The EMBO Journal vol.13 no.16 pp.3861–3872, 1994

Smooth muscle-specific switching of α-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon

Clare Gooding, Gavin C.Roberts, Gilles Moreau[1], Bernardo Nadal-Ginard[1] and Christopher W.J.Smith[2]

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK and [1]Department of Cardiology, Children's Hospital, Department of Cellular and Molecular Physiology, Harvard Medical School, Boston MA 02115, USA
[2]Corresponding author

Communicated by R.J.Jackson

Exons 2 and 3 of α-tropomyosin are spliced in a strict mutually exclusive manner. Exon 3 is a default choice, being selected in almost all cell types where the gene is expressed. The default selection arises from a competition between the two exons, in which the stronger branch point/pyrimidine tract elements of exon 3 win. Exon 2 is selected predominantly or exclusively only in smooth muscle cells. We show here that the basis for the smooth muscle-specific switching of exon selection is inhibition of exon 3. Exon 3 is still skipped with smooth muscle specificity, even in the absence of exon 2. We have defined two conserved sequence elements, one in each of the introns flanking exon 3, that are essential for this regulation. Mutation of either element severely impairs regulated suppression of exon 3. No other exon or intron sequences appear to be necessary for regulation. We have also demonstrated skipping of exon 3 that is dependent upon both regulatory elements in an *in vitro* splicing assay. We further show that both splice sites of exon 3 must be inhibited in a concerted fashion to switch to selection of exon 2. This may relate to the requirement for negative elements on both sides of the exon.

Key words: alternative splicing/gene regulation/RNA splicing/smooth muscle/tropomyosin

Introduction

Alternative splicing of pre-mRNA is a widespread phenomenon which provides a versatile means of regulating gene expression. Many genes give rise to a number of protein products with distinct functions via alternative splicing, and expression of the different isoforms is often regulated with cell type or developmental specificity (Smith *et al.*, 1989a). It is therefore important to understand how alternative splicing is controlled. The basic mechanisms of constitutive splicing have been the subject of intensive study over the past 10 years [for recent reviews see Lamond (1993) and Moore *et al.* (1993)]. The splicing reaction proceeds in two successive *trans*-esterification steps, yielding the spliced mRNA and the excised intron with a looped or 'lariat' structure. Specific *cis* sequence elements at the intron boundaries are required. These are the 5′ splice site consensus sequence and three elements at the 3′ end of the intron, which comprise in 5′ to 3′ order the branch point sequence, the polypyrimidine tract and then the 3′ splice site AG. Various splicing factors, including the U1, U2, U4/6 and U5 snRNPs, are required to assemble with the pre-mRNA into the spliceosome, the 60S complex within which the splicing reaction occurs. The consensus *cis* elements provide binding sites for some of these splicing factors. The 5′ splice site base pairs with U1 snRNA (Zhuang and Weiner, 1986) and the branch point with U2 snRNA (Parker *et al.*, 1987). The polypyrimidine tract is the binding site for at least two non-snRNP splicing factors, U2AF (Ruskin *et al.*, 1988; Zamore *et al.*, 1992) and PSF (Patton *et al.*, 1993). The degree of similarity with the consensus sequences seems to correlate with the functional strengths of splice site elements and with the affinity for associated splicing factors (e.g. Reed and Maniatis, 1988; Zhuang *et al.*, 1989; Lear *et al.*, 1990; Mullen *et al.*, 1991; Zamore *et al.*, 1992). Indeed, the relative balance of splice site strengths is often an important factor in establishing particular patterns of alternative splicing when splice sites are in competition with each other (Smith *et al.*, 1989a).

In attempting to understand regulated alternative splicing, we need to know what additional influences are operating over and above the basic requirements for constitutive splicing. In many cases of alternative splicing a particular splicing pattern occurs in a wide variety of cell types, and this default pattern may simply represent the outcome of the differential interaction of competing splice sites with the constitutive splicing machinery. However, in cases where there is a clear switch in splicing patterns between two different cell types, then some *trans* regulation must be involved. Such *trans* regulation may involve the participation of cell-specific regulatory factors, or it could result from variations in the activities of constitutive factors. Moreover, regulation could operate in either a negative or positive manner. Recently, mechanistic insights into both positive and negative regulation of splicing by specific factors have been obtained using the somatic sex-determining pathway of *Drosophila* (Tian and Maniatis, 1992, 1993; Valcarcel *et al.*, 1993). Also, the ability of the constitutive splicing factor SF2/ASF to influence splice site competition in mammalian introns has been demonstrated (Mayeda and Krainer, 1992; Zahler *et al.*, 1993).

The rat α-tropomyosin (αTM) gene has provided an instructive mammalian model system for looking at mechanisms of alternative splicing. This gene generates a large number of tissue-specific isoforms by a series of regulated alternative splicing events (Wieczorek *et al.*, 1988; Lees-Miller *et al.*, 1990). At the 5′ end of the gene exons 2 and 3 encode interchangeable protein domains and are

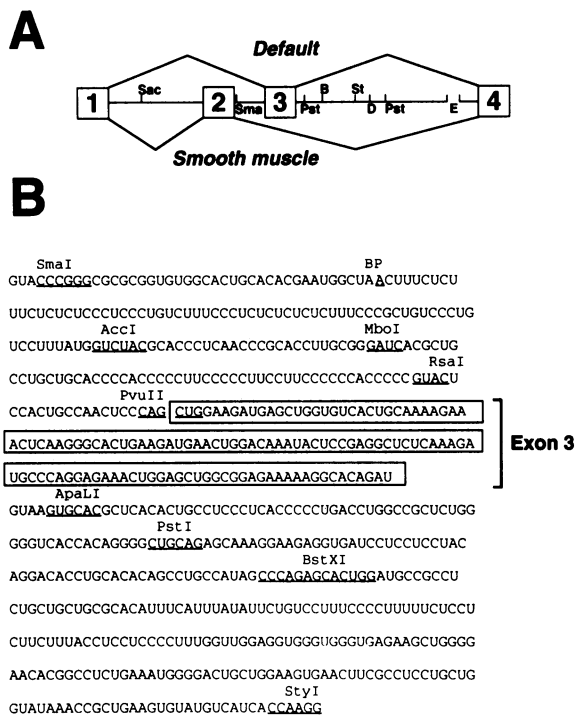

Fig. 1. Mutually exclusive splicing of α-tropomyosin exons 2 and 3. (**A**) Summary of the structure of the 5′ end of the α-tropomyosin gene and default and regulated splicing patterns. The location of some restriction sites is indicated: Sac, *Sac*I; Sma, *Sma*I; Pst, *Pst*I; B, *Bst*XI; S, *Sty*I; D, *Dra*I; E, *Eco*RI. (**B**) Sequence of α-tropomyosin pre-mRNA in the region of exon 3 from *Sma*I to *Sty*I sites. Restriction sites mentioned in the text are underlined, as is the branch point (BP) of exon 3. Exon 3 sequences are boxed.

spliced in a mutually exclusive fashion (Figure 1). The basis for the strict mutually exclusive behaviour is steric interference arising from the 41 nucleotide separation of the exon 3-associated branch point and the 5′ splice site of exon 2 (Smith and Nadal-Ginard, 1989). In almost all cells and tissues examined, exon 3 is chosen either exclusively or predominantly, and so represents the 'default' selection by the splicing machinery (Wieczorek *et al.*, 1988). This default selection is primarily the result of the 3′ splice site elements associated with exon 3, which are functionally stronger than those of exon 2, leading to exon 3 out-competing exon 2 in most cell types. If the polypyrimidine tract or, to a lesser extent, the branch point, of exon 3 are mutationally weakened, or if exon 3 is completely deleted, then exon 2 is efficiently incorporated into mRNA (Mullen *et al.*, 1991). The functional strength of the exon 3 pyrimidine tract has been correlated with a 200-fold higher affinity for the splicing factor U2AF compared with the exon 2 pyrimidine tract (Zamore *et al.*, 1992).

Having explained the basis of mutually exclusive behaviour of exons 2 and 3, and the default selection of exon 3, we have now focused our attention on the tissue-specific selection of exon 2. The only cell type in which exon 2 is predominantly selected is smooth muscle. In smooth muscle tissue exon 2 is incorporated exclusively (Wieczorek *et al.*, 1988), while in cultured smooth muscle cells exon 2 can be found in >50% of αTM mRNA (Rothman *et al.*, 1992). This switch in exon selection could conceivably involve activation of the splice sites of exon 2 or inhibition of exon 3. In this paper we show that the smooth muscle-specific regulation of αTM exons 2 and 3 involves specific inhibition of exon 3. We also demonstrate that two sequence elements are necessary for this negative regulation. These elements lie in each of the introns flanking exon 3. The upstream regulatory element (URE) lies ~50 nucleotides upstream of exon 3, between the polypyrimidine tract and the 3′ splice site. The downstream regulatory element (DRE) lies >130 nucleotides downstream of exon 3. The two elements share sequence motifs and show some phylogenetic conservation between mammalian and avian genes. Deletion of either element severely reduces skipping of exon 3, not only in smooth muscle but also the small proportion of transcripts that usually exclude exon 3 in non-smooth muscle cells. Furthermore, we show that transcripts containing αTM exons 2–3–4 are spliced in HeLa nuclear extract to give predominantly 3–4 spliced products, and that a significant proportion of transcripts follow the 2–4 splicing pathway. Significantly, the production of *in vitro* 2–4 spliced products is dependent upon the integrity of both the regulatory sequences. These transcripts therefore provide an *in vitro* assay for the inhibitory factors that regulate αTM splicing.

Results

Smooth muscle-specific splicing of transfected constructs

To study the mechanism of smooth muscle regulation of TM splicing we have analysed transient RNA expression from constructs containing the regulated region of αTM after transfection into smooth muscle or non-smooth muscle cell lines. We have used various non-smooth muscle cell lines (COS, HeLa, L) and have consistently found no difference between the splicing patterns observed. For smooth muscle cells we have used cultured pulmonary artery smooth muscle cells (PA cells) which have been found (i) to maintain endogenous αTM splicing patterns of at least 50% exon 2 incorporation over large passage numbers, and (ii) to splice transfected constructs similarly (Rothman *et al.*, 1992). This degree of exon 2 inclusion is not as great as in smooth muscle tissue (Wieczorek *et al.*, 1988), but it is higher than in any other cultured cell system we have tested.

Our basic construct design contains αTM genomic regions encompassing exons 1–4, with various deletions in introns 1, 2 and 3. Exons 1 and 4 are fused to 5′ and 3′ SV40 sequences that provide enhancer, promoter and 3′ end processing signals (but no splice sites). We initially transfected the construct TS23D, containing exons 1–4, into PA and COS cells and analysed transient expression by S1 nuclease protection using probes specific for the 124 or 134 splicing patterns (Figure 2). The probes were labelled within SV40 sequences to detect construct but not endogenous αTM RNA. In COS cells, construct TS23D was spliced in the expected default pattern with predominant inclusion of exon 3 and exclusion of exon 2, as indicated by the high ratio of fully protected to partially protected 134 probe (Figure 2A), and the corresponding predominant partial protection of the 124 probe (Figure 2B). In contrast, in PA cells the same construct had a different splicing pattern, with exon 3

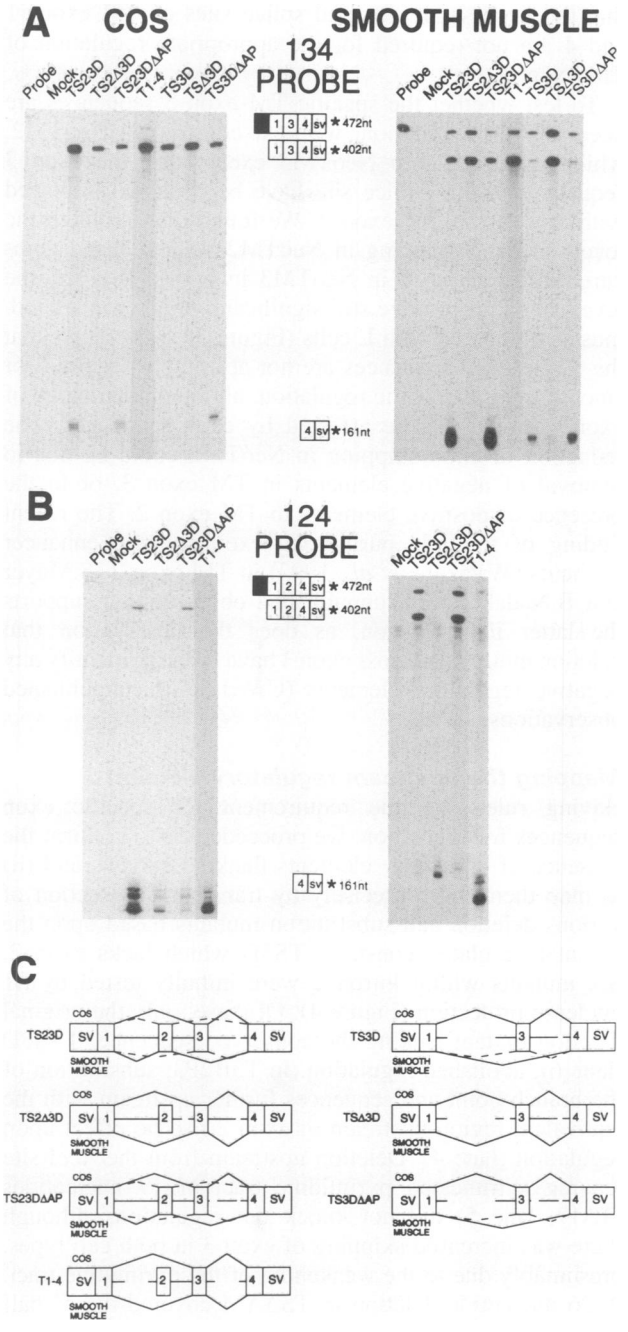

Fig. 2. Smooth muscle-specific inhibition of exon 3. The constructs shown schematically in (**C**) were transfected into COS and PA (smooth muscle) cells, and transiently expressed RNA was harvested 48 h post-glycerol shock. RNA was analysed by S1 nuclease protection using cDNA-derived probes specific for the 134 (**A**) or 124 (**B**) splicing pathways. Probes were labelled within 3′ SV40 sequences as indicated by the asterisk to detect only construct RNA; they had a 70 nucleotide tail to distinguish true full protection (402 nucleotides) from re-annealed probe (472 nucleotides). Partial protection at 161 nucleotides represents exclusion of the mutually exclusive exon present in the probe. The deduced splicing patterns in COS cells are shown above, and those for smooth muscle cells are shown below the diagram of each construct in (C). Full lines represent major or exclusive splicing patterns, while dashed lines represent minor splicing patterns.

now predominantly excluded and exon 2 incorporated instead (Figure 2; Rothman *et al.*, 1992). Thus, smooth muscle regulation can be maintained with transfected αTM minigene constructs.

Smooth muscle regulation involves inhibition of αTM exon 3

The switch in splicing pattern we observe in PA cells could arise from either activation of exon 2 or inhibition of exon 3. Arguing against exon 2 activation is the observation that in the absence of exon 3, exon 2 is incorporated efficiently into RNA in all cell types and in HeLa nuclear extract (Mullen *et al.*, 1991). To distinguish between activation and inhibition, we have transfected a construct in which exon 2 is deleted (TS3D) and which shows full incorporation of exon 3 in most cells (Mullen *et al.*, 1991). If regulation is purely by activation of exon 2, we would expect to see no difference in splicing of this construct between PA and other cells. On the other hand, if regulation involves inhibition of exon 3, then we would expect to see some exclusion of exon 3 in PA cells even in the absence of exon 2. Consistent with an inhibitory mechanism of regulation, we found that in RNA from TS3D exon 3 was still skipped significantly in PA cells, while it was fully incorporated in COS cells (TS3D; Figure 2A and C). Thus, smooth muscle regulation of αTM splicing involves inhibition of exon 3, and this inhibition does not require the presence of exon 2. This result does not preclude the possibility of activation of exon 2 in PA cells, although the fact that exon 2 is used efficiently in the absence of exon 3 argues against a major role for such activation. An alternative explanation, that exon 3 is specifically activated in all cells except smooth muscle, seems intuitively unlikely and is inconsistent with the mapping data below.

Flanking elements are necessary for inhibition of exon 3

The negative regulation of exon 3 must be mediated by specific *cis* sequences within the pre-mRNA. We therefore tested various constructs derived from the regulated constructs TS3D and TS23D. In constructs TS2Δ3D and TSΔ3D (Figure 2) a 49 nucleotide deletion was introduced downstream from the *Acc*I site between the polypyrimidine tract and 3′ splice site of exon 3. The rationale for this deletion was that a phylogenetically conserved sequence was found at the 3′ end of this region (see below). In both constructs this deletion was found to abolish regulation, causing complete inclusion of exon 3 in both cell types. Thus, there is an element between exon 3 and its polypyrimidine tract that is essential for the smooth muscle-specific repression of exon 3. It is worth noting that this deletion also reduces the small amount of exon 3 skipping seen in COS cells.

We next tested whether any intron elements downstream of exon 3 are necessary for smooth muscle regulation. A deletion between *Apa*LI and *Pst*I sites removing 64 nucleotides immediately downstream of exon 3 (constructs TS23DΔAP and TS3DΔAP, Figure 2) did not abolish regulation. In fact this deletion actually appeared to enhance the amount of exon 3 skipping in both COS and PA cells (compare TS3D and TS3DΔAP, Figure 2). Construct T1-4 contains exons 2 and 3, but has deletions in introns 1 and 3 downstream from the *Pst*I site. This construct is entirely unregulated; spliced RNA included exclusively exon 3 in both COS and PA cells. It seems unlikely that the deletion in intron 1 is responsible for the loss of regulation in T1-4, since TS3D had an even more

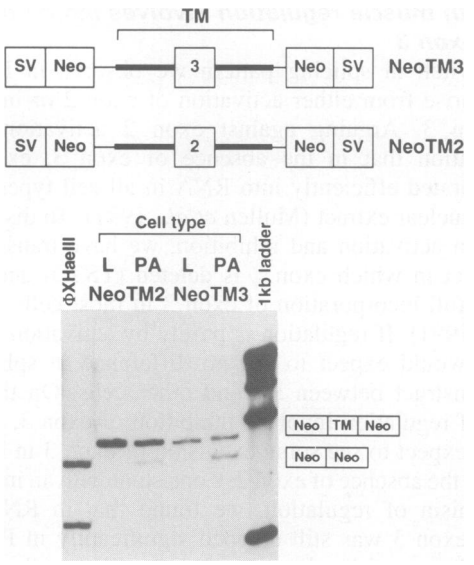

Fig. 3. Regulation is not dependent upon exon sequences or flanking splice sites. Constructs NeoTM3 and NeoTM2 are shown schematically. Intron sequences shown by the thin line are derived from human β-globin intron 1, and those shown by thick lines are the introns flanking TM exon 3. The TM insert in NeoTM3 is from the *Sma*I to *Dra*I sites (Figure 1). In NeoTM2, the exon 3 sequences and 5′ splice site have been cleanly replaced with the equivalent TM exon 2 sequences. RNA from transfected L and PA cells was analysed by nested RT-PCR. Bands corresponding to exon inclusion or skipping are at 1520 and 1394 nucleotides respectively (the top ϕX*Hae*II marker is 1353, the flanking '1 kb ladder' markers are 1018 and 1635). In this experiment, the degree of exon skipping for NeoTM3 was 11% in L and 43% in PA cells, and for NeoTM2, 1.6% in L and 15% in PA cells. Thus, regulation is maintained when flanking exon sequences and splice sites are substituted and when the regulated exon sequences are changed, although the exon sequences can affect the absolute levels of skipping.

extensive deletion of intron 1 and exon 2, but retained regulation. It seems likely that deletion of intron 3 to the 3′ side of the *Pst*I site 67 nucleotides downstream of exon 3 is responsible for loss of regulation, and that there is therefore a second element essential for negative regulation of TM exon 3 lying downstream. This conclusion was confirmed by further deletion analysis (see below). We refer to these two essential regulatory elements as the URE and DRE (for upstream and downstream regulatory elements).

Specific exon sequences are not essential for regulation

We next tested whether the sequences of TM exon 3 or of the flanking exons 1 and 4 and their associated splice sites were necessary for regulation. To this end, we first analysed the splicing of construct NeoTM3 in L and smooth muscle cells (Figure 3). In this construct the bacterial *Neo*^{r} gene has been split by two hybrid introns flanking TM exon 3. The external splice sites and intron sections are derived from the human β-globin intron 1, while TM exon 3 is flanked by its native intron sequences containing the regulatory elements. Transiently expressed RNA was analysed by reverse transcription followed by nested PCR. We clearly observed smooth muscle-specific regulation of splicing of NeoTM3, with increased skipping of exon 3 in smooth muscle cells (Figure 3). This confirms that the exon sequences and splice sites of TM exons 1 and 4 are not required for the appropriate regulation of TM exon 3.

To test whether the specific TM exon 3 sequences are necessary for regulation, we next constructed NeoTM2, which is identical to NeoTM3 except that the exon 3 sequences and 5′ splice site have been cleanly replaced with those from TM exon 2. We found that although the levels of exon skipping in NeoTM2 were reduced compared with that seen in NeoTM3 in both cell types, the levels of skipping were still significantly higher in smooth muscle compared with L cells (Figure 3). This shows that the TM exon 3 sequences are not absolutely required for smooth muscle-specific regulation, although the degree of exon skipping can be affected by exon sequences. The reduction of exon skipping in NeoTM2 could be due to removal of negative elements in TM exon 3, or to the presence of positive elements in TM exon 2. The recent finding of multiple purine-rich exon splicing enhancer elements (Watakabe *et al.*, 1993) in TM exon 2 (S.Mayer and B.Nadal-Ginard, unpublished observations) supports the latter interpretation, as does the observation that deletion mutations across exon 3 have failed to identify any negative regulatory elements (C.W.J.Smith, unpublished observations).

Mapping the upstream regulatory element

Having ruled out the requirement for specific exon sequences for regulation, we proceeded (i) to confirm the presence of regulatory elements flanking exon 3, and (ii) to map them more precisely by transient transfection of various deletion and substitution mutants based upon the parental regulated construct TS3D which lacks exon 2. Six mutants within intron 2 were initially tested by S1 nuclease protection (Figure 4). Of these, only the original deletion mutant lacking the conserved sequence, TSΔ3D (lane 6), abolished regulation. In T3B2P3, substitution of the branch point and sequences further upstream with the equivalent region upstream of exon 2 had no effect upon regulation (lane 4). Deletion upstream from the *Acc*I site leaving a truncated pyrimidine tract of 17 nucleotides (TΔ17, lane 5) did not knock out regulation, although there was increased skipping of exon 3 in both cell types, presumably due to the weakening of the pyrimidine tract. A 26 nucleotide deletion in TSΔAM covered the 5′ half of the deletion in TSΔ3D. Significantly, this deletion, which left the phylogenetically conserved sequence intact, did not abolish regulation (lane 7). Finally, in intron 2, TSΔRP (lane 8) had a 16 nucleotide deletion just upstream of the 3′ splice site, and TSR2 (lane 9) had clustered point mutations within the 27 nucleotides preceding the 3′ splice site. Again, both of these constructs were still regulated, showing increased skipping of exon 3 in PASM cells. From these data we conclude that there is an essential negative regulatory element (the URE) between 27 and 72 nucleotides upstream of exon 3. The URE contains a 21 nucleotide region to the 3′ side of and including the *Mbo*I site (Figure 1) that is conserved between the rat and chicken genes and which contains three repeats of a UGC motif (Figure 5A; Ruiz-Opazo and Nadal-Ginard, 1987; Lemonnier *et al.*, 1991).

We confirmed the importance of the UGC motif region

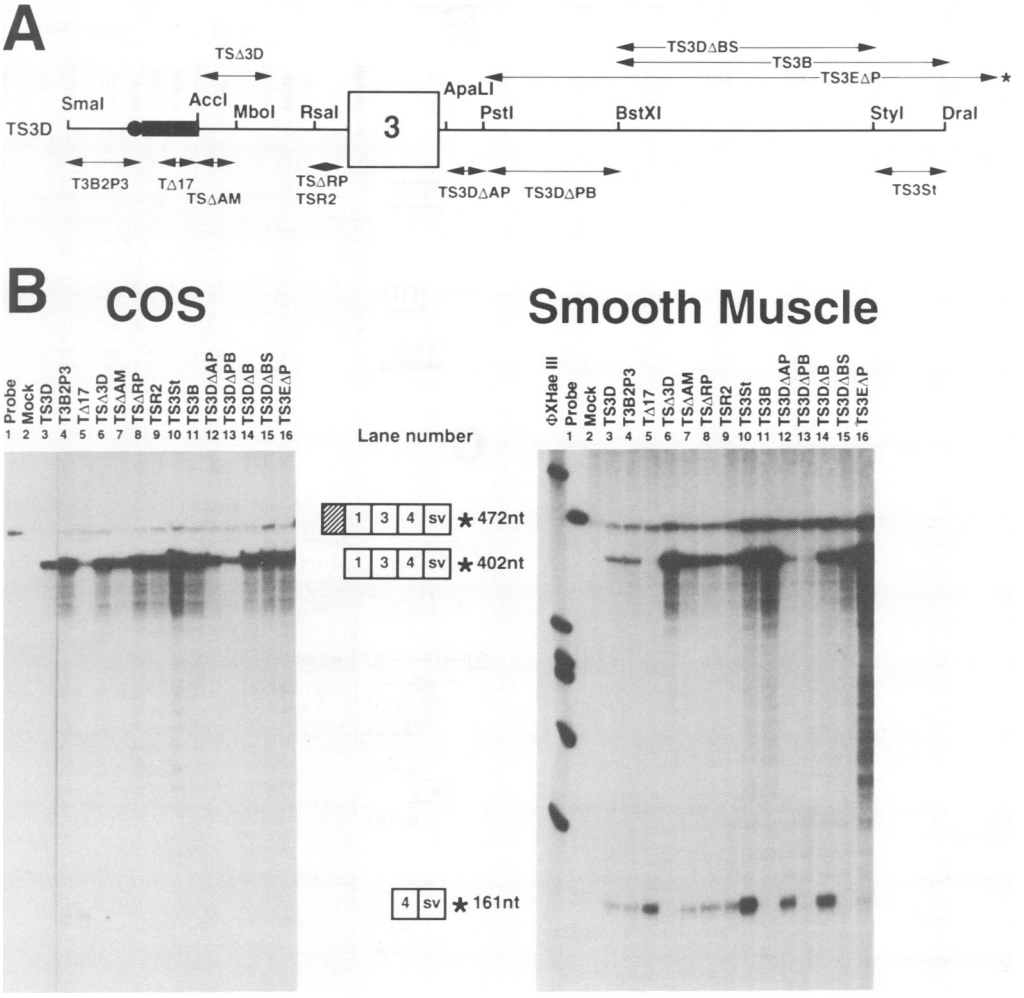

Fig. 4. Negative elements flank TM exon 3. A series of mutant constructs based upon the regulated parental construct TS3D (Figure 2) was tested in transient transfection of COS and PA cells. (**A**) Map of the region flanking TM exon 3 in construct TS3D showing the restriction sites used in the generation of mutants. The branch point and pyrimidine tract of exon 3 are shown by the black circle and rectangle, respectively. The positions of deletions and point mutations made in the various constructs are denoted by the arrows. The 3′ end of the deletion in mutant TS3EΔP is marked with a small asterisk as the deletion extended downstream of the *Dra*I site and the intron contained an additional downstream *Pst*I−*Eco*RI fragment (see text). Mutants that disrupted regulation are shown above the map and those that did not significantly affect regulation are shown below. (**B**) S1 nuclease protection analysis of constructs shown in (A). The conditions were as described in Figure 2, except that only the 134 probe was used. Lanes are numbered identically for the constructs in each panel. Smooth muscle regulation is indicated by increased ratios of partial protection (161 nucleotides), indicative of exon 3 suppression in smooth muscle cells. The data show that an upstream regulatory element lies between the *Mbo*I and *Rsa*I sites of intron 2, and a downstream element lies between the *Bst*XI and *Sty*I sites of intron 3.

by making three further mutations (Figure 5A and B). In UREΔ18 we made an 18 nucleotide deletion that removed the UGC repeats from the parental regulated construct TS3St and left a unique *Apa*I site into which we could clone mutant oligonucleotides. In UREM1 we replaced the UGC repeats with a 23 nucleotide sequence, and in UREM2 we re-inserted an 18 nucleotide sequence containing nine point mutations. These constructs were transfected into L and smooth muscle cells and the expressed RNA analysed by nested RT-PCR. All three mutant constructs showed severely impaired regulation. This confirms the importance of the phylogenetically conserved UGC repeat region of the URE in smooth muscle regulation of splicing.

Mapping the downstream regulatory element

We next proceeded to define more precisely the DRE by making both processive deletions upstream from the *Dra*I site and also internal deletions. All constructs retain the 3′ splice site and branch point region of exon 4. In construct T1-4, in which the loss of regulation was attributed to the deletion downstream of exon 3 (see above, Figure 2), the distance between exons 3 and 4 was reduced to 133 nucleotides. To distinguish whether the loss of regulation was due to deletion of specific sequences or the extreme intron shortening, we made construct TS3EΔP (Figure 4, lane 16). This construct had a deletion between *Pst*I sites downstream of exon 3, but had additional downstream intron 3 sequences resulting in an intron between exons 3 and 4 of 481 nucleotides compared with 514 nucleotides in TS3D. This construct was also fully unregulated. Truncation of the intron from the 3′ side showed that the regulation was retained when deletions were made to a *Sty*I site (TS3St; Figure 4, lane 10) 336 nucleotides downstream of exon 3 (intron length 402 nucleotides). This confirms that the loss of regulation in

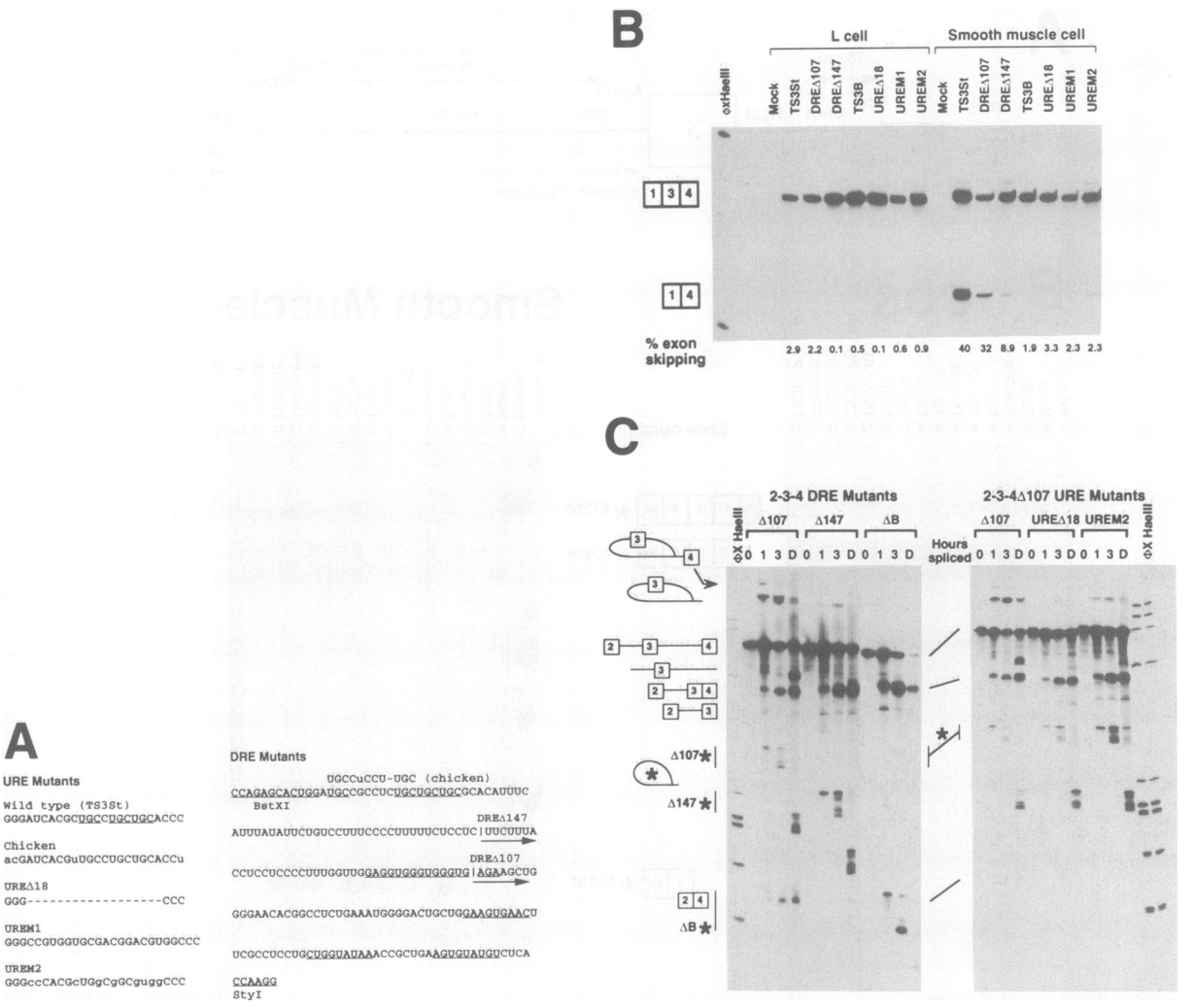

Fig. 5. *In vivo* and *in vitro* skipping of TM exon 3 is dependent upon both the URE and DRE. A series of finer scale mutations within the defined URE and DRE regions were tested for their effects upon exon 3 skipping in transfected cells and in HeLa cell nuclear extracts. (**A**) The sequences of the URE and DRE are shown and compared with the equivalent regions of the chicken gene (Lemonnier *et al.*, 1991); mismatches between chicken and rat are shown in lower case in the chicken sequence. UGC motifs, potential cryptic 5′ splice sites and relevant restriction sites are underlined. Also shown are the URE and DRE mutants based upon the parental regulated construct TS3St (Figure 4). UREΔ18 had an 18 nucleotide deletion of the conserved UGC repeat region of the URE, introducing a unique *Apa*I site. UREM1 and M2 were generated by cloning mutant oligonucleotides into the *Apa*I site. Mismatches in UREM2 are shown in lower case. DRE mutants Δ107 and Δ147 were generated by exonuclease III deletion upstream from the *Sty*I site, as indicated by the arrows. (**B**) RT-PCR analysis of transiently expressed RNA in L and PA cells. Bands corresponding to inclusion or skipping of exon 3 run at 467 and 341 nucleotides, respectively, and are indicated schematically to the side of the gel (the flanking φX*Hae*II markers are at 603 and 310 nucleotides). The degree of exon skipping in each lane, as determined for this experiment, is shown below each lane. (**C**) *In vitro* splicing of transcripts containing exons 2–3–4 with various mutations of the URE and DRE. Splicing was for the times indicated and, in addition, a sample of the 3 h time point was debranched in lanes marked 'D'. The parental transcript 2–3–4Δ107 has the Δ107 deletion of intron 3 displayed in (A) and shows products of both 3–4 and 2–4 splicing. The identities of various bands are shown schematically to the left of the gels. The lariat products of 3–4 splicing vary in size between different transcripts and appear to be subject to some degradation; their positions are indicated by the asterisks and vertical lines to the left of the gels. The transcript marked Δ147 contains the Δ147 deletion shown in (A), while ΔB is deleted to the *Bst*XI site equivalent to TS3B in (B). The right-hand panel shows splicing of 2–3–4Δ107 as well as related transcripts harbouring the UREΔ18 and M2 mutations [see (A)]. In general, the lariat product running above the pre-mRNA is the clearest diagnostic marker for 2–4 splicing. The data show that mutations of the URE or DRE that disrupt *in vivo* regulation (B) also reduce *in vitro* skipping of exon 3 (C).

TS3EΔP was due to deletion of specific intron sequences rather than to a reduction in intron size. Regulation was abolished when the intron was deleted to the *Bst*XI site (TS3B, lane 11), 135 nucleotides downstream of exon 3. Internal deletion between the *Bst*XI and *Sty*I sites (TS3DΔBS, lane 15) also abolished regulation, confirming that an essential regulatory element lies between these sites. A deletion of four nucleotides at the *Bst*XI site had no effect upon regulation (TS3DΔB, lane 14). Construct TS3DΔAP (lane 12) confirmed that deletion of nucleotides nine to 72 downstream of exon 3 did not abolish regulation (Figures 2 and 4). In construct TS3DΔPB (lane 13), the region between the *Pst*I and *Bst*XI sites was deleted. We routinely found that expression from this construct was lower than from all the other constructs and was sometimes difficult to detect by S1 nuclease analysis (Figure 4). Nevertheless, analysis by RT-PCR confirmed that transcripts from this construct were still regulated, with enhanced skipping in smooth muscle cells (data not shown). From these data we conclude that the essential

downstream regulatory element lies between the *Bst*XI and *Sty*I sites, 135–336 nucleotides downstream of exon 3. Notable features within this region are as follows. (i) A short sequence adjacent to the *Bst*XI site with four UGC repeats, similar to the URE (Figure 5A). There is a similar sequence in the chicken gene at the corresponding location, though the phylogenetic conservation is not as striking as with the URE (Figure 5A). (ii) Just downstream of the UGC repeats is a pyrimidine-rich tract of 50–60 nucleotides, including a 27 nucleotides unbroken stretch of pyrimidines. (iii) Between the pyrimidine region and the *Sty*I site there are six sequences with a 5/9 or better match to the consensus 5′ splice site (underlined in Figure 5A). Clustered cryptic 5′ splice sites have been shown to play a role in cell-specific splicing inhibition in the P-element transposase gene (Siebel *et al.*, 1992).

We further narrowed down the necessary sequences of the DRE with constructs DREΔ107 and DREΔ147, which have deletions of 107 and 147 nucleotides, respectively, upstream from the *Sty*I site in intron 3 (Figure 5A and B). Splicing of RNA from these constructs was analysed by nested RT-PCR. Construct DREΔ107 still showed substantial regulation in transfected cells, while DREΔ147 had partially impaired regulation (Figure 5B). As before, deletion to the *Bst*XI site (TS3B) abolished regulation. The DREΔ107 mutant lacks three of the six cryptic 5′ splice sites, while DREΔ147 lacks all six sites and part of the pyrimidine-rich tract just downstream of the *Bst*XI site. This suggests that the pyrimidine tract and/or cryptic 5′ splice sites may have a role in DRE function, though further mutations will be necessary to define precisely their roles (see Discussion).

Skipping of TM exon 3 in vitro is dependent upon URE and DRE sequences

We next proceeded to investigate whether we could observe any skipping of TM exon 3 that was dependent upon the defined regulatory elements in an *in vitro* splicing assay. We first tested transcripts containing TM exons 2, 3 and 4 with 881 nucleotides of the intron immediately downstream of exon 3 and containing the intact regulatory elements flanking exon 3. Since exons 2 and 3 remain mutually exclusive due to steric interference, these transcripts have a simple binary choice between splicing of exon 2 or 3 to exon 4. Although we could see some 3–4 splicing in our initial transcripts, the major *in vitro* splicing event used a cryptic 5′ splice site between the *Bst*XI and *Sty*I sites downstream of exon 3 (data not shown). We therefore tested the effects of the DREΔ107 and Δ147 mutations which lack the cryptic site used *in vitro* (Figure 5C, transcripts 2–3–4Δ107, Δ147). We found that with the 2–3–4Δ107 transcripts the major *in vitro* splicing pattern in HeLa extracts was splicing of exon 3 to 4, and that a significant amount of 2–4 splicing was also observed. Thus exon 3 not only decisively wins the competition between 3′ splice sites in HeLa extract (Mullen *et al.*, 1991), but also has an advantage in the 5′ splice site competition. We analysed the branch point being used by exon 4 in the splicing reaction of transcript 2–3–4Δ107 and mapped a single A residue within a reasonable branch point consensus, CTCTGAG, 23 nucleotides upstream of the 3′ splice site of exon 4 (data not shown). This is a conventional 'consensus' location for a branch point compared with the distant upstream branch points of exons 2 and 3 (Smith and Nadal-Ginard, 1989; Mullen *et al.*, 1991). We found that the amount of the 2–4 products was less in the Δ147 mutant which had partially impaired *in vivo* regulation, and was abolished in the 2–3–4ΔB mutant which is deleted to the *Bst*XI site, a deletion which abolishes *in vivo* regulation (TS3B, Figure 5B). Both of these mutants continued to splice 3–4 very efficiently (Figure 5C). Thus, the 2–4 splicing detectable in HeLa extracts is dependent upon an intact DRE.

We next tested whether mutation of the URE sequences would have a similar effect upon *in vitro* skipping of exon 3. We therefore made 2–3–4Δ107 transcripts that had the UREΔ18 and M2 mutations that abolished smooth muscle regulation *in vivo* (Figure 5B). We found that these mutations also reduced the 2–4 splicing pathway in HeLa extract (Figure 5C). Thus, mutations in the URE cause a reduction of *in vitro* 2–4 splicing that parallels the impairment of *in vivo* regulation. This is an even more striking finding than with the DRE mutants, since mutant M2 has only nine clustered point mutations within the URE compared with the wild type, and these are >140 nucleotides from either of the 5′ splice sites of exons 2 and 3. These data convincingly demonstrate that skipping of TM exon 3 *in vitro* is dependent upon both of the regulatory elements that are necessary for smooth muscle regulation of TM splicing.

Exon switching requires concerted inhibition of both splice sites of exon 3

The 2–3–4 transcripts have a simple 5′ splice site competition between exons 2 and 3, in which exon 3 has the competitive advantage but in which exon 2 is also selected in a significant proportion of transcripts. This is convenient for the analysis of *in vitro* splicing, but it is an artificial situation since endogenous transcripts also have the 3′ splice site competition between exons 2 and 3, and additionally the 2–4 splice is in direct competition with the 1–3 splice. We therefore tested whether our ability to detect significant amounts of 2–4 splicing in the 2–3–4 transcripts was due to the lack of competition from the strong 1–3 splicing event (Mullen *et al.*, 1991) by splicing transcripts containing exons 1–2–3–4 (Figure 6). The transcripts had a 490 nucleotide deletion in intron 1, reducing the length of this intron to 276 nucleotides, as well as the DREΔ107 deletion in intron 3. These transcripts showed abundant bands that corresponded to the products of both 3–4 splicing and 1–3 splicing (Figure 6). In contrast to the 2–3–4 transcripts, there was no evidence of any 2–4 splicing, as indicated by the lack of the diagnostic lariat product for 2–4 splicing. There was a small amount of a product that may correspond to full 1–3–4 splicing. Although this is exactly the same length as a fully spliced 124 product, the lack of any other bands corresponding to 1–2 or 2–4 splicing argues in favour of this band being 134 product. Therefore, the 1–3 splicing event, which removes exon 2 from the transcript, outcompetes any 2–4 splicing in the *in vitro* splicing system. Our ability to observe significant amounts of 2–4 splicing with the 2–3–4Δ107 transcript is most likely due to the lack of stringent competition from the 1–3 splicing event.

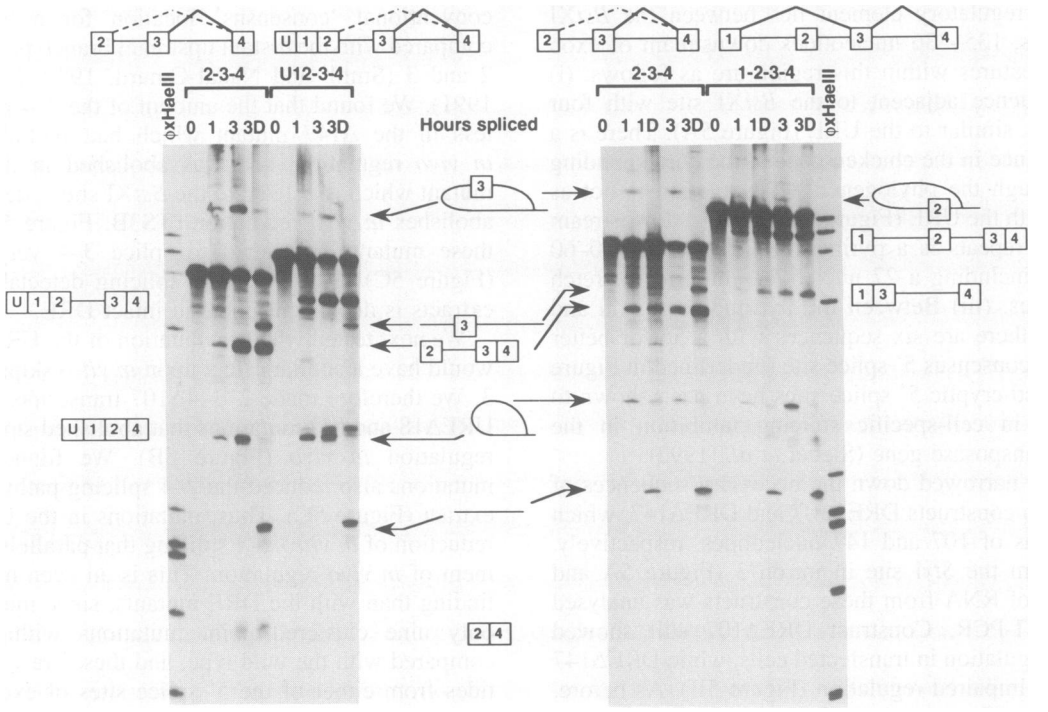

Fig. 6. Exon 2 selection requires inhibition of both splice sites of exon 3. Splicing of transcripts containing exons 2–3–4 with the Δ107 deletion in intron 3 were analysed in HeLa nuclear extract. Splicing was for the times indicated above each lane, with a 'D' denoting debranching of the sample. The structures of the transcripts and the deduced splicing patterns are shown above each set of lanes. The transcript labelled 2–3–4 is the same as the one labelled Δ107 in Figure 5C. Major splicing pathways are shown by solid lines and minor pathways by dashed lines. The structures of some of the diagnostic bands are shown to the sides of the gels. The left panel compares splicing of 2–3–4 transcripts with U12–3–4 transcripts in which exon 1 has been effectively 'pre-spliced' to exon 2. Both transcripts show a mixture of 3–4 and 2–4 splicing products. (NB The diagnostic product bands for U12–3–4 are shown schematically to the left side of the panel.) The right-hand panel compares splicing of 2–3–4 transcripts with 1–2–3–4 transcripts that have intron 1 present (in truncated form). The 1–2–3–4 transcripts show bands diagnostic of both 1–3 and 3–4 splicing, but no bands corresponding to splicing of exon 2. In particular, the diagnostic lariat product band from 2–4 splicing is absent.

This is consistent with the dominant role of the branch point/pyrimidine tract of TM exon 3 in determining the default 134 splicing pattern (Mullen *et al.*, 1991).

We next tested whether inhibition of the normally strong 1–3 splice, allowing 1–2 splicing, would be sufficient to commit the transcript to the 1–2–4 splicing pathway by affecting the subsequent 5′ splice site competition between exons 2 and 3. Such an 'ordered pathway' has been described in the rat β-tropomyosin gene, where splicing of one of a pair of mutually exclusive exons to the upstream common exon is dependent upon first splicing to the downstream constitutive exon (Helfman *et al.*, 1988). To this end we made transcript U12–3–4, which had exon 1, including the complete 5′ untranslated region, effectively 'pre-spliced' to exon 2. We found that transcript U12–3–4 spliced almost indistinguishably from transcript 2–3–4 with mainly 3–4 splicing but some 2–4 splicing (Figure 6, left panel; note that the schematic diagram of U12–3–4 products is to the left of the panel, although lanes with U12–3–4 are to the right). The product of the 3–4 splice has the structure U12–34, retaining the unspliceable intron between exons 2 and 3, which would result in nuclear retention *in vivo* (Smith and Nadal-Ginard, 1989). Prior joining of exon 1 to 2 is therefore unable to dictate the outcome of the subsequent competing 2–4 and 3–4 splicing events. Mutually exclusive splicing in this system does not show the kind of ordered pathway observed in the rat β-tropomyosin gene (Helfman *et al.*, 1988). This result therefore points to the importance of inhibiting both splice sites of exon 3 in a concerted fashion to switch to exon 2 selection. Inhibition of either splice site alone would be insufficient to cleanly switch to selection of exon 2.

Discussion

We have explained previously the strict mutually exclusive nature of exons 2 and 3 of the rat αTM gene (Smith and Nadal-Ginard, 1989), and also the default selection of exon 3 (Mullen *et al.*, 1991). In this paper we have shown that the basis for the selection of exon 2 in smooth muscle cells is the specific inhibition of exon 3. In the absence of exon 2 from test constructs, exon 3 was still skipped with smooth muscle specificity (Figures 2–5). Our data do not rule out the possibility that exon 2 may also be specifically activated in smooth muscle cells. However, given that exon 2 can be fully incorporated into spliced RNA in all cell types if exon 3 is deleted (Mullen *et al.*, 1991), it does not seem necessary to propose such activation to fully explain the tissue specificity of exon selection. Exon 2 has recently been shown to contain purine-rich 'exon splicing enhancer' elements (S.Mayer and B.Nadal-Ginard, unpublished observations; Watakabe *et al.*, 1993). However, there is no evidence that these mediate a cell-specific effect and it is more likely that they play a role in defining the default strength of the

exon 2 splice sites. In support of this, we found that substitution of exon 2 for exon 3 sequences led to a reduction of exon skipping in both smooth muscle and fibroblast cells (Figure 3). Moreover, while in other systems it has been shown that splicing enhancer effects are mediated by SR proteins (Lavigueur *et al.*, 1993; Sun *et al.*, 1993), we found no effect of smooth muscle SR proteins on the 5′ splice site competition between exons 2 and 3 (C.Gooding and C.W.J.Smith, unpublished observations).

We have defined two sequence elements that are necessary for the negative regulation of TM exon 3. The URE lies midway between the polypyrimidine tract and 3′ splice site of TM exon 3. It consists of an ~20 nucleotide conserved sequence containing three UGC (or CUGC) repeats. Point mutations within this region severely impaired regulation (Figure 5). Immediately downstream lies a pyrimidine-rich sequence, although preliminary deletion data suggest that this does not play a major role in regulation (C.W.J.Smith, unpublished observations). The DRE lies >130 nucleotides beyond exon 3. It contains, in 5′ to 3′ order, a short sequence with four UGC motifs (Figure 5A), a pyrimidine-rich tract and a region containing a number of cryptic 5′ splice site sequences. The pyrimidine region and cryptic splice sites may be functionally important, since partial deletion impairs regulation (Figure 5B). The DRE does not show such obvious phylogenetic conservation as the URE (Figure 5A). The chicken gene has a region with some homology to the UGC repeat region of the DRE at an almost identical position, but the pyrimidine-rich region and cryptic splice sites are not present in the chicken gene. Interestingly, we have recently found that the equivalent chicken exons are not well regulated in rat smooth muscle cells, suggesting that this difference in DRE sequences may reflect a genuine mechanistic divergence (G.C.Roberts and C.W.J.Smith, unpublished observations). Given the similar organization of the URE and DRE, it will be of interest to see whether the two elements are interchangeable.

Efficient regulation requires the presence of both elements; deletion of either severely reduces exon 3 skipping. However, we have observed that many URE or DRE mutants still show slightly enhanced skipping of exon 3 in smooth muscle cells (Figure 5B), so it is possible that either element alone can mediate a very low level of tissue-specific exon skipping but that full regulation requires cooperation between the two elements. We are in the process of constructing double mutants of both DRE and URE to see if these show complete abolition of exon skipping.

It may be significant that the regulatory elements are able to mediate a background low level of exon 3 skipping in all cell types (Figures 2, 4 and 5B) and in HeLa nuclear extract (Figure 5C). Using transcripts that retain the 5′ splice site competition between exons 2 and 3 *in vitro*, but lacking competition from the strong 1–3 splice, we were able to see significant levels of exon 3 skipping that were dependent upon both the URE and DRE. However, while deletion of the DRE appeared to totally abolish 2–4 splicing in these transcripts, mutation of the URE reduced but did not completely knock out exon 3 skipping (Figure 5C). This may suggest that the principal function of the URE is in the inhibition of the strong branch point/pyrimidine tract of exon 3. Indeed, it seems likely that the requirement for regulatory elements on both sides of TM exon 3 is related to the need to inhibit both the 5′ and 3′ splice sites to fully switch to exon 2 selection (Figure 6). The low level of regulatory element-dependent skipping of exon 3 in all cells suggests that at least part of the negative regulatory apparatus may be constitutive. Possibly, a low constitutive level of an inhibitory factor is specifically increased only in smooth muscle. Alternatively, regulation may involve a two or more component system, with a smooth muscle-specific factor that is able to potentiate the relatively ineffective constitutive inhibition.

Comparison with other TM genes

The rat (Helfman *et al.*, 1988, 1990; Guo *et al.*, 1991; Mulligan *et al.*, 1992) and chicken βTM genes (Clouet-d'Orval *et al.*, 1991; Libri *et al.*, 1991, 1992; Gallego *et al.*, 1992), and the human α_sTM gene (Graham *et al.*, 1992) have also been used to study mutually exclusive splicing. None of these genes is the homologue of the rat αTM gene used here, and the pair of mutually exclusive exons studied in each case corresponds to a more 3′ part of the gene, in which the upstream exon is selected by default while the downstream exon is skeletal muscle-specific. In these cases it is found that the skeletal muscle exon, although having an extensive polypyrimidine tract, is repressed in non-striated muscle cells. With the chicken βTM gene, it has been shown that competition between the two exons based on branch point/pyrimidine tract strengths can occur if the downstream exon is mutationally derepressed (Libri *et al.*, 1992). The default repression of the βTM skeletal exon involves a sequence element in the upstream intron between the polypyrimidine tract and the 3′ splice site (Helfman *et al.*, 1990; Gallego *et al.*, 1992; Libri *et al.*, 1992), and also elements within the exon itself (Guo *et al.*, 1991; Libri *et al.*, 1991). In the chicken gene a short hairpin structure within the exon is essential for repression (Libri *et al.*, 1991). In the human α_sTM gene sequences within the muscle-specific exon are sufficient for mediating its repression in non-muscle cells (Graham *et al.*, 1992). There do not appear to be any essential elements for repression within the downstream intron of the rat βTM (Guo *et al.*, 1991), although there is the potential for extensive RNA secondary structure formation between the two introns flanking the chicken skeletal muscle exon (Clouet-d'Orval *et al.*, 1991). In contrast, there is no obvious conserved secondary structure that would form involving the URE and DRE in αTM. These genes all have mechanism(s) for repressing the muscle-specific exon as a default condition; tissue-specific regulation then involves derepression by an as yet uncharacterized mechanism. In contrast, the rat αTM default exon 3 is regulated by strong repression only in smooth muscle. The upstream element in the βTM gene is in a similar position to the URE that we have located. There is no obvious sequence similarity between the two elements, although both are relatively enriched in pyrimidines (Helfman *et al.*, 1990). In the rat βTM, the hnRNP protein PTB (Patton *et al.*, 1991) has been found to bind to the intron regulatory element and may be associated with the default repression of the muscle exon (Guo *et al.*, 1991; Mulligan *et al.*, 1992). We have found no evidence for

PTB binding to the conserved UGC core of the URE, but we have observed it cross-linking to the pyrimidine-rich regions within the DRE and adjacent to the URE (C.Gooding and C.W.J.Smith, unpublished observations). We are investigating the possibility that PTB may have a role in the low level constitutive skipping of TM exon 3.

Comparison with other negative splicing regulation

Other examples of negative regulation of splicing may be instructive in considering the regulation of αTM. The *Drosophila* P-element transposase gene has an intron that is spliced efficiently in germline but not somatic cells. Splicing is inhibited in somatic cells by the assembly of non-productive complexes containing U1snRNP and a group of somatic cell proteins on two pseudo 5′ splice sites just upstream of the authentic 5′ splice site (Siebel *et al.*, 1992). The 200 nucleotide part of the intron downstream of TM exon 3, bounded by the *Bst*XI and *Sty*I sites that contains the DRE, also has six sequences with a match of 5/9 or better to the 5′ splice site consensus (Figure 5A). Three of these are deleted in the DREΔ107 mutant which is still regulated, though with slightly lower levels of exon 3 skipping than the parental TS3St construct (Figure 5B). It does retain the sequence GAGGUGGGUGGGUGAGA, which has three overlapping potential cryptic 5′ splice sites with matches of 7/9, 6/9 and 6/9 to the consensus. We have some deletion mutations that are unregulated but which retain all of the cryptic 5′ splice sites (C.W.J.Smith, unpublished data), so clearly they are not sufficient for DRE function. Nevertheless, it is possible that they may play a role in regulation. It remains a challenge to explain how cryptic 5′ splice sites could be active at such a large distance from the inhibited splice sites of exon 3.

The *Drosophila* sex-lethal (Sxl) protein is also an inhibitory regulator of splicing of both its own pre-mRNA and that of the *tra* gene. Sxl protein binds to a specific U_8C sequence within the polypyrimidine tract of the regulated 3′ splice site of *tra*. This blocks binding of U2AF to that site and redirects it to the downstream female-specific site (Valcarcel *et al.*, 1993). Auto-regulation of *sxl* splicing is more complex as it involves the repression by Sxl protein of splicing of a discrete exon, as for TM regulation. Moreover, in Sxl auto-regulation, although the 3′ splice site of the repressed exon has a U_8C sequence within the pyrimidine tract, it is not absolutely required for regulation since there are a series of functionally redundant U-rich sequences flanking the regulated exon, from ~900 nucleotides upstream to ~300 nucleotides downstream (Sakamoto *et al.*, 1992; Horabin and Schedl, 1993). It is conceivable that Sxl binding to these elements over a wide region of the pre-mRNA completely masks the regulated exon in female cells. Such a mechanism contrasts with the more precise targeted inhibition of *tra* pre-mRNA by Sxl (Valcarcel *et al.*, 1993). Exon 3 of αTM also has flanking regulatory elements, but unlike *sxl* pre-mRNA, where both splice sites of the regulated exon are flanked by the regulatory U-rich sequences, the regulated TM branch point/pyrimidine tract is upstream of both the characterized regulatory sequences. Moreover, the negative regulation of TM must be precisely targeted since assembly of splicing factors at the pyrimidine tract and branch point of exon 3, 50–100 nucleotides upstream of the URE, must be impeded, while the 5′ splice site of exon 2 only 41 nucleotides further upstream remains functional.

How could the TM regulatory elements mediate such precise targeting? One possibility is that proteins bound to the URE and DRE interact, looping out the intervening RNA containing exon 3. If so, this might have the effect of bringing the cluster of cryptic 5′ splice sites into proximity with the branch point/pyrimidine tract of exon 3. This could possibly result in the formation of non-productive complexes between splicing factors bound to the exon 3 branch point/pyrimidine tract and cryptic 5′ splice sites, preventing productive spliceosome formation. Such a model could explain how the regulatory elements effect repression at the branch point of exon 3, but not at the nearby 5′ splice site of exon 2. Repression of the 5′ splice site of exon 3 may involve a similar mechanism, although the precise targeting does not need to be invoked since there are no nearby splice sites that must remain active. More precise dissection of the essential DRE sequences should help to test this model.

Whatever the precise mechanism, we expect that the negative regulation of TM splicing is mediated by binding of proteins to the URE and DRE sequences. We have detected a number of proteins by UV cross-linking to short RNAs containing the conserved core of the URE (G.Moreau and C.Gooding, unpublished observations). Binding of some of these proteins appears to be sequence-specific. We are currently testing the activity of these proteins as well as other protein fractions from smooth muscle tissue in the *in vitro* alternative splicing assay.

Materials and methods

Constructs

Constructs used as templates for *in vitro* transcription or for transfection were all prepared by standard cloning procedures (Sambrook *et al.*, 1989). Constructs for transfection used a pUC-derived vector containing promoter, enhancer, origin of replication and 3′ end processing sequences from SV40, and tropomyosin exons 1 and 4 fused to the SV40 sequences at the *Nco*I site at codon 1 of exon 1, and at the *Bal*I site of exon 4. The basic constructs have been described previously (Mullen *et al.*, 1991). Most constructs were derived from a parental construct, pTMSR, lacking exons 2 and 3, with a single *Eco*RI site introduced at the *Sac*I site within intron 1, and at the *Rsa*I site just upstream of exon 4 (Mullen *et al.*, 1991). This allowed the re-insertion of various fragments containing exons 2 and/or 3 as linker-ligated *Eco*RI fragments. The main parental constructs are pTS3D (which has the exon 3-containing *Sma*I–*Dra*I fragment) and pTS23D (which has the exon 2- and 3-containing *Sac*I–*Dra*I fragment). Constructs TSΔ3D and TS2Δ3D contained a deletion of 49 nucleotides downstream from the *Acc*I site between exons 2 and 3. This deletion was introduced using the construct Δ44-157 (Smith *et al.*, 1989b). TS3DΔAP and TS23DΔAP had a deletion between the *Apa*LI and *Pst*I sites downstream of exon 3. T1-4 had a 732 nucleotide *Sma*I–*Pst*I insert containing exons 2 and 3 inserted into pTMSR. TSR2 was derived from the construct pR2 (Smith *et al.*, 1989b) and had the sequence immediately upstream of exon 3 mutated by 12 point mutations to GACGCGTACGGCACGGAAAACGAACAG. Other constructs in Figure 4 were derived from TS3D by simple deletions and substitutions as described in the text and figure legends. Constructs pNeoTM3 and pNeoTM2 contained the bacterial neomycin resistance gene interrupted by a β-globin-derived intron at the *Bal*I site. This intron contained a unique *Xho*I site just upstream of the branch point into which the *Sma*I–*Dra*I exon 3-containing fragment was inserted. For pNeoTM2 a related fragment, with the exon 3 sequences and 5′ splice site replaced with those from TM exon 2, was inserted. The exon 2 fragment was cloned from the substitution mutant $1_{B3P3}23$ (Mullen *et al.*, 1991) and linked to the intron downstream of exon 3 by using the

natural *Kpn*I site adjacent to exon 2 and attaching *Kpn*I linkers to the *Apa*LI site just downstream of exon 3. Construct UREΔ18 was obtained by reverse PCR using TS3St as the parental template. The deletion was verified by sequencing. UREM1 and M2 were generated by cloning oligonucleotides into the unique *Apa*I site of UREΔ18 as shown in Figure 5A. DREΔ107 and Δ147 were generated by exonuclease III deletion using TS3St as a parental construct. Unique *Bgl*II and *Sac*I sites were first introduced adjacent to the *Sty*I site in intron 3, and exonuclease deletions were then made in a 5′ direction with the *Sac*I site acting as the block to 3′ deletion. A number of deletion mutants were identified by sequencing. Transcripts for *in vitro* splicing were cloned in pGem vectors (Promega). The 2–3–4Δ107, Δ147 and ΔB constructs were cloned into the polylinker *Hin*dIII site via *Hin*dIII linkers on the *Xho*I site of exon 2 and the *Bal*I site of exon 4. Deletions within intron 3 were the same as for the transfected constructs. Construct U12–3–4 was derived from 2–3–4Δ107 by linking exon 2 to a cDNA containing the complete exon 1 including the full 5′ untranslated sequence. Construct 1–2–3–4 was generated from 2–3–4Δ107 by the addition of exon 1 and intron 1 sequences from the *Nco*I site at codon 1 of exon 1 to the *Kpn*I site just after exon 2. A deletion of 490 nucleotides between *Dra*I and *Dra*III sites in intron 1 was also present in this construct. Further details of constructs can be obtained upon request from the authors.

In vitro transcription and splicing reactions

^{32}P-labelled RNA transcripts were transcribed from pGEM vectors with SP6 (Pharmacia) or T7 RNA polymerase using a G(5′)ppp(5′)G dinucleotide primer as described previously (Smith and Nadal-Ginard, 1989; Mullen *et al.*, 1991). HeLa cell nuclear extracts were prepared using the modifications of Abmayr *et al.* (1988). Standard splicing reactions contained 20–50 fmol [^{32}P]RNA transcript, 2 mM $MgCl_2$, 500 μM ATP, 20 mM creatine phosphate, 2500 U/ml RNasin, 12 mM Tris (pH 7.9), 12% (v/v) glycerol, 60 mM KCl, 0.12 mM EDTA, 0.3 mM DTT, 2.6% polyvinyl alcohol and 60% nuclear extract, and were incubated at 30°C. Departures from standard conditions are noted in the figure legends. Splicing reactions were then subjected to proteinase K digestion and phenol/chloroform extraction. Further characterization of reaction intermediates and products was carried out by debranching in HeLa cell S-100 cytoplasmic fraction at 30°C for 30 min. Reaction products were analysed by electrophoresis in 8 M urea, 4% polyacrylamide gels, followed by autoradiography.

Cell culture and transfection

COS, HeLa, L and PA cells were all grown under standard conditions in DMEM supplemented with 10% fetal calf serum. Transient transfection was carried out by calcium phosphate coprecipitation as described previously (Smith and Nadal-Ginard, 1989; Mullen *et al.*, 1991).

Detection of expressed RNA

Transiently expressed RNA was analysed by either S1 nuclease protection (Figures 2 and 4) or nested reverse transcription followed by RT-PCR (Figures 3 and 5B). Cytoplasmic RNA was harvested by the hot phenol method, and conditions and probes used for S1 analysis were as described previously (Smith and Nadal-Ginard, 1989; Mullen *et al.*, 1991). For RT-PCR (Figures 3 and 5B), 2–5 μg of cytoplasmic RNA was denatured at 65°C for 5 min, snap cooled on ice and then mixed with 50 ng of RT primer (SV3′3) in the presence of 2 mM dNTPs, 500 mM Tris−HCl, pH 8.3, 60 mM $MgCl_2$ and 400 mM KCl. Annealing was at 42°C for 2 h, after which ~10 U of AMV reverse transcriptase (Promega) were added, and incubation was carried out for a further 45 min at 42°C. For PCR, 2 μl of the RT reaction were used as template in 25 μl reactions. PCR buffer was as supplied by Boehringer, but with $MgCl_2$ at 2.5 mM. Unlabelled 5′ primer SV5′1 was added to 50–100 ng, and 5′ ^{32}P-labelled 3′ primer SV3′1 to 5–10 ng. A hot start protocol was used in which the labelled 3′ primer and Taq polymerase (1 U; Boehringer) were added to the remainder of the reaction at 80°C. Cycling conditions were: 94°C, 30 s, 62°C, 30 s, 72°C, 1 min, 30 cycles. Samples were then phenol/chloroform-extracted, ethanol-precipitated and separated on 8 M urea 4–6% polyacrylamide gels. The relative ratios of bands in PCRs were determined using a Molecular Dynamics Phosphorimager. In control reactions we determined that the relative ratios of bands corresponding to inclusion and exclusion of TM exon 3 did not vary widely over 20–35 cycles of PCR. Moreover, the fact that RT-PCR results agreed with S1 nuclease data gives additional confirmation that amplification has not distorted significantly the ratios of bands. Sequences of oligonucleotides used for RT-PCR are: SV3′3, GCAAACTCAGCCACAGGTCTGTACC; SV3′1, CTCACTGCGTTCCAGGCAATGCT; and SV5′1, GAGCTATTCCAGAAGTAGTGAGGAG.

Acknowledgements

We thank Tammy Chu for technical assistance and Steve Mayer for critically reading the manuscript. This work was supported by grants from the British Heart Foundation to C.W.J.S. and from the National Institutes of Health and Muscular Dystrophy Association of America to B.N.-G.

References

Abmayr,S.M., Reed,R. and Maniatis,T. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 7216–7220.
Clouet-d'Orval,B., d'Aubenton-Carafa,Y., Sirand-Pugnet,P., Gallego,M., Brody,E. and Marie,J. (1991) *Science*, **252**, 1823–1828.
Gallego,M.E., Balvay,L. and Brody,E. (1992) *Mol. Cell. Biol.*, **12**, 5415–5425.
Graham,I.R., Hamshere,M. and Eperon,I.C. (1992) *Mol. Cell. Biol.*, **12**, 3872–3882.
Guo,W., Mulligan,G.J., Wormsley,S. and Helfman,D.M. (1991) *Genes Dev.*, **5**, 2096–2107.
Helfman,D.M., Ricci,W.M. and Finn,L.A. (1988) *Genes Dev.*, **2**, 1627–1638.
Helfman,D.M., Roscigno,R.F., Mulligan,G.J., Finn,L.A. and Weber,K.S. (1990) *Genes Dev.*, **4**, 98–110.
Horabin,J.I. and Schedl,P. (1993) *Mol. Cell. Biol.*, **13**, 7734–7746.
Lamond,A. (1993) *BioEssays*, **15**, 595–603.
Lavigueur,A., La Branche,H., Kornblihtt,A.R. and Chabot,B. (1993) *Genes Dev.*, **7**, 2405–2417.
Lear,A.L., Eperon,L.P., Wheatley,I.M. and Eperon,I.C. (1990) *J. Mol. Biol.*, **211**, 103–115.
Lees-Miller,J.P., Goodwin,L.O. and Helfman,D.M. (1990) *Mol. Cell. Biol.*, **10**, 1729–1742.
Lemonnier,M., Balvay,L., Mouly,V., Libri,D. and Fiszman,M. (1991) *Gene*, **107**, 229–240.
Libri,D., Piseri,A. and Fiszman,M. (1991) *Science*, **252**, 1842–1845.
Libri,D., Balvay,L. and Fiszman,M. (1992) *Mol. Cell. Biol.*, **12**, 3204–3215.
Mayeda,A. and Krainer,A.R. (1992) *Cell*, **68**, 365–375.
Moore,M.J., Query,C.C. and Sharp,P.A. (1993) In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
Mullen,M.P., Smith,C.W.J., Patton,J.G. and Nadal-Ginard,B. (1991) *Genes Dev.*, **5**, 642–655.
Mulligan,G.J., Guo,W., Wormsley,S. and Helfman,D.M. (1992) *J. Biol. Chem.*, **267**, 25480–25487.
Parker,R., Siliciano,P.G. and Guthrie,C. (1987) *Cell*, **49**, 229–239.
Patton,J.G., Mayer,S.A., Tempst,P. and Nadal-Ginard,B. (1991) *Genes Dev.*, **5**, 1237–1251.
Patton,J.G., Porro,E.B., Galceran,J., Tempst,P. and Nadal-Ginard,B. (1993) *Genes Dev.*, **7**, 393–406.
Reed,R. and Maniatis,T. (1988) *Genes Dev.*, **2**, 2168–2176.
Rothman,A., Kulik,T.J., Taubman,M.B., Berk,B.C., Smith,C.W.J. and Nadal-Ginard,B. (1992) *Circulation*, **86**, 1977–1986.
Ruiz-Opazo,N. and Nadal-Ginard,B. (1987) *J. Biol. Chem.*, **261**, 4755–4765.
Ruskin,B., Zamore,P.D. and Green,M.R. (1988) *Cell*, **52**, 207–219.
Sakamoto,H., Inoue,K., Higuchi,I., Ono,Y. and Shimura,Y. (1992) *Nucleic Acids Res.*, **20**, 5533–5540.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual*. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Siebel,C.W., Fresco,L.D. and Rio,D.C. (1992) *Genes Dev.*, **6**, 1386–1401.
Smith,C.W.J. and Nadal-Ginard,B. (1989) *Cell*, **56**, 749–758.
Smith,C.W.J., Patton,J.G. and Nadal-Ginard,B. (1989a) *Annu. Rev. Genet.*, **23**, 527–577.
Smith,C.W.J., Porro,E.B., Patton,J.G. and Nadal-Ginard,B. (1989b) *Nature*, **342**, 243–247.
Sun,Q., Mayeda,A., Hampson,R.K., Krainer,A.R. and Rottman,F.M. (1993) *Genes Dev.*, **7**, 2598–2608.
Tian,M. and Maniatis,T. (1992) *Science*, **256**, 237–240.
Tian,M. and Maniatis,T. (1993) *Cell*, **74**, 105–114.
Valcarcel,J., Singh,R., Zamore,P.D. and Green,M.R. (1993) *Nature*, **362**, 171–175.
Watakabe,A., Tanaka,K. and Shimura,Y. (1993) *Genes Dev.*, **7**, 407–418.
Wieczorek,D.F., Smith,C.W.J. and Nadal-Ginard,B. (1988) *Mol. Cell. Biol.*, **8**, 679–694.

Zahler,A.M., Neugebauer,K.M., Lane,W.S. and Roth,M.B. (1993) *Science*, **260**, 219–222.
Zamore,P.D., Patton,J.G. and Green,M.R. (1992) *Nature*, **355**, 609–614.
Zhuang,Y. and Weiner,A.M. (1986) *Cell*, **46**, 827–835.
Zhuang,Y., Goldstein,A.M. and Weiner,A.M. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 2752–2756.

Received on May 4, 1994; revised on June 9, 1994