *Ascaris in vitro* system (Hannon *et al.*, 1991). The *trans*-splice donor sites are AG/GUAAA in SL1 RNA (Krause and Hirsh, 1987) and AG/GUACG in SL2 RNA (Huang and Hirsh, 1989), good matches to the AG/GUAAG consensus for 5′ *cis*-splice sites in *C.elegans* (Fields, 1990). The intron 3′ splice site consensus sequence (Fields, 1990) is the same as the *trans*-splice acceptor site consensus (T.Blumenthal, unpublished observations) and an intron 3′ splice site is competent to act as a *trans*-splice acceptor if placed in the proper context (Conrad *et al.*, 1991).

Previous work in our lab has been directed at determining the signals that promote *trans*-splicing. Our experiments have suggested that no specific sequence is required. When we transplanted an intron lacking a splice donor sequence into the 5′ untranslated region of a conventional gene, its product was *trans*-spliced to SL1 (Conrad *et al.*, 1991). This result suggested the simple model that the presence of an intron-like sequence at the 5′ end of a transcript, terminated by any splice acceptor site, is a sufficient signal to specify *trans*-splicing to SL1. A prediction of this model is that a normally *trans*-spliced gene could be converted into a *cis*-spliced gene by introduction of an exon or a splice donor site upstream of the splice acceptor site. In this paper evidence is presented demonstrating that introduction of a consensus splice donor sequence into the outron of a collagen gene, *rol-6* (Kramer *et al.*, 1990), does indeed convert it into a functional conventional gene. Interestingly, the introduced donor site was less effective at some locations than at others, indicating the importance of sequence context. In some cases both

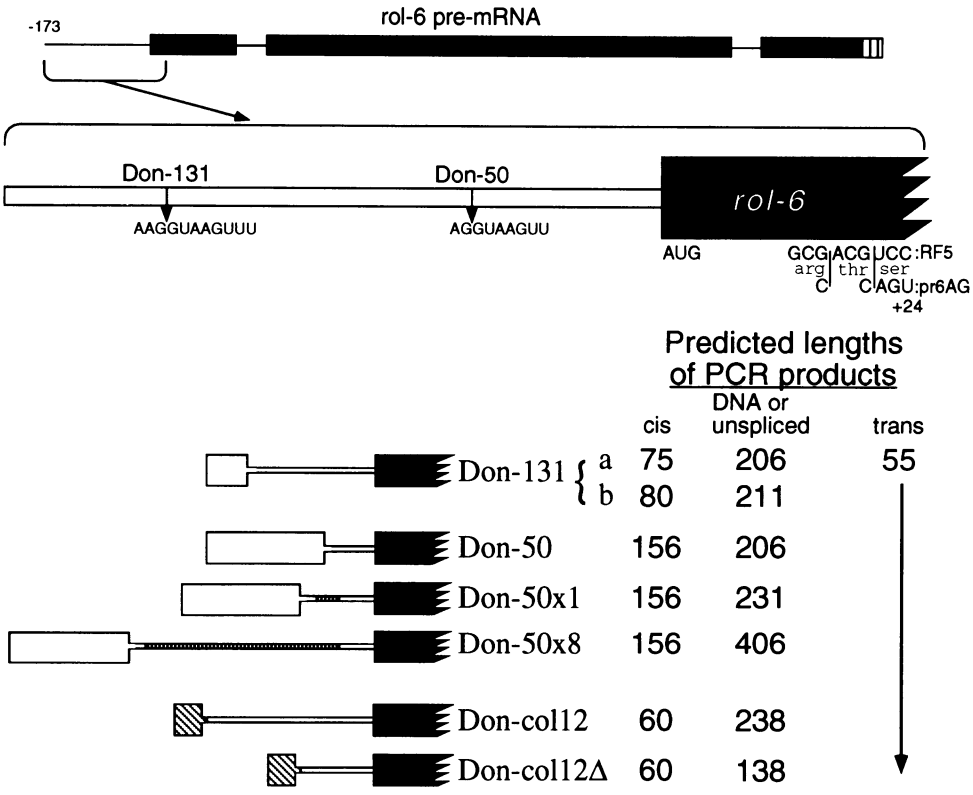

Fig. 1. Summary of test genes and PCR analysis. A schematic representation of the *rol-6* transcript is shown at the top. This gene has a 173 nt outron; two introns of 51 and 50 nt; and three exons of 141, 786 and 153 nt. The bracketed area at the 5′ end of the transcript is shown enlarged underneath. The region shown represents the entire outron plus 24 nt of the first exon/coding region. In the exon are shown the translational start and the 5 nt changes made to create pr6AG from pRF5. The (preserved) amino acid identities in the region of the base changes are also shown. In the outron region of the enlargement are shown the changes made to create the mutations Don-50 and Don-131a. The numbers indicate the positions of the introduced splice sites. The bottom of the figure schematically depicts the changes to Don-50 and Don-col12 to create Don-50×1, Don-50×8 and Don-col12Δ. The vertical cross-hatching represents the AT-rich oligonucleotide-duplex inserts and the angled cross-hatching regions of transplanted col-12 sequence. The diagrams are all drawn to scale. To the right is a table presenting the predicted lengths of the PCR products from each construct. For both cDNA synthesis and PCR, the downstream primer was specific for the region modified in the pRF5 to pr6AG conversion; it is referred to as the r6AG primer. Three upstream primers were used, one to amplify *trans*-spliced products and two to amplify *cis*-spliced (and unspliced or DNA) products. The first is equivalent to the 5′ 20 nt of the SL1 spliced leader (SL1−20). The latter two are either equivalent to the 5′ 20 nt of the *rol-6* transcript (r6−5′e), or to the 5′ 20 nt of the *col-12* transcript (Dc12−5′e). Note that the *trans*-spliced products are all 55 nt long.

processes occurred, creating a competition between *cis*- and *trans*-splicing.

Results

Insertion of an exon upstream of the rol-6 outron

Previously we reported the conversion of a conventional gene into a *trans*-spliced gene simply by providing it with an artificial outron (Conrad *et al.*, 1991). This experiment demonstrated that a splice acceptor site, which normally functions in *cis*-splicing, could act as a *trans*-splice acceptor, and suggested it might be possible to convert a *trans*-spliced gene into one that is not, simply by providing an upstream splice donor. The gene we selected for these experiments is one widely used as a co-transformation marker for *C.elegans*, the *rol-6* collagen gene (Kramer *et al.*, 1990). The pre-mRNA of this gene is *trans*-spliced to SL1 at a site 173 nt downstream of its transcription initiation site. Our strategy was to insert potential splice donors between the transcription inititiation site and the *trans*-splice site in an attempt to replace *trans*-splicing with *cis*-splicing of the introduced site to the *trans*-splice site.

A dominant *rol-6* allele, *su1006* (Mello *et al.*, 1991), which confers an easily identifiable rolling phenotype in transformed worm progeny, was used as the test gene. The plasmid pRF5, which contains the *rol-6* gene on a 2.2 kb *Hin*dIII fragment, produces rolling transformants at high frequency. The endogenous gene contains an amber mutation, so most of its RNA product is degraded, but since any residual RNA would interfere with our analysis, pRF5 was modified to produce pr6AG, containing a 5 nt change (see Figure 1) to provide a specific priming region for reverse transcription (RT) PCR experiments (data not shown).

As an initial test of the hypothesis, a 27 nt untranslated exon from a different collagen gene, *col-12*, (Park and Kramer, 1990) was inserted at the 5′ end of *rol-6*. This exon plus the consensus 5′ splice site (/GTAAG) was inserted precisely at the 5′ end of the *rol-6* outron, creating Don-col12. This modified gene was injected into the gonadal syncytium and rolling transformants were selected, indicating gene function was retained. Genomic Southern blots of two strains showed approximately equal amounts of the transgene (data not shown).

To determine the fate of the transgene RNA in each transformant, PCR amplification was performed using single-stranded cDNA made by extending a transgene-specific primer on poly(A) RNA prepared from mixed-stage

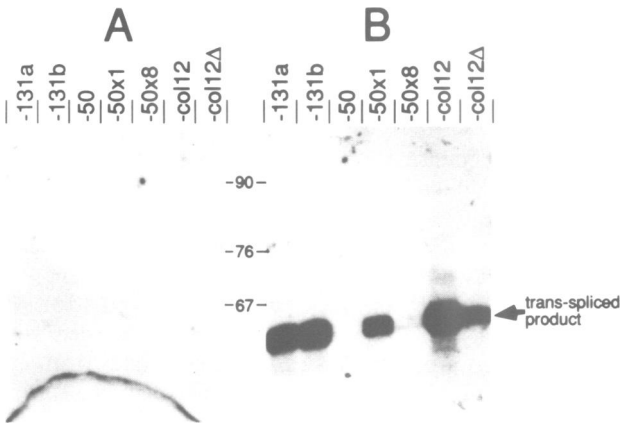

Fig. 2. Analysis of *trans*-spliced products by PCR. RNA isolated from strains transformed with the constructs shown in Figure 1 was reverse transcribed and PCR was performed as described in Materials and methods. Each lane contains the PCR products from RNA from a different transformed strain. PCR products were electrophoresed on denaturing 10% polyacrylamide gels and electroblotted as described, then probed with an end-labelled oligonucleotide equivalent to positions 1−17 of *rol-6*. Reactions were performed with the common downstream primer (r6AG) and the SL1−20 upstream primer, yielding products from *trans*-spliced RNAs. **A.** Shows the PCRs using the control annealing reactions (RNase added instead of reverse transcriptase). **B.** Shows the PCRs using the cDNA reactions—any new bands showing up on this side are RNA-dependent. All *trans*-spliced products are the same expected size and no DNA signal is expected.

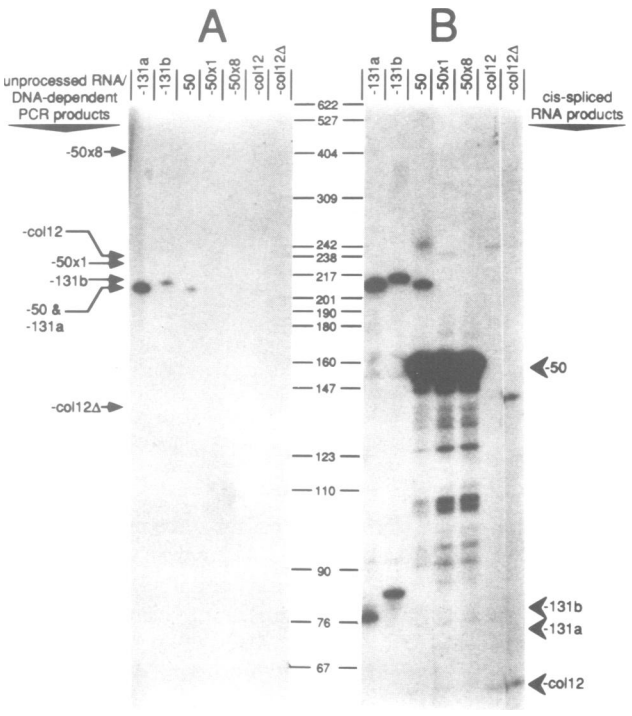

Fig. 3. Analysis of *cis*-spliced products by PCR. Analysis and presentation are the same as in Figure 2, except the denaturing gel was 6% polyacrylamide and the products shown are from PCR using the 5′ end primers (both for *rol-6*, r6-5′e, and *col-12*, Dc12-5′e) in conjunction with the r6AG primer. **A.** Shows the PCRs using the control annealing reactions (RNase added instead of reverse transcriptase). The bands expected from amplification of unprocessed RNA or contaminating DNA are indicated on the left side. **B.** Shows the PCRs using the cDNA reactions—all expected positions for processed RNA-dependent bands are indicated on the right side.

worm populations. Two possible 5′ ends might be found on the processed mRNA: either SL1, the spliced leader normally found on *rol-6* mRNA or the 5′ end of the initial transcript, if no processing or *cis*-splicing had occurred. For Don-col12, this would be the same as the 5′ end of *col-12* mRNA. To determine which splicing events had occurred, two separate PCRs were performed with the cDNA, one with an upstream primer equivalent to the SL1 sequence and one with the upstream primer equivalent to the first 20 nt of *col-12*. The expected products are shown in Figure 1. The *trans*-spliced product should be 55 nt long and dependent on SL1 primer. PCR products for *cis*-spliced RNA should be 60 nt for Don-col12. Unspliced RNA products will give the same size products as would contaminating DNA and this size is also shown in Figure 1.

The results of the PCR experiments are shown in Figure 2 for *trans*-spliced product and Figure 3 for *cis*-spliced and unspliced products. Bands present in both panels of Figure 3 presumably represent products resulting from copying DNA contaminants. Clearly in the case of Don-col12, the presence of the upstream exon has failed to suppress *trans*-splicing, as the expected *trans*-spliced product is present in abundance. One possible explanation for this result is that the *col-12* exon is not transcribed. However, this does not appear to be the case, since a product of the size expected for the unspliced pre-mRNA is present at a low level, and only when reverse transcriptase is present (Figure 3). Furthermore, a small amount of *cis*-spliced product can be seen at the appropriate location (Figure 3B) so at least some of the mRNA must contain the introduced exon. We conclude that the presence of the *col-12* exon at the 5′ end of the *rol-6* outron results in a competition between the two processes and that *trans*-splicing is still highly favored.

One reason why *trans*-splicing is still favored with Don-col12 pre-mRNA might be that a small exon (27 bp) followed by a relatively large intron (178 bp) is an unfavorable context for *cis*-splicing, so *trans*-splicing is able to compete effectively with it. To test this idea we deleted 100 bp of the outron−intron to create a 78 bp potential intron (called Don-col12Δ). As can be seen in Figures 2 and 3, this change did not have a dramatic effect: *trans*-splicing still predominated. However, the level of *trans*-spliced product did decrease and the level of *cis*-spliced product did increase somewhat. Thus, as expected, *cis*-splicing became a better competitor when the intron was made smaller, closer to the length of the typical 50 nt *C.elegans* intron.

Introduction of a splice donor site within the outron

In order to test whether *cis*-splicing competes with *trans*-splicing more effectively if we simply insert a match to the *C.elegans* splice donor consensus within the outron, we modified pr6AG by replacing 10 nt to produce a consensus *C.elegans* splice donor site, AAG/GUAAGUU (Fields, 1990), at a site 131 nt upstream of the *trans*-splice site (Don-131a). This site was chosen because introns are more AU-rich than exons in *C.elegans* (Blumenthal and Thomas, 1988) and this splice site is inserted in a region of a transitional shift, in which a lower A + U region exists upstream of a higher A + U region. We found that most of the pre-mRNA was *trans*-spliced (Figure 2), although we did observe some *cis*-splicing between the introduced donor site and the *trans*-splice acceptor (Figure 3), demonstrating

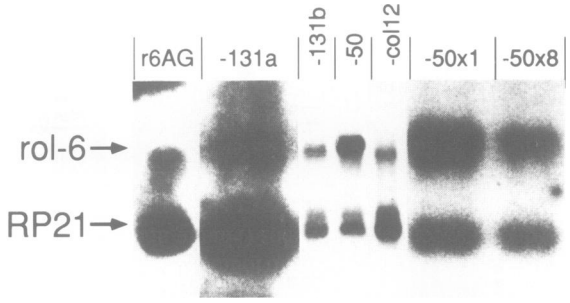

Fig. 4. Accumulation of rol-6 mRNA in transgenic strains. Poly(A)$^+$ RNA was electrophoresed on denaturing (formaldehyde) agarose gels and capillary-blotted onto nylon membranes as described. The Figure represents a composite of several different sized gels. All blots were probed together with a mixture of *rol-6* and *rp21* probes, as described in Materials and methods.

that this is a functional intron, but that for some reason *trans*-splicing still competes effectively. A possible explanation for this result is that *cis*-splicing did occur at a higher level, but the *cis*-spliced product was unstable. Such instability might be caused by failure to translate the correct reading frame due to the presence of an upstream, out-of-frame intitiation codon present just 5′ of the introduced donor site. To test this possibility, we destroyed the AUG in Don-131a by site-directed mutagenesis to create Don131b, which has also had its 5′ 'exon' extended by 5 nt relative to Don-131a. However this alteration made no perceptible difference in the ratio of *trans*- to *cis*-spliced products (Figures 2 and 3). Thus we conclude that a donor site introduced 131 nt upstream of the *trans*-splice site competes only inefficiently with *trans*-splicing.

The experiment with Don-col12Δ suggested that intron length might be an important variable, so we introduced a splice donor consensus sequence at a site only 50 bp upstream of the *trans*-splice acceptor. Don-50 creates the optimal size *C.elegans* intron and it is also at a transitional boundary between upstream non-A + U rich RNA and downstream A + U-rich RNA. The results, shown in Figures 2 and 3 demonstrate that introduction of the splice donor at this site completely suppressed *trans*-splicing, replacing it with efficient *cis*-splicing between the introduced donor and the *trans*-splice site. We also tested for *trans*-splicing to SL2 with this and all other constructs, and in no case was any SL2 splicing observed (data not shown), eliminating the possibility that the specificity was changed. These results along with our earlier findings (Conrad *et al.*, 1991) allow us to conclude that *trans*-splice sites and 3′ *cis*-splice sites are functionally equivalent. They demonstrate that the important variable that determines if a pre-mRNA is *trans*-spliced is related to whether a suitable donor site is recognized upstream of an acceptor site. Clearly the presence of an upstream donor is capable of converting a *trans*-spliced gene into a conventional gene.

Intron length versus splice site context

Why is the splice site at −50 recognized efficiently while the identical site at −131 is mostly ignored? In order to test the possibility that intron length is the key variable, we increased the length of the intron in Don-50. Copies of an oligonucleotide cassette, rich in A + T to simulate intron, were inserted into a *Xho*I site created at position −35/−36. To avoid secondary structure, one strand was made rich in As, the other in Ts, by using an A_4T_2 (or T_4A_2) motif. Two extensions, called Don-50×1 (extended by one 25 nt insert) and Don-50×8 (extended by eight inserts) were tested. Both are predominantly A-rich (only the 3′-most insert in Don-50×8 is T-rich). The data clearly show that both of these constructs are *cis*-spliced at high levels (Figure 3). Indeed *trans*-splicing is undetectable even when the intron is 250 nt long (Figure 2). Surprisingly, the construct with the 75 nt intron is *trans*-spliced at a detectable level (Figure 2). However, in a primer extension experiment with poly(A) RNA from both Don-50 and Don-50×1 (data not shown), we saw only single bands with a size consistent with the *cis*-spliced product. This result demonstrates that very little *trans*-spliced product has accumulated in these cases and also that no major products other than the ones we see by PCR were made. This eliminates the possibility of alternative transcription start sites that make *cis*-spliced or unspliced RNAs. We conclude that intron length *per se* does not affect the efficiency of *cis*-splicing, at least in cases such as this where other factors involved in generating a functional intron appear to be optimal.

RNA accumulation

To determine if *cis*-splicing of the 5′ end of a normally *trans*-spliced gene had any effect on the stability of the transgene RNA, poly(A) RNA from all of the transformants was subjected to Northern analysis. In addition to a *rol-6*-specific probe, a probe for the *rp21* ribosomal protein gene was hybridized to the blot, as a loading control. To ensure that any differences were not due simply to a biased population, two separate RNA preparations of mixed-stage populations from the four primary transformants were made and one of each is shown in Figure 4. It is clear that RNA accumulation levels are not reduced by the lack of *trans*-splicing. Indeed there is some suggestion that RNA levels in Don-50 and its derivatives, whose products are largely *cis*-spliced, are even higher than the others, whose products are largely *trans*-spliced.

Discussion

We demonstrated previously that a *cis*-splice acceptor can act as a *trans*-splice acceptor if moved into the proper context. Here we have demonstrated the converse: a *trans*-splice acceptor site can function efficiently as a *cis*-splice acceptor when it is linked to a splice donor in the proper context. Additionally, our results demonstrate that part of an outron can be used as a 5′ exon, replacing the spliced leader, without obviously interfering with gene function. Clearly a splice acceptor near the 5′ end of the pre-mRNA can function either as a *trans*-splice acceptor or a *cis*-splice acceptor in *C.elegans*, depending on sequence context of potential splice donors and possibly other sequences within the intron as well. Most of our constructs resulted in competition between *cis*- and *trans*-splicing. Some recent results suggest that such competition may occur with pre-mRNA products of endogenous *C.elegans* genes as well (Nonet and Meyer, 1991; W.Li, R.K.Herman and J.Shaw, personal communication).

Poor splicing of the col-12 exon in an artificial construct

We expected the 27 nt non-coding *col-12* 5′ exon to function well when placed upstream of the *rol-6* outron at the

transcription start site, because this exon is from a member of the same gene family (collagens). Hence it was surprising that it was recognized and spliced at what is normally the *rol-6 trans*-splice site only at a very low level. (It should be noted that the possibility that the *col-12* exon is not always transcribed has not been eliminated, although its presence 30 bp 3′ of the *rol-6* TATA box makes this improbable.) It seems reasonable that *cis*-splicing might be favored over *trans*-splicing, since the latter requires two substrate molecules. Thus the dominance of *trans*-splicing in the Don-col12 construct may indicate the presence of an impediment to *cis*-splicing. Since splice site choice is still not well understood, we can conclude only that the *rol-6 trans*-splice site and the *col-12* splice donor site functioned together very poorly. Initially we thought that intron length might be the primary determinant, since there is a strong tendency for *C.elegans* introns to be quite short, ~50 nt (Blumenthal and Thomas, 1988). However, when we deleted 100 nt of the *rol-6* outron to create Don-col12Δ, only a small increase in the ratio of *cis*- to *trans*-spliced product was observed. This increase does suggest, however, that intron length can influence the competition between *cis* and *trans*-splicing.

Switching from trans- to cis-splicing

By introducing a 5′ splice site an appropriate distance upstream of the *trans*-splice acceptor site of *rol-6*, we hoped to replace *trans*-splicing with *cis*-splicing. This would convert the upstream section of the outron into an exon and provide a new 5′UTR in place of the spliced leader. Since *C.elegans* exons are typically ~50% A + U, while introns are ~70% A + U (Blumenthal and Thomas, 1988), we searched for an appropriate site within the *rol-6* outron to introduce a 5′ splice site, by looking for coincidental boundaries between an upstream low A + U region and a downstream higher A + U region. There are two such regions within the *rol-6* outron, a gradual one at −125 to −146 and a sharper one at −47 to −58. We introduced the perfect match to the extended *C.elegans* splice site consensus at each of these sites. Introduction of the splice donor site at −131 did result in *cis*-splicing, but *trans*-spliced products still predominated. We do not know why this construct was not efficiently *cis*-spliced. The exon length should be adequate: the 42 or 47 nt is substantially larger than the naturally occurring *col-12* first exon. Evidently some block to efficient *cis*-splicing exists. One possibility is that the transition between low A + U and higher A + U is not sufficiently sharp for efficient recognition of the introduced donor site by U1 snRNP.

When the splice donor site was introduced at −50, efficient *cis*-splicing occurred. Thus it is clear that when a consensus splice donor is introduced into the *rol-6* outron it can be recognized and function efficiently, completely suppressing the *trans*-splicing that normally occurs. There are two obvious differences between the −131 site and the −50 site: first, the latter creates a shorter intron and longer exon; and secondly, it occurs at a steeper transition in A + U richness. Either or both of these differences could mitigate towards more effective utilization of the Don-50 donor site.

Is intron length the key variable?

Fields (1990) reported small differences in the intron boundary consensus sequences for long and short (<75 nt) introns in *C.elegans*, suggesting the possiblity that there are differences in recognition of these two types of introns. Also, since 75% of *C.elegans* introns are close to 50 nt in length (Blumenthal and Thomas, 1988), short intron length has been strongly selected for. When we tested the possibility that intron length is the crucial difference between Don-50 and Don-131 by extending the Don-50 intron with intron-like sequence, we found that both constructs tested, which had introns of 75 and 250 nt, were *cis*-spliced efficiently. We conclude that efficient *cis*-splicing of an intron close to the 5′ end of the pre-mRNA does not require that the intron be short. Rather, some other variable(s) must be responsible for the inefficient usage of the Don-131 (a and b) and Don-col12(Δ) splice donor sites. Possibly some characteristic of the region between the −50 and −131 positions inhibits *cis*-splicing if it is located in the potential intron. Perhaps recognition sequences for RNA-binding proteins or semi-stable three-dimensional structures interfere with the spliceosome-catalyzed association of the donor and acceptor splice sites. Thus both A + U richness and shortness would increase the chances of collaboration between 5′ and 3′ splice sites. Although when the intron was expanded with simple sequence, no protein binding sites or secondary structure-forming regions were introduced, events that would increase intron length during evolution might well increase the complexity of the intron and hence might be selected against.

The equivalence of cis- and trans-splice acceptor sites

The experiments presented here, considered with those reported previously (Conrad *et al.*, 1991), demonstrate that *cis*- and *trans*-splice acceptor sites are functionally interchangeable in *C.elegans*. What then determines which process occurs in each individual instance? It seems most likely that recognition of any potential splice donor sites present upstream of the splice acceptor site is the determining variable. When a good 5′ splice site is present, *cis*-splicing predominates—U1 binds and initiates the *cis*-splicing event (Ruby and Abelson, 1988; Seraphin and Rosbash, 1989). However, *trans*-splicing presumably does not rely on U1 recognition of the 5′ splice site to trigger spliceosome formation; instead a different component of the splicing machinery presumably must identify the 3′ splice site, the *trans*-splice acceptor and the SL snRNP must be bound subsequently. This splice acceptor-site recognizing machinery is common to both types of splicing, since the acceptor sites have the same consensus sequences, and elimination of U2, U4 or U6 in the nematode *in vitro* system inactivates both *cis*- and *trans*-splicing (Hannon *et al.*, 1991). Perhaps binding of U1 to the donor site causes a searching process to occur so that as soon as the branchpoint and splice-acceptor site are recognized, '*cis*-spliceosome' assembly occurs. When no splice donor is recognized or something impedes the search, the 3′ splice site '*trans*-spliceosome' assembly is triggered. Recently several instances of naturally-occurring splice acceptor sites that can act as *cis*- or *trans*-splice acceptors have been discovered (Nonet and Meyer, 1991; W.Li, J.Shaw and R.K.Herman, personal communication; J.Culotti, personal communication; C.S.Rubin, personal communication). This suggests the possibility that the choice between these two alternative splicing modes may be made in many splicing events in *C.elegans*, presumably in a way similar to the artificial situations we have created.

Table I. List of oligonucleotides used in mutageneses

Oligonucleotide name	Oligonucleotide sequence (5'-3')	Positions
r6AG	TGACCCTAACTACGGCCACCAGTGGCGCCATTGTATTTT	2–40
Don-50	CTCGTCCAAAAAAGGTAAGTTAAAATTTCTGGA	−64 to −32
Don-131a	TCACAATGACATCCTAAGGTAAGTTTATTCTTTTTATTGTT	−149 to −109
Don-131b	TTAGCGGCAATCACAACGTTACCACATCCTAAGGTAAG	−159 to −127
*Xho*I site in Don-50	TAAGTTAAAATTTCTCGAGATATCTTTAGATCTA	−49 to −16
Don-col12	TGAAGGATATTTCACTTGGCTTCTAAAGTCCAGTGAGAGGTAAGTGGAGAATATAC	−185 to −162
Don-col12Δ	GAGAATATACGATTAGCGGCCGCAATTAAAAATTTCTGGATA	−171 to −28
Don50xA	TCGATTAAAATTAAAATTAAAATTA	n.a
Don50xT	TCGATAATTTTAATTTTAATTTTAA	n.a.

The 'positions' column indicates the region mutagenic oligonucleotides annealed to according to the numbering scheme in Figure 1 (where applicable); Don-131b was used with Don-131a-mutated DNA, *Xho*I in Don-50 with Don-50, Don-col12Δ with Don-col12.
n.a., not applicable.

Materials and methods

Worm transformation

The procedure described by Mello *et al.* (1991), with minor modifications, was used to generate transgenic strains of *C.elegans*. N2 *var* Bristol worms carrying an amber *rol-6 (n1178)* allele that is slightly dumpy in appearance (J.Kramer, unpublished) were injected with DNA at 200 μg/ml in injection buffer (Fire, 1986). DNA was purified by a large scale alkaline lysis procedure (Birnboim, 1983). The DNA was further purified by LiCl purification, digestion with RNase and proteinase K, extraction with phenol−chloroform and precipitated with ethanol (Fire and Waterston, 1989). Final pellets were resuspended in TE or water and OD_{260} taken to determine concentration, which was verified by comparison with standards on agarose gels.

Test genes

The various test genes were constructed *in vitro* from a plasmid, pRF5, carrying the *rol-6(su1006)* gene of *C.elegans* (Cox *et al.*, 1980). pRF5 was made by cloning a 2.2 kb *Hin*dIII fragment (from pRF4; Mello *et al.*, 1991) containing the entire *rol-6* gene preceded by 747 5' flanking bp into the phasmid vector pTZ19U (Mead *et al.*, 1986). This plasmid was modified by site-directed mutagenesis to create pr6AG as shown in Figure 1. The following constructions were made by oligonucleotide-directed mutagenesis (Nakamaye and Eckstein, 1986) using the oligonucleotides listed in Table I: r6AG, Don-50, Don-131a, Don-131b, creation of a *Xho*I site in the Don-50 intron, Don-col12 and Don-col12Δ. The Don-50×1 and −50×8 constructions were made as follows. The plasmid carrying the Don-50 mutation with a unique *Xho*I site introduced into its intron was cut with *Xho*I. 8 μg of each of two complementary oligonucleotides (Don50xA and Don50xT, Table I) were mixed in a 20 μl vol of kinase buffer (Maniatis *et al.*, 1982) that also contained 10 μM ATP and 0.5 units/μl T4 polynucleotide kinase. This mixture was incubated for 30 min at 37°C and 5 μl mixed with ~0.1 μg of the *Xho*I-cut plasmid (molar insert-to-plasmid ratio ~8:1). The DNAs were ligated in the presence of 1×ligase buffer (Bethesda Research Laboratories) and 1 unit of T4 DNA ligase in 20 μl total vol for 16 h at 16°C. Following transformation of *Escherichia coli* HB101, colonies containing inserts were identified by hybridization with one of the two oligonucleotides. Number and orientations of oligonucleotide inserts were determined by sequencing.

Preparation of poly(A) RNA

Total RNA was extracted from mixed populations of worms as described previously by Conrad *et al.* (1991). Poly(A) RNA was then prepared using the poly(A+) tract system from Promega (Madison, WI).

PCR analysis of RNA from transformants

Products of the *rol-6* transgenes present in the poly(A) preparations were analyzed by PCR using the procedure described previously by Conrad *et al.* (1991) with the following variations: 1 μg of a downstream primer (antisense) from position 33−18 (thus extending from the r6AG-specific region) was used in the RT reaction. The control reaction had 50 μg of RNase A added prior to the 37°C incubation. The 12 μl RT (actual reverse transcription or control) reactions were diluted to 200 μl with water and 10 μl of each used in the subsequent PCR. 0.1 μg of each primer was used in each PCR. PCRs were performed for 25 cycles (each cycle was 1 min at 92°C, 1 min at 50°C, 30 s at 72°C). The precipitated PCR products were dissolved in 7 μl formamide with 0.04% bromophenol blue and xylene cyanol and electrophoresed, after heating at 95°C for 2 min, on either 10% (20:1) acrylamide−7 M urea−TBE gel for *trans*-spliced products or 6% (30:0.8) acrylamide−7 M urea−TBE gel for *cis*-spliced products, at a field density of 40 V/cm. Southern blots of these gels were probed with an end-labelled oligonucleotide equivalent in sequence to *rol-6*, from 1−17.

Northern blots

1−10 μg samples of poly(A) RNA were electrophoresed on 1% agarose−6% formaldehyde−MOPS gels (Selden, 1987). Gels were capillary-blotted to Hybond-N membrane (Amersham, Arlington Heights, IL) in 20×SSC and probed with a specific *rol-6* riboprobe to positions 193−265 (J.Kramer, personal communication) and a random-primed probe made from a PCR product containing part of a ribosomal protein gene, *rp21* (A.Fire, personal communication). Prehybridization was performed for at least 2 h at 42°C in 50% formamide, 5×SSC, 5×Denhardt's solution, 0.1% SDS, 50 mM Na_3PO4 (pH 6.5) and 25 μg/μl sonicated salmon sperm or calf thymus DNA, followed by hybridization overnight in the same conditions, but with 2×Denhardt's solution and 200 μg/μl sonicated calf thymus DNA. Blots were washed three times for 15 min with 0.2% SDS−2×SSC at room temperature and once for 30 min at 50°C in 0.2% SDS−0.1×SSC.

Acknowledgements

We thank John Spieth and Tom Donahue for their critical reading of this manuscript. We also wish to thank Jim Kramer for his gift of the parent *rol-6* plasmid, pRF4. This work was supported by grants GM30870 and GM42432 from the National Institute of General Medical Sciences.

References

Agabian,N. (1990) *Cell*, **61**, 1157−1160.
Bektesh,S.L. and Hirsh,D.I. (1988) *Nucleic Acids Res.*, **16**, 5692.
Birnboim,H.C. (1983) *Methods Enzymol.*, **100**, 243−255.
Blumenthal,T. and Thomas,J. (1988) *Trends Genet.*, **4**, 305−308.
Conrad,R., Thomas,J., Spieth,J. and Blumenthal,T. (1991) *Mol. Cell. Biol.*, **11**, 1921−1926.
Cox,G.N., Laufer,J.S., Kusch,M. and Edgar,R.S. (1980) *Genetics*, **95**, 317−339.
Donelson,J.E. and Zeng,W. (1990) *Parasitol. Today*, **6**, 327−334.
Emmons,S. (1988) *The Nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 56−58.
Fields,C. (1990) *Nucleic Acids Res.*, **18**, 1509−1512.
Fire,A. (1986) *EMBO J.*, **5**, 2673−2680.
Fire,A. and Waterston,R.H. (1989) *EMBO J.*, **8**, 3419−3428.
Green,M.R. (1991) *Annu. Rev. Cell Biol.*, **7**, 559−599.
Hannon,G.J., Maroney,P.A. and Nilsen,T.W. (1991) *J. Biol. Chem.*, **266**, 22792−22795.
Huang,X.-Y. and Hirsh,D.I. (1992) In Setlow,J.K. (ed.), *Genetic Engineering*. Plenum Press, NY, vol 14, pp. 211−229.
Keller,M., Tessier,L.-H., Chan,R.L., Fournier,R., Weil,J.-H. and Imbault,P. (1992) *Nucleic Acids Res.*, **20**, 1711−1715.
Kramer,J.M., French,R.P., Park,E.-C. and Johnson,J.J. (1990) *Mol. Cell. Biol.*, **10**, 2081−2089.
Krause,M. and Hirsh,D. (1987) *Cell*, **49**, 753−761.
Maniatis,J., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A*

Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Mead,D.A., Szczesna-Skorupa,E. and Kemper,B. (1986) *Protein Engng*, **1**, 67–74.
Mello,C.C., Kramer,J.M., Stinchcomb,D. and Ambros,V. (1991) *EMBO J.*, **10**, 3959–3970.
Murphy,W.J., Watkins,K.P. and Agabian,N. (1986) *Cell*, **47**, 517–525.
Nakamaye,K.L. and Eckstein,F. (1986) *Nucleic Acids Res.*, **14**, 9679–9698.
Nilsen,T.W. (1989) *Exp. Parasitol.*, **69**, 413–416.
Nonet,M.L. and Meyer,B.J. (1991) *Nature*, **351**, 65–68.
Park,Y.-S. and Kramer,J.M. (1990) *J. Mol. Biol.*, **211**, 395–406.
Rajkovic,A., Davis,R.E., Simonsen,J.N. and Rottman,F.M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8879–8883.
Ruby,S.W. and Abelson,J. (1989) *Trends Genet.*, **7**, 79–85.
Selden,R.F. (1987) In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seraphin,B. and Rosbash,M. (1989) *Cell*, **59**, 349–358.
Sutton,R.E. and Boothroyd,J.C. (1986) *Cell*, **47**, 527–535.
Thomas,J.D., Conrad,R.C. and Blumenthal,T. (1988) *Cell*, **54**, 533–539.
Van Doren,K. and Hirsh,D. (1988) *Nature*, **335**, 556–559.

Received on November 6, 1992