Co-transcriptional commitment to alternative splice site selection

Gavin C. Roberts, Clare Gooding, Ho Yi Mak*, Nicholas J. Proudfoot* and Christopher W. J. Smith*

Department of Biochemistry, 80 Tennis Court Road, Old Addenbrookes Site, University of Cambridge, Cambridge CB2 1GA, UK and 1Sir William Dunn School of Pathology, South Parks Road, University of Oxford, Oxford OX1 3RE, UK

Received October 2, 1998; Revised and Accepted November 3, 1998

ABSTRACT

Production of mRNA in eukaryotic cells involves not only transcription but also various processing reactions such as splicing. Recent experiments have indicated that there are direct physical connections between components of the transcription and processing machinery, supporting previous suggestions that pre-mRNA splicing occurs co-transcriptionally. Here we have used a novel functional approach to demonstrate co-transcriptional regulation of alternative splicing. Exon 3 of the α-tropomyosin gene is specifically repressed in smooth muscle cells. By delaying synthesis of an essential downstream inhibitory element, we show that the decision to splice or repress exon 3 occurs during a limited window of opportunity following transcription, indicating that splice site selection proceeds rapidly after transcription.

INTRODUCTION

A functional coupling between transcription and pre-mRNA splicing has long been suspected. Cytological observations of nascent transcript processing indicated that transcription and RNA splicing occur simultaneously (1–3). Nevertheless, an obligatory functional link between RNA polymerase II (RNA pol II) transcription and pre-mRNA processing seemed unlikely since efficient in vitro splicing and 3′-end processing reactions are observed with transcripts that have been previously synthesized by bacteriophage RNA polymerases. Since the in vitro reactions are usually carried out with simplified model substrates, it remains possible that larger, more complex transcripts may require more carefully orchestrated processing in vivo. Moreover, the complex processes of alternative splicing may require co-ordination between RNA pol II and the processing machinery for appropriate regulation.

Recent experiments have supported the concept of a functional link between RNA pol II transcription and all three major RNA processing operations (capping, splicing and polyadenylation), and have provided insights into the physical interactions which form the basis of this co-ordinated activity (reviewed in 4–6). The unique C-terminal domain (CTD) of the largest subunit of RNA pol II provides a physical connection between transcription and RNA processing. Capping enzymes (7–10), elements of the cleavage and polyadenylation apparatus (11,12), splicing factors such as SR proteins (13), and even spliceosomes (14–16), are all found to associate with the CTD. Some of these interactions occur preferentially or solely with the hyperphosphorylated CTD, which is associated with actively elongating polymerase. In addition to the in vitro biochemical data, splicing factors are found to localize at sites of transcription in vivo (17,18). Most importantly, a functional correlation to these observations is provided by transfection experiments in which deletion of the CTD was found to inhibit RNA splicing, 3′-end processing and transcription termination (12). In addition, the RNA pol II, or its CTD alone, has recently been characterized as an essential 3′-end processing factor in vitro in the absence of ongoing transcription (19). This activity was independent of the phosphorylation state of the CTD. Taken together, these observations suggest that the CTD serves to integrate the processes of transcription and RNA processing by providing a platform upon which RNA processing complexes assemble ready to catalyze the maturation of the pre-mRNA as it is synthesized.

Most of the recent investigations have probed the physical interactions between elements of the processing and transcription machineries. With so much evidence pointing towards a close physical and temporal association between transcription and RNA splicing, it would appear that the regulation of alternative splicing must also necessarily occur very closely after transcription. In this paper, we use a novel functional approach to probe the temporal coupling between the processes of transcription and alternative splicing in vivo. Our results show that splice site selection for an alternative tropomyosin exon is made almost immediately following its transcription, and imply a very tight temporal coupling between the processes of transcription,
splicing and alternative splicing. These data have implications for the possible mechanisms of alternative splicing regulation in vivo.

MATERIALS AND METHODS

Splicing constructs

All in vitro splicing constructs were derivatives of pTS3St (20). Insertion of spacers and pause elements were made at the BstEII site downstream of exon 3. The spacer insertions of 129, 290 and 391 nt were made by inserting fragments of 147, 308 and 409 nt into the vector cut with BstEII and PsrI, which removes 18 nt downstream of α-TM exon 3. These nucleotides are not required for regulation of splicing (20). The spacer fragments for insertion were derived from the 5′ end of intron 1 of the α-TM gene by PCR, including nucleotides from position 18 of the intron onwards. Intron 1 contains no sequences involved in regulation of splicing (20). The 5′ end of the fragments was engineered to contain a BstEII site, while the 3′ ends had PsrI sites to allow directional cloning into the vector. The remaining constructs were made by inserting into the BstEII site which had been blunted by filling in with Klenow DNA polymerase fragment (21). The 301 nt spacer fragment was a PvuII–HincII fragment of pBlue-script-II SK+, with the PvuII end nearest to exon 3. The MAZ4 C2 and c2 elements were excised as BamHI–SalI fragments derived from the plasmids A3SPA/MAZ4, SPAC2a+, and SpA×1+, respectively (22,23) and blunt-ended before ligation. For each of the characterized pause elements, constructs were recovered which contained the insert in either its original (+) or reverse (−) orientation. For the MAZ4c3′ constructs the MAZ4 fragment was cloned into the EcoRI site downstream, between the downstream regulatory element and exon 4. Constructs for in vitro splicing were made by subcloning into the △107 plasmid (20).

Analysis of splicing

For in vivo analysis, PAC-1 smooth muscle (SM) cells were transfected with splicing constructs by the calcium phosphate method, as described previously (20,24). In most experiments, the reactions were scaled down proportionately for transfection of cells growing in 4 cm wells, rather than 10 cm dishes. Total RNA was harvested with Tri-Reagent (Sigma), then reverse transcription and PCR were performed as described previously (20,24,25). Note that quantitation of the PCR reactions involved measuring the ratio of the bands corresponding to exon 3 inclusion or skipping. We have previously shown that this ratio is maintained during multiple cycles of PCR (25). Therefore PCR was carried out over 30 cycles to maximize the yield of product. These conditions allow a reliable assessment of the ratio of the two spliced products in a single lane, but the absolute quantities between lanes cannot be reliably compared. In practice, the gels were loaded to give approximately equal amounts of the 1–4 spliced product. For in vitro analysis, RNA transcripts were synthesized and incubated in splicing reactions with HeLa nuclear extract as described previously (20,24). Standard in vitro splicing reactions were incubated for 3 h. The results of splicing in vivo and in vitro were quantitated using a Molecular Dynamics Phosphor-Imager and ImageQuant software as before (24). For the in vitro splicing data the ratio of 2–4 to 3–4 splicing was calculated, taking into account the different content of radiolabelled bases in the two spliced products.

RESULTS

We have used the α-tropomyosin (α-TM) gene as a model system for studying alternative splicing (20,24,26–28). Exons 2 and 3 of α-TM are mutually exclusive, with exon 3 selected in most cells and tissues, and exon 2 specifically selected in SM cells. We have previously shown that the switch to selection of exon 2 in SM is the result of specific inhibition of exon 3, mediated by cis-elements in the flanking introns (Fig. 1; 20,24,29). These elements, known as the upstream and downstream regulatory elements (URE and DRE, respectively), are both required for SM repression of exon 3. An inhibitory complex involving factors bound to these elements and to the upstream polypyrimidine tract of exon 3 is thought to prevent the assembly of functional splicing complexes on the exon (24,28–30). Although not fully characterized, one component of the complex appears to be the polypyrimidine tract binding protein (PTB). Mutations of optimal PTB binding sites in either the polypyrimidine tract of exon 3 or a pyrimidine-rich segment of the DRE known as DY, impair regulated splicing in vivo and reduce the affinity of PTB binding in vitro (24,29).

We reasoned that if the choice between splicing or repression of exon 3 occurs co-transcriptionally, delaying synthesis of the inhibitory DRE after the exon has been transcribed might antagonize regulation by allowing splicing factors to assemble on the exon before the full inhibitory complex could be formed. To test this hypothesis we designed a series of constructs based on pTS3St (Fig. 2A), a previously characterized test construct comprising α-TM exons 1, 3 and 4. The effect of SM regulation is to exclude exon 3 from a proportion of the RNAs expressed after transient transfection of cultured cells (20,24,28). Spacer elements were inserted 50 nt downstream of exon 3 between the 5′ splice site and the DRE, which lies a further 90 nt downstream. The site of insertion was chosen because it is sufficiently distant from the consensus splice sites and regulatory elements and is in a region that is not required for splicing or regulation of splicing (20,24). The spacers are expected to delay synthesis of the DRE by the time taken for RNA polymerase to transcribe them.
The constructs were transfected into PAC-1 SM cells (31) and transiently expressed RNA was analyzed by reverse transcription followed by PCR. While the control construct pTS3St showed normal regulation (48% exon 3 skipping in Fig. 2B, lane 1), the degree of regulated exon 3 skipping decreased proportionately with the length of spacer inserted [Fig. 2B (lanes 2–4) and C]. The original three spacers (129, 290 and 391, lanes 2–4, Fig. 2B) were a nested set containing the same sequence at the 5′ end. To verify that these spacers were neutral we tested an unrelated 301 nt spacer derived from pBluescript SK+. This spacer had an almost identical effect to the 290 spacer (Fig. 2C).

The decrease in regulated exon skipping observed after insertion of spacer elements could arise from the delay in synthesis of the DRE by RNA pol II. Alternatively, it might result from the increased physical distance between the RNA regulatory elements. To distinguish between these possibilities we made constructs in which DNA elements known to cause pausing of RNA polymerase (22,23,32) were inserted in place of the spacers. The effect of polymerase pausing between exon 3 and the DRE might be to allow further opportunity for splicing factors to assemble on exon 3 before the DRE is transcribed. This would promote commitment of exon 3 to the splicing pathway and consequent inclusion in the resulting mRNA. An effect on regulated exon skipping above that expected by virtue of spacing alone would indicate that insertions at this position impair regulation by delaying transcription, and not by increasing displacement of the regulatory elements. The C2 and α2 elements are characterized pause sites situated downstream of the human complement C2 and α2 globin genes (22,32). Each of these genes is followed by an unusually short intergenic space before the next downstream gene and the role of the pause elements, in conjunction with their associated upstream 3′-end processing sites, is to promote efficient transcriptional termination. Deficiencies in transcription termination could lead to inhibition of transcription initiation at the downstream gene (33,34). The C2 pause site contains a binding site for the MAZ zinc finger protein (23). The MAZ4 constructs contain four tandem repeats of a G5AG5 consensus binding site for MAZ. This produces a very strong synthetic RNA pol II pause site.

All three pause elements caused a reduction in the amount of regulated exon skipping that was greater than would be expected upon the basis of their size alone (Fig. 2B and C). This is indicated most clearly in Figure 2C where the percentage of exon skipping, compared to the wild type pTS3St construct, is plotted against the size of the inserted spacer element. Note that the points corresponding to all the pause elements fall well below the line corresponding to the ‘neutral’ spacers. The C2 and α2 elements acted equally well in either orientation. In poly(A) site competition assays these elements were found to be stronger in the native (+) orientation (22). However, in transcription interference assays they were found to be equally active in either orientation (33). Consistent with its performance in assays for pausing activity, the synthetic MAZ4 element caused the most dramatic decrease in regulation, resulting in an average of only 8% (±5%) of the exon skipping observed for the wild type construct (Fig. 2B, compare lanes 1 and 5). This effect significantly exceeds that expected to result simply from displacement by neutral spacing (Fig. 2C). Inversion of the MAZ4 element led to a lesser, though still discernible effect, reducing regulation to 34% (±21%) of wild type levels [Fig. 2B (lane 6) and C]. This finding is consistent with observations that the MAZ element has reduced activity in

Figure 2. Co-transcriptional regulation of α-TM splicing in vivo. (A) Diagram of the basic splicing construct, pTS3St, comprising α-TM exons 1, 3 and 4. Exon 3 is flanked by the regulatory elements URE and DRE. The URE lies between exon 3 and its upstream branch point and poly(A) polyuridine tract (shown by the black circle and rectangle, respectively) and comprises a ~20 nt conserved region containing three UGC motifs. The DRE lies ~150 nt downstream of exon 3. After transient transfection of SM cells, a proportion of mRNAs specifically exclude exon 3. This effect is dependent on both regulatory elements (20,24,29). The position at which inserts are introduced is indicated. (B) Sample analysis of test construct splicing in SM cells. Gel shows results of transient transfections of the constructs indicated, followed by RT–PCR analysis of expressed mRNA using primers to invariant flanking regions as described previously (20). The larger product includes exon 3, while the smaller species corresponds to regulated exclusion of exon 3. The percentage of exon 3 exclusion in this experiment is indicated beneath each lane. Note that PCR was carried out under conditions that allow for reliable analysis of the ratio of the two products, but not for comparison of absolute amounts between lanes. The gel was loaded to give approximately equal amounts of the 1–3–4 product in each lane. (C) Graphical representation of multiple transfection experiments. Three repeats of the experiment described in (B) were performed, and the amount of regulated exon skipping of each construct was calculated as a percentage of that observed for the wild type pTS3St construct in that experiment. Average percentages across experiments were computed (y axis), with standard deviations (error bars), and plotted against the length of spacer introduced (x axis). Introduction of progressively longer neutral spacers (129, 290, 301 and 391 nt) resulted in a decrease in exon skipping that was proportional to the spacer length (open squares, joined by line). Introduction of progressively longer neutral spacers (129, 290, 301 and 391 nt) resulted in a decrease in exon skipping that was proportional to the length of spacer inserted [Fig. 2B (lanes 2–4) and C]. The original three spacers (129, 290 and 391, lanes 2–4, Fig. 2B) were a nested set containing the same sequence at the 5′ end. To verify that these spacers were neutral we tested an unrelated 301 nt spacer derived from pBluescript SK+. This spacer had an almost identical effect to the 290 spacer (Fig. 2C).

The decrease in regulated exon skipping observed after insertion of spacer elements could arise from the delay in synthesis of the DRE by RNA pol II. Alternatively, it might result from the increased physical distance between the RNA regulatory elements. To distinguish between these possibilities we made constructs in which DNA elements known to cause pausing of RNA polymerase (22,23,32) were inserted in place of the spacers. The effect of polymerase pausing between exon 3 and the DRE might be to allow further opportunity for splicing factors to assemble on exon 3 before the DRE is transcribed. This would promote commitment of exon 3 to the splicing pathway and consequent inclusion in the resulting mRNA. An effect on regulated exon skipping above that expected by virtue of spacing alone would indicate that insertions at this position impair regulation by delaying transcription, and not by increasing displacement of the regulatory elements. The C2 and α2 elements are characterized pause sites situated downstream of the human complement C2 and α2 globin genes (22,32). Each of these genes is followed by an unusually short intergenic space before the next downstream gene and the role of the pause elements, in conjunction with their associated upstream 3′-end processing sites, is to promote efficient transcriptional termination. Deficiencies in transcription termination could lead to inhibition of transcription initiation at the downstream gene (33,34). The C2 pause site contains a binding site for the MAZ zinc finger protein (23). The MAZ4 constructs contain four tandem repeats of a G5AG5 consensus binding site for MAZ. This produces a very strong synthetic RNA pol II pause site.

All three pause elements caused a reduction in the amount of regulated exon skipping that was greater than would be expected upon the basis of their size alone (Fig. 2B and C). This is indicated most clearly in Figure 2C where the percentage of exon skipping, compared to the wild type pTS3St construct, is plotted against the size of the inserted spacer element. Note that the points corresponding to all the pause elements fall well below the line corresponding to the ‘neutral’ spacers. The C2 and α2 elements acted equally well in either orientation. In poly(A) site competition assays these elements were found to be stronger in the native (+) orientation (22). However, in transcription interference assays they were found to be equally active in either orientation (33). Consistent with its performance in assays for pausing activity, the synthetic MAZ4 element caused the most dramatic decrease in regulation, resulting in an average of only 8% (±5%) of the exon skipping observed for the wild type construct (Fig. 2B, compare lanes 1 and 5). This effect significantly exceeds that expected to result simply from displacement by neutral spacing (Fig. 2C). Inversion of the MAZ4 element led to a lesser, though still discernible effect, reducing regulation to 34% (±21%) of wild type levels [Fig. 2B (lane 6) and C]. This finding is consistent with observations that the MAZ element has reduced activity in
pausing RNA polymerase when transcribed in the opposite orientation \((22,23)\). The residual activity observed after inversion of the MAZ\(_4\) element also shows that its effect is not a function of the RNA sequence introduced at this position, but is more likely to result from activity at the level of DNA. As a control to show that the effects of the spacer/pause elements depended upon their position upstream of the DRE, we inserted the most potent element, MAZ\(_4\), downstream of the DRE in both orientations. In these two constructs (MAZ\(_4\)3\(^{\prime}\) and -) the length of the introns between exons 3 and 4 were identical to the MAZ\(_4\) constructs, but the MAZ\(_4\) element would be synthesized after the DRE and so would not delay its synthesis. As expected, insertion of the MAZ\(_4\) element in either orientation downstream of the DRE had no effect on exon skipping [Fig. 2 (lanes 7 and 8) and C], showing that its activity depends on its position, and not merely its presence in the pre-mRNA.

To confirm that the effects of the spacer and pause elements in vivo resulted from the coupling of transcription and splicing, some of the elements were tested in an in vitro splicing assay. Splicing of pre-synthesized RNAs in vitro is necessarily uncoupled from transcription, so the effects of the various insertions on splicing alone could be determined (Fig. 3). The A107 pre-mRNA comprises exons 2, 3 and 4 of \(\alpha\)-TM, offering a binary choice between 2–4 and 3–4 splicing (Fig. 3A). In HeLa nuclear extracts both 2–4 and 3–4 spliced products are observed in similar quantities (Fig. 3B, lane 1; 20, 24). A background of regulatory activity is known to cause selection of exon 2 in a proportion of transcripts, and this regulated pathway is absolutely dependent on the integrity of the DRE. All mutations within the DRE that impair regulated splicing in SM cells have been shown to have the same effect upon the 2–4 splicing pattern in vitro (20, 24). This activity of the DRE correlates with its ability to bind PTB (24). Insertion of the 290 and 391 nt neutral spacers, the C2+ and MAZ\(_4\) pause elements between the 5′ splice site and DRE of Δ107 had no significant effect on the ratio of 2–4 and 3–4 splicing from these pre-mRNAs (lanes 1–5). Neither did insertion of MAZ\(_4\) downstream of the DRE (lane 6). This is best indicated by the relative proportions of 2–4 and 3–4 spliced products, since these spliced products have identical structure when processed from the various mutant transcripts (Fig. 3B), unlike the lariat products which vary considerably in size and radiolabel content. The fact that the MAZ\(_4\) element produced a 10-fold reduction in regulated exon skipping in vivo (Fig. 2B and C), but had no discernible effect in vitro, strongly suggests that the effects of the spacer and pause insertions upon alternative splicing result indirectly from primary effects upon transcription.

**DISCUSSION**

The data that we have presented is consistent with the growing body of evidence that the processes involved in mRNA synthesis—transcription, capping, 3′-end formation, splicing and regulation of splicing—are closely coupled, either physically and/or temporally. Previous investigations have shown that increasing the distance between competing processing sites can alter the balance of competition between two poly A sites, or between a splicing event and 3′-end processing (35–37). Here we have increased the displacement between a negative splicing regulatory element and the exon that it regulates, and shown that increased displacement decreases the regulated exclusion of the exon. In default splicing of \(\alpha\)-TM, constitutive splicing factors bind to the splice site signals of exon 3, committing it to the splicing pathway (27). In regulated splicing, an alternative set of factors, including PTB, are thought to bind to elements in the flanking introns, forming an inhibitory complex which blocks assembly of the constitutive splicing apparatus, thereby repressing the exon (24, 28, 29).
experiments we have described use insertions to delay the transcription of the essential DRE inhibitory element required for regulation of TM exon 3. By using previously characterized RNA polymerase pause sites (Fig. 2), and by comparing the effects of the spacer insertions in vivo and in vitro (Figs 2 and 3), we have been able to show that the effects of the spacers were due to delaying the synthesis of the DRE. The relative effect of the pause sites in disrupting regulated splicing (MAZ > C2, α2) correlates roughly with their activity in other assays (22; N.J.P., unpublished observations). Splice site selection and commitment to splicing presumably occur within the period of this delay and must therefore occur very rapidly after transcription. Likewise, formation of the negative regulatory complex that suppresses exon 3 must therefore occur very rapidly after transcription. Likewise, formation of the negative regulatory complex that suppresses exon 3 must occur immediately after transcription of the relevant sequences. Our results therefore indicate that not only splicing factors (17,18), but also regulators of splicing must be available at the sites of gene transcription.

Extrapolation of the effects of neutral spacers suggests that an insertion of 600 bp would result in complete abolition of regulation (Fig. 2C). In these circumstances, full commitment of the exon to splicing will have occurred before synthesis of the DRE. This observation allows a rough estimate of the distance behind the transcribing polymerase at which commitment to splicing occurs: in this case, the displacement between the splice site furthest upstream (i.e. the branchpoint sequence of exon 3) and the transcribing polymerase would be ~1 kb. Therefore these data suggest that the lag between transcription and splice site selection in this system is of the order of 1 kb transcription distance. This figure agrees with previous measurements of the rates of splicing in vivo, where 50% of introns were found to have completed the entire process of splicing within 5 kb transcription distance after their appearance (1).

Decisions between alternative processing pathways are increasingly observed to depend on a fine balance of splice site strengths and relative separations of splice site and regulatory sequences (reviewed in 14 and 38) and our work re-emphasizes this emerging theme. A novel aspect of our findings here is the demonstration that the relative distance between competing elements can affect kinetic competition between alternative processing pathways co-transcriptionally. This supports previous suggestions that in principle some instances of regulated RNA processing could actually arise from regulated pausing of RNA polymerase rather than from a direct RNA-based regulatory system (39). However, as yet there is no evidence that such pausing actually plays a role in regulating the splicing of TM, or of any other characterized pre-mRNA.

While our data highlight the temporal coupling between transcription and alternative splicing, a recent report has indicated that alternative exon inclusion may be influenced by the identity of the promoter which drives transcription (40), implying that the precise composition of the initiation complex formed at different promoters may influence the array of splicing factors that associate with the elongating polymerase. SR proteins are known to influence alternative splicing decisions and to associate with the CTD of RNA pol II, suggesting the possibility of coordinated interplay between the regulation of expression of individual genes at the level of both transcription and splicing.

ACKNOWLEDGEMENTS

We thank Gilles Moreau for help in construction of the original spacer constructs and Justine Southby for critically reading the manuscript. This work was supported by grants from the Medical Research Council and the British Heart Foundation to C.W.J.S. and from the Wellcome Trust (032773) to N.J.P.

REFERENCES