Regulated Tissue-specific Alternative Splicing of Enhanced Green Fluorescent Protein Transgenes Conferred by α-Tropomyosin Regulatory Elements in Transgenic Mice*

Peter D. Ellis, Christopher W. J. Smith‡, and Paul Kemp§

From the Department of Biochemistry, Tennis Court Road, University of Cambridge, Cambridge CB2 1QW, United Kingdom

The mutually exclusive exons 2 and 3 of α-tropomyosin (αTM) have been used as a model system for strictly regulated alternative splicing. Exon 2 inclusion is only observed at high levels in smooth muscle (SM) tissues, whereas striated muscle and non-muscle cells use predominantly exon 3. Experiments in cell culture have shown that exon 2 selection results from repression of exon 3 and that this repression is mediated by regulatory elements flanking exon 3. We have now tested the cell culture-derived model in transgenic mice. We show that by harnessing the intronic splicing regulatory elements, expression of an enhanced green fluorescent protein transgene with a constitutively active promoter can be restricted to SM cells. Splicing of both endogenous αTM and a series of transgenes carrying regulatory element mutations was analyzed by reverse transcriptase-PCR. These studies indicated that although SM-rich tissues are equipped to regulate splicing of high levels of endogenous or transgene αTM RNA, other non-SM tissues such as spleen, which express lower amounts of αTM, also splice significant proportions of exon 2, and this splicing pattern can be recapitulated by transgenes expressed at low levels. We confirm the importance in vivo of the negatively acting regulatory elements for regulated skipping of exon 3. Moreover, we provide evidence that some of the regulatory factors responsible for exon 3 skipping appear to be titratable, with loss of regulated splicing sometimes being associated with high transgene expression levels.

Alternative splicing is a major form of gene regulation in multicellular eukaryotes. It allows the repertoire of expressed proteins to far exceed the number of available genes by the generation from single genes of multiple mRNAs encoding functionally distinct proteins (1–3). Bioinformatic analyses coupling expressed sequence tag and genomic sequence data bases (4) and, more recently, splice-sensitive microarrays (5) have indicated that most human genes are alternatively spliced. In addition to allowing generation of qualitatively distinct proteins, a large number of human alternative splicing events lead to quantitative regulation of gene expression by generating RNAs that are substrates for degradation by non-sense-mediated decay (NMD)) (6). Alternative splicing is therefore of interest as a basic mechanism that allows regulated gene expression. It is likewise of interest as a technical tool that could be used to provide regulated restriction of transgene expression, perhaps in combination with regulated promoters. Most studies of regulated mammalian alternative splicing have used tissue culture lines to analyze splicing of transfected minigene constructs, in combination with in vitro analysis in cell extracts (7). Although this approach has allowed much progress to be made in the characterization of regulatory RNA sequences and the trans-acting factors with which they interact, it is an approach with some limitations. For example, many alternative splicing events are regulated in tissue types for which there are no reliable cultured cell lines. Smooth muscle is a good example of such a tissue. Smooth muscle cells are important in various locations in the body, and smooth muscle dysfunction underlies a number of pathologies (8). Vascular smooth muscle is responsible for controlling blood flow and systemic blood pressure. Alterations in the state of differentiation of vascular smooth muscle cells (VSMCs) from a contractile to a proliferative phenotype are associated with the development of atherosclerotic plaques and restenosis (9). The transition between the normal contractile phenotype and the proliferative VSMC phenotype is accompanied by various alterations in gene expression at the transcriptional and splicing levels (8). Despite the biomedical importance of SMCs, there are no cultured SMC lines that reliably maintain the differentiated phenotype.

We have previously used the α-tropomyosin (TM) gene as a model system for regulated splicing (10–17) (Fig. 1). Exons 2 and 3 are spliced in a mutually exclusive manner, with predominant inclusion of exon 2 only in smooth muscle tissues, whereas exon 3 is used almost exclusively in striated muscles (cardiac and skeletal) and most other tissues where αTM expression could be observed (18, 19), including proliferating VSMCs (20). Likewise, transfection of minigene constructs into almost all cell types resulted in predominant or exclusive inclusion of exon 3 (10). The exception was in PAC-1 cells, a VSMC line derived from embryonic rat pulmonary artery (11, 21). The combination of restriction of exon 2 selection to smooth muscle tissues and PAC-1 cells for transfected constructs led to the concept that exon 3 represents the “default” splicing pathway, with exon 2 selection resulting only from specific regulation in SM cells.

Received for publication, May 14, 2004
Published, JBC Papers in Press, June 11, 2004, DOI 10.1074/jbc.M405380200

* This work was supported by Project Grant PG/2001031 from the British Heart Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
† To whom correspondence may be addressed. Tel.: 44-1223-333655; Fax: 44-1223-766002; E-mail: cwa1@cam.ac.uk.
‡ To whom correspondence may be addressed. Tel.: 44-1223-333627; Fax: 44-1223-766002; E-mail: pk@mol.e.bio.cam.ac.uk.
§ To whom correspondence may be addressed. Tel.: 44-1223-333627; Fax: 44-1223-766002; E-mail: pk@mol.e.bio.cam.ac.uk.

1 The abbreviations used are: NMD, nonsense-mediated decay; αTM, α-tropomyosin; SM, smooth muscle; EGFP, enhanced green fluorescent protein; RT, reverse transcriptase; Q-RT-PCR, quantitative real time PCR; URE, upstream regulatory element; DRE, downstream regulatory element; GFP, green fluorescent protein; VSMCs, vascular smooth muscle cells; PTB, polypyrimidine tract-binding protein; nt, nucleotide.
PAC-1 cells typically show ∼20–50% inclusion of exon 2 in both their endogenous αTM gene and in transiently expressed minigene, and have therefore facilitated characterization of the regulatory mechanisms responsible for switching selection between αTM exons 2 and 3 (11–15). The regulatory switch involves inhibition of exon 3 and does not require the presence of exon 2. Regulation is mediated by four sequence elements (11). The polypyrimidine tract of exon 3 (P3), one of the essential 3’ splice site elements, contains three optimal binding sites for polypyrimidine tract-binding protein (PTB), a splicing repressor (22) (Fig. 1B). Midway between the P3 pyrimidine tract and exon 3, which are separated by an unusually long distance (100 nt), there is an upstream regulatory element (URE) consisting of a short cluster of UGC (or CUG) repeats (11, 13). Downstream of the exon there is an essential downstream regulatory element (DRE). The DRE is composed of both a short stretch of UGC/CUG motifs, known as DUGC and a pyrimidine tract (DY). Whereas both the P3 and DY pyrimidine tracts are known to bind the repressor protein PTB (12, 14, 16, 22), the factors that bind to the URE and DUGC elements are, as yet, uncharacterized. These regulatory elements are capable of conferring PAC-1 cell-regulated splicing to heterologous exons (15), suggesting that they may be sufficient to restrict gene expression in transgenic animals.

Whereas PAC-1 cells have proven very useful in characterizing these regulatory elements, their partial and variable state of differentiation calls into question the extent to which these findings can be generalized, especially in other smooth muscle tissues in vivo. We therefore decided to analyze control of αTM alternative splicing in transgenic mice, using reporter constructs based on enhanced green fluorescence protein (EGFP), in which EGFP expression was dependent upon the regulated exclusion of the αTM exon. Our data confirm the importance of the previously characterized regulatory elements, but also challenge the simple concept of a regulated “smooth muscle” splicing pattern and a default splicing pattern elsewhere. Finally, our data demonstrate that the intrinsic elements flanking αTM exon 3 can be harnessed to drive highly tissue-restricted gene expression in transgenic animals, and so could be a useful technical tool, especially in conjunction with tissue-restricted promoters with overlapping specificity.

MATERIALS AND METHODS

Plasmids and Cloning—The vector pCAGGS was a generous gift of Jun-ichi Miazaki (University of Osaka). To ensure appropriate initiation of translation, the Kozak sequence in the EGFP gene of pCAGGS was converted to a consensus sequence (GCCACCATGG) by PCR. The introns were inserted between bases 5 and 6 of the EGFP open reading frame in a two-step process. First, a short sequence containing a consensus 5’ splice site and 3’ splice site separated by unique AflII and SmaI sites was inserted. pGTM2 and pGTM3 were generated by amplification of the α-TM exon 2 or 3 containing cassettes from pNeoTM2 and pNeoTM3, respectively (15). In the case of pGTM2, the XhoI site at the 5’ end of exon 2 was filled, resulting in a 4-bp frameshifting insertion. In pGTM3 a G to A mutation at nt 106 of exon 3 introduced a TGA stop codon in the EGFP open reading frame. The P3Δ123, ΔURE, ΔDRE, and 3URE mutations were cassette-cloned into the pGTM3 background from the mutant TM constructs reported previously (12–14).

Generation of Transgenic Mice—Transgenic mice were generated by pronuclear injection of Sal I-linearized plasmid DNA into CBA × C57Bl/6 fertilized eggs. Viable offspring were genotyped by PCR of tail samples for the presence of the EGFP transgene.

Isolation of Tissues and RNA—Tissues were harvested from 8- to 10-week-old transgenic mice killed by CO₂ asphyxiation. To ensure consistent parts of each tissue were used, the same dissection and tissue preparation protocol was used for all animals. All tissues were placed into RNA-Later solution immediately after dissection except for the heart and aorta which were removed as a single unit and placed into phosphate-buffered saline. Excess fatty tissue around the junction of the heart and aorta was removed before the aorta was cut away from the heart. The aorta was stripped of fat before immersion in RNA-Later solution. The aorta were removed, and the ventricles were placed into RNA-Later. All tissue was stored at −20 °C prior to processing. RNA was isolated using the RN-Easy kit according to the manufacturer’s instructions. The RNA was quantified by fluorescence using the ribogreen kit (Molecular Probes).

RT-PCR—250 ng of RNA was reverse-transcribed in a 20-μl reaction using SuperScript II (Invitrogen) according to the manufacturer’s instructions. After heating to 70 °C to inactivate the enzyme, the reactions were diluted to 100 μl. For PCR detection of GTM2 splicing (Fig. 2B), 0.5 μl of diluted RT reaction was used in 25-μl PCRs with a 1 mm MgCl₂-containing buffer. An 80 °C hot start was followed by 30 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s with the primers 5’-GCCGAAGCTCGCCCAAACTT-3’ and 5’-GGGTGGCTCCTCAAGCTT-3’. Exon inclusion gave a band of 506 bp, and exon skipping gave a band of 376 bp. For endogenous α-tropomyosin splicing (Fig. 2A), the PCRs used 3 μl of RT template in 25-μl reactions. An 80 °C hot start was followed by 40 cycles of 94 °C for 30 s, 63 °C for 30 s, and 72 °C for 30 s with the primers mTM1 (5’-AGAAAGCCCTT-GGATCGA-3’) and mTM4 (5’-AACAAGCCCATGCCATCACTC-3’). PCR products were then phenol/chloroform-extracted, ethanol-p precipitated, and then digested with XhoI, which cuts specifically in exon 2. exon 3 products were 204 bp, whereas exon 2 products were 150 and 54 bp.

Quantitative Real Time PCR—Q-RT-PCR was carried out using SYBR green detection for the endogenous α-tropomyosin transcripts and spliced EGFP transcripts.

For the endogenous α-tropomyosin transcripts, 5 μl of RT product was added to a mixture containing 12.5 μl of SYBR Green Master Mix (Applied Biosystems), 2.5 μl of primer E (10 pmol/μl) or F (10 pmol/μl), and 2.5 μl of primer G (10 pmol/μl) in a final volume of 25 μl. Primer E detected transcripts including α-TM exon 3 and had the sequence CA-AATCTCCGAGGCTCTCAAAG. Primer F detected transcripts α-TM exon 2 and had the sequence GCTGGAGGAGCTGCACAAG. Primer G had the sequence GCCATCCAACTCTCTCTCACA. Samples were incubated at 50 °C for 2 min and heated to 95 °C for 10 min. PCR was performed at 40 cycles of 95 °C for 15 s and 60 °C for 1 min in an ABI 7000.

Ribosomal RNA was quantified in parallel reactions using 12.5 μl Universal Master Mix (Applied Biosystems), 0.125 μl of RNA forward primer, 0.125 μl of RNA reverse primer, 0.125 μl of VIC-labeled RNA Taqman probe, and 5 μl of cDNA in a total volume of 25 μl.

Each PCR was shown to produce a single correct PCR product with no primer dimer formation. To test the ability of each PCR to detect the correct transcript in the presence of the alternative form of α-tropomyosin cDNA, mixtures of cDNAs consisting of α-TM exons 1-2-4 cDNA was 100, 75, 50, 25, and 0% with the remainder being α-TM exons 1-3-4 were prepared from known concentrations of each cDNA. The total concentration of DNA in these mixtures ranged from 100 to 0.1 fg, and the percentage of α-TM exons 1-2-4 cDNA was 100, 75, 50, 25, and 0% with the remainder being α-TM exons 1-3-4 cDNA. This analysis showed that neither the PCR for αTM2 nor the PCR for αTM3 detected the other cDNA. Quantitation of these data showed that the amplifications produced linear standard curves for each of the cDNAs. For the EGFP transcripts 5 μl of RT product was added to a mixture containing 12.5 μl of Universal Master Mix (Applied Biosystems), 0.125 μl of RNA forward primer, 0.125 μl of RNA reverse primer, 0.125 μl of VIC-labeled RNA Taqman probe, 2.5 μl of primer A (10 pmol/μl) or B (10 pmol/μl), 2.5 μl of primer C (10 pmol/μl), and 1 μl of probe D in a final volume of 25 μl. The RNA primers are pre-designed Taqman primers (Applied Biosystems). Primer A detected transcripts including the transgene TM exon 3 and had the sequence ACGGAGAAAAAG-CCACAG. Primer B detected transcripts excluding the transgene TM exon 3 and had the sequence GGATCGAGGGTCTGACCAAG. Primer G had the sequence GCCATCCAACTCTCTCTCACA. Samples were incubated at 50 °C for 2 min and heated to 95 °C for 10 min. PCR was performed using 40 cycles of 95 °C for 15 s and 60 °C for 1 min in an ABI 7000.

To validate this PCR, plasmids containing the EGFP gene spliced to include or skip the central exon were prepared, diluted to known concentrations (10 ng/μl to 160 fg/μl), and mixed. Q-RT-PCR of the mixtures showed that the PCR produced linear standard curves over this range. Because of the similarity in the sequences and the requirement to use identical upstream primers, both PCRs detect the other form of the transcript. However, in both cases the sensitivity of detection of the central exon form was ∼1000-fold greater than for the intended target. As a result, the presence of either form below 0.1% of the total product could not be determined.
corresponds to the cycle number at which the fluorescence reaches a threshold above the base-line value. Each reaction was performed in triplicate, and the Ct value for each test reaction was normalized to the rRNA value for the same sample (for determination using FAM-labeled probes) or to the rRNA value from parallel experiments (when SYBR green detection was used).

**Fluorescence Microscopy**—Tissues from mice were removed and embedded in Cryo-M-Bed. Ten-micron sections were cut, thaw-mounted onto superfrost plus slides, and mounted in glycerol. Sections were analyzed on a Nikon Eclipse E600 microscope. For confocal microscopy the sections were fixed in 4% paraformaldehyde for 10 min. The nuclei were stained with 4,6-diamidino-2-phenylindole. Images were captured using a Leica single photon confocal microscope with Leica TLS NT software.

**RESULTS**

**Regulation of Transgene Expression by Alternative Splicing**—The intronic region upstream of \(\alpha\)-TM exon 3, including the branch point sequence, P3 pyrimidine tract, and URE, together with the downstream intron including the DRE are sufficient to cause enhanced skipping of \(\alpha\)-TM exon 3 or a heterologous exon in PAC-1 cells (15). We decided to determine whether the characterized intronic elements are able to confer appropriate tissue-specific regulation upon splicing of a heterologous exon in transgenic mice. Our reporter constructs were all based upon EGFP under the control of the ubiquitously expressed CAGGS promoter (23). In a control construct the EGFP was present as a single exon (pCAGE), whereas in the other constructs the EGFP was present as a mini-gene consisting of three exons (Fig. 1C). The first exon contained the translational initiator ATG, whereas the third exon contained the remainder of the EGFP open reading frame. The central exon was flanked by the intron regulatory elements that usually surround \(\alpha\)-TM exon 3. Inclusion of this exon would disrupt the EGFP open reading frame, whereas regulated exon skipping...
would lead to an mRNA with an intact open reading frame for EGFP.

Most of our constructs used TM exon 3 as the central exon, but as an initial test we used αTM exon 2 (construct GTM2, Fig. 1C). This presents a particularly stringent test, because αTM exon 2 contains exon splicing enhancer elements stronger than those in exon 3 (24) and is usually selected in smooth muscle as a consequence of regulated repression of exon 3. Although the exon 2 exon splicing enhancers do not show cell-specific activity in cultured cells, it remains possible that they could mediate SM-specific activity in vivo, which would antagonize the negative regulation imposed by the intronic elements. In order to disrupt the EGFP open reading frame upon exon inclusion, a 4-bp insert in αTM exon 2 was created by infill of an XhoI site. This resulted in a frameshift such that the downstream GFP exon was out of frame.

The two initial constructs (pCAGE and pGTM2) were used to generate lines of transgenic mice by pronuclear injection. Consistent with previous reports pCAGE, transgenic mice expressed EGFP in all the tissues examined (23). These mice could be genotyped by examination under fluorescent light, and all tissues showed strong green fluorescence (data not shown). Conversely, pGTM2 transgenic mice showed no green fluorescence in the skin. Furthermore, no fluorescence was detected in the majority of tissues either as whole organs or in tissue sections. However, EGFP fluorescence was detected in the smooth muscle cells of the gut, bladder, uterus, and vasculature but not within other cell types in these tissues (Fig. 2A). This pattern of fluorescence was observed in three independent transgenic lines. These observations are consistent with restriction of EGFP protein expression to smooth muscle cells by regulated alternative splicing, despite the widespread transcriptional activity of the CAGGS promoter.

In order to verify that the GTM2 transgene was transcribed widely and that restricted EGFP expression was because of regulated splicing, RT-PCR was carried out using RNA harvested from various tissues (Fig. 2B). RNA from the transgene could be detected in all the tissue samples, although the levels of expression appeared to be low in the aorta and liver samples. In smooth muscle tissues (e.g. aorta, bladder, gut, and uterus) products corresponding to both inclusion and skipping of the transgene central exon were observed (Fig. 2B). In the kidney and heart no exon skipping was observed, whereas the lung sample showed predominant inclusion of the exon, with a small proportion of skipping. Surprisingly, in the liver both forms of the mRNA were detected even though EGFP was not detectable, presumably due to the low levels of expression. An additional strong band was observed in the kidney, heart, and lung samples about 100 bp above the exon included band. Sequencing of this product showed that it contained an additional 107-nt insert between the TM exon and the downstream GFP exon. This insert corresponds to a pseudo-exon, flanked by a consensus 5’ splice site and a non-consensus GAG 3’ splice site just downstream of the DRE. The pseudo-exon has been shown to be activated by a number of simple mutations in transfected constructs.2

To compare splicing of the endogenous gene with the transgene, the samples were analyzed by PCR for αTM using primers in exon 1 and exon 4. The amplicons generated from both αTM mRNA isoforms containing exons 2 or 3 are identical in size but can be distinguished by digestion with XhoI, which cleaves within exon 2 but not 3. This analysis showed that the exon 2-containing isoform was detectable in all the smooth muscle tissues and the liver but was not detected in the heart, kidney, or lung indicating that the transgene-splicing pattern mimicked the splicing pattern of the endogenous gene (Fig. 2C).

The only exception was that there was no evidence of the pseudo-exon containing band with the endogenous gene.

These initial data from the GTM2 transgenic mice clearly indicated that the regulatory elements flanking αTM exon 3 are sufficient to confer tissue-specific repression of splicing upon a heterologous exon and that this can be harnessed to allow regulated gene expression. We therefore decided to analyze in more detail the regulation of αTM splicing in vivo and the role of the regulatory elements previously defined in transient PAC-1 cell transfections.

Quantitative Analysis of Endogenous αTM Splicing—Our initial RT-PCR analysis had indicated an unexpected degree of exon 2 selection in liver (Fig. 2C), which is not enriched in smooth muscle cells, calling into question the cell type specificity of αTM splicing. Before further analyzing splicing of the EGFP transgenes, we therefore decided to quantitatively ana-
lyze splicing of exons 2 and 3 of αTM across a range of tissues. This had previously been surveyed qualitatively by nuclease protection and was restricted primarily to those tissues with high levels of expression (18, 19).

The levels of αTM mRNAs containing either exon 2 or exon 3 were determined by quantitative real time PCR and compared with the levels of rRNA in the same mouse tissue RNA samples. The proportion of αTM mRNA containing exon 2 (the "regulated" or SM pathway) was plotted against αTM mRNA expression level relative to rRNA (Fig. 3). Tissues with the highest levels of expression (heart, skeletal muscle, bladder, and aorta) showed the most extreme variations in splicing pattern. The striated muscle tissues, heart and skeletal muscle, which had the highest level of αTM expression, showed less than 0.2% exon 2 selection. In contrast, aorta and bladder, with the next highest levels of αTM expression, had more than 95% exon 2 selection. The next highest amount of αTM mRNA was present in the small intestine and lung, two tissues with a significant smooth muscle content. The amount of the exon 2-containing αTM in these samples was 80% for the small intestine and 29% in the lung. The brain had the lowest level of exon 2 inclusion (7%) apart from heart and skeletal muscle. Surprisingly, the other three tissues, none of which have a major smooth muscle component, showed much higher proportions of exon 2 inclusion: kidney 31%, spleen 55%, and liver 75%. Notably, these tissues had the lowest levels of αTM expression, with levels in liver more than 100-fold lower than in heart and skeletal muscles. These expression data indicate that the concept of a regulated splicing pattern in smooth muscle cells, which splice exclusively or predominantly exon 2, and a default inclusion of exon 3 in all other cell types needs to be re-examined (see "Discussion").

Variable Regulation of αTM Exon 3 Splicing in Different Mouse Lines—Having established the splicing patterns of endogenous αTM, we proceeded to generate a series of transgenic mouse lines based upon the pGTM3 construct (Fig. 1C), in which the central exon was αTM exon 3 flanked by its own regulatory introns. A single point mutation (G→A) at position 110 of exon 3 generated a stop codon in the EGFP reading frame. Three independent transgenic mouse lines were initially generated using the GTM3 transgene. The pattern of splicing in a range of tissues was determined by reverse transcription followed by Q-RT-PCR for exon-containing and skipped RNA isoforms. Examination of the products by agarose gel electrophoresis showed a single band for both reactions in most tissues (data not shown). The exon inclusion PCR also detected small amounts of product containing exon 3 spliced to the pseudo-exon in heart, skeletal muscle, and brain. Because this product contains exon 3, it was scored as exon inclusion for the purposes of quantitation. All three transgenic GTM3 lines (denoted r, s, and x) showed clear regulation of splicing (Fig. 4 and Table I). This was most evident in the negligible levels of exon skipping in heart, skeletal muscle, and brain (<1.2%) and very high levels of exon 3 skipping in bladder (>78%). Splicing patterns in the other tissues were much more variable. Aorta gave high levels of exon skipping (>46%), but with greater variability than bladder, whereas levels of exon skipping in the small intestine varied between 5 and 86%. The splicing patterns in liver, kidney, spleen, and lung were intermediate between the striated and smooth muscle tissues but with a high degree of variability. When the degree of transgene exon skipping was plotted against the percentage of endogenous α-TM exon 2 splicing (the equivalent splicing pattern to transgene exon skipping) in the same tissues, none of the mouse lines produced a straight line, which would be the expected pattern if transgene and endogenous α-TM splicing were regulated identically. Rather, all three mouse lines produced a curved plot, with lower regulation of transgene splicing compared with endogenous α-TM splicing in all tissues examined (data not shown). This suggests that whereas small intestine, liver, spleen, kidney, and lung are able to splice the endogenous αTM gene to give significant levels of exon 2 incorporation, they do not provide a "robust" regulatory environment that is able to support similar levels of regulated splicing of overexpressed transgenes (see below).

To analyze further the variability of transgene splicing, five
more independent pGTM3 transgenic mice were generated, and splicing patterns were determined in RNA isolated from the aorta, lung, small intestine, and spleen. Transgene exon skipping was highly variable in small intestine, kidney, and spleen but less variable in aorta (Fig. 5, A–D, and Table I). A weak trend of decreased transgene exon skipping associated with increased transgene expression was observed in some, but not all, tissues examined. For each tissue the mouse line with the highest expression levels gave one of the lowest levels of exon skipping (line r in aorta and small intestine, z in lung, and y in spleen, see Fig. 5).

The variability in tissue specificity of GTM3 transgene splicing between individual mouse lines posed a potential problem for our subsequent goal of analyzing the effects of mutations in the splicing regulatory elements. However, the extreme splicing patterns of exon skipping in bladder and inclusion in skeletal muscle were robustly maintained in all the GTM3 mouse lines, irrespective of expression levels (Table I and Fig. 4). We therefore adopted a regulation index based upon the arithmetic difference in exon skipping between bladder and skeletal muscle (Table III). The alternative “fold difference” approach is potentially more meaningful in terms of the underlying mechanism and tissue specificity of regulation, but it is vulnerable to large fluctuations arising from variations in the very small values of transgenic exon skipping determined in skeletal muscle. Conversely, these variations have negligible impact upon the arithmetic index; the arithmetic regulation index for GTM3 mice was 79.2 ± 2.2, whereas the fold regulation index was 239 ± 90 (Table III). Use of the arithmetic regulation index subsequently allowed a meaningful analysis of the effects of mutations in splicing regulatory elements despite the line-to-line variability of transgenic splicing in other tissues.

Effect of Mutations to the Known Splicing Regulatory Sequences—To analyze the role of the regulatory elements in the introns flanking exon 3, transgenic mice were generated containing a series of derivatives of pGTM3 with mutations in the defined regulatory elements (Fig. 1C). The P3Δ123 mutant contains four pyrimidine transitions in the three optimal PGB-binding sites within the polypyrimidine tract of exon 3, ∆URE contains a 15-nucleotide deletion of the URE, and ∆DRE has a 90-nucleotide deletion of the DRE (Fig. 1B). These three mutations all cause a decrease of αTM exon 3 skipping in PAC-1 cells (12, 14). In contrast, the 3URE mutant contains two additional copies of the URE and in various cell lines increases the levels of exon 3 skipping (13).

Consistent with the previously reported effects in transfected cells, all four mutations had obvious effects in vivo. The P3Δ123, ∆URE, and ∆DRE mutations caused a progressively severe reduction of transgene exon skipping in all cell types, reflected in a marked reduction in the regulation index (Table III). Deletion of the DRE (three lines) had a greater effect than deletion of the URE (three lines) in all the transgenic mice studied. In all tissues deletion of the DRE reduced exon skipping to less than 1% (Table II). This strong effect is consistent with the fact that the DRE consists of two separable regulatory elements, the UGC region and the DY pyrimidine tract. The ∆URE and P3Δ123 mutations had a quantitatively similar effect as ∆DRE in a number of non-smooth muscle tissues, reducing exon skipping to below 1%. In tissues with a significant smooth muscle content (bladder, aorta, and small intestine), there was significant residual transgene exon skipping with ∆URE (as high as 9% in bladder and 22% in aorta of line “j”), whereas P3Δ123 mutants retained up to 28% exon skipping in the bladder and 32% in the aorta. The increasingly severe impact of the mutations (P3Δ123 < ∆URE < ∆DRE) was particularly evident in the arithmetic regulation index (Table III). When regulation was assessed by the fold difference in exon skipping, the P3Δ123 and ∆URE mutations had a less obvious effect, whereas the more severe ∆DRE mutation impaired regulation using either index. These data clearly demonstrate that the regulatory elements identified by transfection of cultured cells play a crucial role in regulation of αTM splicing in vivo.

Multimerization of the URE—Multimerization of the URE enhances TM exon 3 skipping in PAC-1 and various other cell lines (13), with a concomitant decrease in regulation because of the elevated skipping in the non-SM cells. We therefore generated transgenic mice expressing a spliced transgene containing three tandem copies of the URE (3URE). Widespread EGFP fluorescence was observed in all lines of the 3URE mice (data not shown), suggesting reduced restriction of transgene expression via regulated splicing. Analysis of transgene RNA showed

### Table I

| Tissue          | Exon skipping | S.D  | Range |
|-----------------|---------------|------|-------|
| Aorta           | 69.5          | 14   | 48–87 |
| Bladder         | 79.5          | 2    | 78–82 |
| Small intestine| 48            | 31   | 5–86  |
| Lung            | 28            | 22   | 6–62  |
| Heart           | 1             | 0.1  | 0.9–1.2 |
| Skeletal muscle | 0.4           | 0.2  | 0.6–0.2 |
| Brain           | 0.4           | 0.2  | 0.6–0.2 |
| Liver           | 9             | 15   | 1–26  |
| Kidney          | 8             | 11   | 1–22  |
| Spleen          | 13            | 14   | 2–35  |

Fig. 5. Relationship of GTM3 transgene splicing and expression levels. The percentage of GTM3 transgene exon skipping is plotted against transgene expression level relative to rRNA for lung (A), spleen (B), small intestine (C), and aorta (D) from eight lines of GTM3 mice. Although there was no clear relationship, in each tissue the highest amount of exon skipping was in relatively low expressing lines, whereas the highest expresser gave low levels of skipping. As a reference, the horizontal dashed line indicates the splicing pattern of endogenous αTM (for which % exon 3 skipping is equivalent to exon 2 inclusion).


Table II

| Tissue          | GTM3 | P3A123 | ΔURE | ΔDRE | 3URE |
|-----------------|------|--------|------|------|------|
| Aorta           | 48–87| 12–32  | 2–22 | 0.3–0.4| 8–86 |
| Bladder         | 78–82| 14–28  | 3–9  | 0.3–0.4| 61–74|
| Small intestine | 5–86 | 2–17   | 1–3  | 0.1–0.2| 15–82|
| Lung            | 6–62 | 0.3–0.6| 0.1–0.4| 0.1–0.2| 5–46 |
| Heart           | 0.9–1.2| 0.2–0.3| 0.01–0.1| 0.05–0.1| 3–20 |
| Skeletal muscle | 0.2–0.6| 0.07–0.3| 0.04–0.1| 0.01–0.15| 0.2–2 |
| Brain           | 0.3–0.7| 0.08–0.11| 0.06–0.1| 0.04–0.06| 2–2.6|
| Liver           | 1–26 | 0.1–0.2| 0.08–0.18| 0.02–0.6| 10–46|
| Kidney          | 1–22 | 0.2–0.6| 0.04–0.24| 0.04–0.2| 21–39|
| Spleen          | 2–35 | 0.2–1.1| 0.1–0.4| 0.1–0.13| 7–52 |

Table III

| Transgene       | Bladder – skeletal muscle | Bladder/skeletal muscle |
|-----------------|---------------------------|-------------------------|
| Endogenous aTM  | 99.1                      |                         |
| GTM3            | 72.2 ± 2.2                | 239 ± 90                |
| P3A123          | 20.7 ± 7.1                | 219 ± 60                |
| ΔURE            | 6.5 ± 2.9                 | 97 ± 16                 |
| ΔDRE            | 0.25 ± 0.0.3              | 4.7 ± 2.3               |
| 3URE            | 74.3 ± 13.5               | 157 ± 140               |

* Arithmetic regulation index = % exon 3 skipping in bladder minus exon 3 skipping in skeletal muscle.
* Fold regulation index = % exon 3 skipping in bladder divided by exon 3 skipping in skeletal muscle.

By contrast, in the aorta there was a relationship between the amount of regulated splicing of the endogenous aTM exon 2 and exon skipping in the 3URE transgene (Fig. 7C). The mouse lines with the highest level of 3URE expression also had the lowest levels of regulated splicing of both the transgene and the endogenous aTM (Fig. 7D). Moreover, when aTM exon 2 selection was plotted against transgene expression levels, a clear decrease in exon 2 selection was observed to accompany increased expression of the 3URE transgene in aorta (Fig. 7D). In contrast, none of the GTM3 lines showed any variation in aTM exon 2 splicing (Fig. 7, C and D). These observations suggest that over the expression range analyzed in the aorta, the 3URE, but not the GTM3, transcript is able to titrate factors that are involved in regulation of the endogenous aTM splicing.

**DISCUSSION**

Our data from transgenic mice have validated earlier findings by using the partially differentiated PAC-1 cells and have
also provided important new insights not available by using cultured cells. The first important observation was that the splicing of endogenous αTM does not follow a simple smooth muscle versus default model. This was evident both from analysis of the endogenous αTM gene (Fig. 3) and from the GTM3 lines that most closely recapitulated native regulation (Fig. 4, line α). Whereas the smooth muscle rich tissues such as bladder and aorta showed nearly complete exon 2 inclusion and the striated muscles showed full inclusion of exon 3, a number of other tissues with little smooth muscle content, such as liver and spleen, showed significant percentages of exon 2 inclusion. Notably, these tissues had very low levels of αTM expression, which had not been detectable using nuclease protection or Northern blot (18). Subsequent analyses of GTM3 transgene splicing indicated that bladder and aorta provided a robust regulatory environment that was able to support significant levels of transgene exon skipping in all mouse lines, even with high levels of transgene expression. The degree of regulation was not as great as with the endogenous gene (79% compared with 99%, Table III), possibly resulting from the lack of a competing mutually exclusive exon in the transgene (Fig. 1), or from the truncations in the flanking introns. In contrast to bladder and aorta, the other tissues were more sensitive to the levels of transgene expression and could only support exon skipping at lower levels of expression. Expression of exon 2 containing αTM isoforms has been detected previously in murine ES cells at stages of embryonic development preceding the formation of SM tissues (25). However, that analysis did not look at the contemporaneous expression of exon 3-containing isoforms, so it is not clear whether the detected exon 2-containing αTM was a major or minor isoform. Similar challenges to simple default versus regulated models for splicing mechanisms have also been reported in other systems (26, 27). Although the concept of smooth muscle and default splicing patterns does not accurately describe the ratio of exon 2- and 3-containing αTM isoforms across various tissues, it neverthe-

less appears to reflect a property unique to smooth muscle of having the capacity to confer regulated splicing even to high levels of both endogenous and ectopically expressed RNA. Likewise, the striated muscle tissues (skeletal and cardiac muscle) are able to robustly splice high levels of endogenous or transgene RNA to include exclusively αTM exon 3. Exon 3 inclusion does not require specific regulatory elements (Tables II and III) (11–13), so it appears most likely that striated muscles contain very low levels of the regulators responsible for the SM-splicing pathway. This could result from very low absolute levels. Alternatively, the low activity of regulatory factors in striated muscles could result from their titration by the very high levels of αTM expression in these tissues.

Various lines of evidence have indicated that use of heterologous promoters can sometimes alter the regulation of downstream alternative splicing events (reviewed in Ref. 28). This could present a serious impediment to the approach that we adopted, in which a widely active promoter drives expression of an alternatively spliced transgene. Therefore, it was gratifying that, despite the considerable line-to-line variability, in some mouse lines splicing of the GTM3 transgene was regulated in parallel with endogenous αTM splicing (Fig. 4). The variability in αTM splicing behavior between individual mouse lines probably arises from more than one cause. First, variations in absolute levels of expression in individual tissues such as spleen appeared to give rise to titration effects, such that higher levels of transgene expression were accompanied by lower proportions of exon skipping (Figs. 5 and 7). Further variability may have arisen in tissues such as small intestine that have a mixture of SM and other cell types. Although the pCAGGS promoter is active in most cell types, it nevertheless shows variable activity between tissues. Furthermore, we observed that the hierarchy of activity in different tissues varied between individual mouse lines. Therefore, within tissues such as small intestine higher levels of transgene expression in the non-SM cells would lead to an apparent unregulated splicing pattern, even though the fluorescence microscopy showed that EGFP expression was restricted to the SM cells of the gut (Fig. 2). The combination of these two factors may explain the observation that small intestine showed the greatest variation in splicing pattern of the GTM3 transgene (Table II, 5–86% exon skipping). Such line-to-line variability that we observed could be circumvented by targeted insertion of the transgene by homologous recombination into the Hprt locus (29). However, integration at the Hprt locus facilitates high level constitutive transcription, so it is possible that mouse lines would not express the transgene at sufficiently low levels to recapitulate appropriate tissue-specific regulation.

The effects of transgene expression levels upon regulated splicing of both transgene and endogenous αTM indicate that regulatory factors distinct from the general splicing machinery can be titrated by the transgene RNA. The strongest titration effects were observed with the 3URE transgene, which at high levels led to lower exon 2 selection for endogenous αTM in the aorta (Fig. 7). At lower expression levels in non-SM tissues, the 3URE construct generally led to higher levels of exon skipping compared with GTM3, consistent with the concept that URE multimers can help to recruit limiting amounts of regulatory factors (13). In contrast, in bladder and aorta, where these factors are presumably not limiting, 3URE behaved similarly to GTM3, at least at low levels of expression. The URE consists of a short region of UGC or CUG repeats, and thus far the regulatory factors that bind to it have not been identified. Obvious candidates include members of the CELF family of RNA-binding proteins, which bind to CUG or UG repeats (30, 31). However, overexpression of known family members (CUG-BP, ETR3, and CELF4) antago-

![Fig. 7. Alteration of endogenous αTM splicing in aorta of 3URE but not GTM3 transgenic mice.](http://www.jbc.org/Downloaded from http://www.jbc.org/)
ized regulation of αTM splicing, suggesting that the tested proteins were not authentic regulators of this system (13). In the aorta, expression of 3URE appeared to titrate the factors responsible for regulating splicing of both the transgene and endogenous αTM. In contrast, the situation in the spleen was more complex. Although the 3URE transgene showed a clear decrease in exon skipping as its level of expression increased, there was no apparent effect upon splicing of the endogenous αTM. This suggests that in this case there are distinct pools of regulatory factors, one of which was used by the 3URE transgene and another by the endogenous αTM.

Knock-out of αTM in transgenic mice has shown that homozygous null mutants are embryonic lethal but that loss of a single allele produces no phenotype, probably due to compensating mechanisms that maintain the quantity of the αTM protein (32, 33). There have been no reports of isoform-specific knockouts, although this was the intention of at least one previous investigation (33). An interesting consequence of the titratability of some of the splicing regulators is that in the highest expressing 3URE mouse line the level of exon 2 inclusion in the aorta was reduced from the usual −90% to less than 10%. This suggests that depletion of the exon 2 isoform does not cause a lethal phenotype, although it is possible that more detailed investigation of such animals might reveal a specific phenotype in the vasculature or other smooth muscle tissues. It is also possible that the reduction in splicing of the TM2 isoform resulted in up-regulation of αTM expression to compensate for the splicing dependent loss of αTM. Consistent with this suggestion, the mice with the highest levels of 3URE expression and only 4% TM exon 2 inclusion had higher overall levels of TM transcripts in the aorta, so that the reduction in absolute levels of the TM exon 2 isoform was only ∼4-fold.

In addition to exon inclusion and skipping, a third splicing pathway was observed with both the GTM2 and GTM3 constructs. This involved both normal exon inclusion as well as an additional 107 nt from the downstream intron. We have observed splicing of this 107-nt “pseudo-exon” in transfected mutant constructs in HeLa and PAC-1 cells, resulting from simple mutations that activate its splice sites. Inclusion of the pseudo-exon results in the introduction of in-frame premature termination codons in RNA from both construct and endogenous αTM. Because these premature termination codons are more than 55-nt upstream of the splice junction between the pseudo-exon and the GFP exon (or αTM exon 4), they are predicted to provoke NMD (34, 35). This means that in our quantification of transgene splicing, we will have underestimated the level of exon inclusion, because products containing the pseudo-exon are likely to be at levels that under-represent the actual frequency of the splicing event. Nevertheless, because pseudo-exon inclusion mainly occurred in tissues such as heart and skeletal muscle, in which exon inclusion is predominant, NMD of pseudo-exon-containing products is unlikely to cause a gross distortion of the total quantitated data. Despite the clear inclusion of the pseudo-exon in the construct RNA, its inclusion in endogenous αTM RNA is much more difficult to detect. There are a number of possible explanations for the apparent contrast between the behavior of the pseudo-exon in the transgene and endogenous αTM. First, the smaller size of the intron downstream from the pseudo-exon in the transgene (174 bp) compared with the endogenous αTM gene (−12 kb) may favor pseudo-exon inclusion. Second, in the transgene there is a single intron downstream of the pseudo-exon stop codon, whereas in the endogenous gene there are 7–8. Although a single intron provoked NMD as efficiently as 2–4 downstream introns in the TPI gene (36), the additional exon junction complexes in αTM might lead to much more efficient NMD of the endogenous product compared with the transgene.

Whereas investigations of alternative splicing in genetically tractable organisms such as Drosophila melanogaster and Caenorhabditis elegans have routinely used transgenic animals, studies of mammalian alternative splicing have typically focused on cell culture in order to analyze regulatory elements. However, transgenic mice have been used in some analyses of alternative splicing mechanisms. The earliest case involved widespread expression of a calcitonin/CGRP transgene, showing that the majority of tissues were able to produce calcitonin as a default product, even though it is usually only produced in the thyroid. In contrast, transgene processing to CGRP was limited to neurons (37). More recently, an artificial transgene with an arrangement of competing 3′ end processing and splice sites, analogous to the immunoglobulin μ (Igμ) heavy chain gene, was found to be co-regulated with the Igμ gene in transgenic mice, arguing against the need for sequence-specific regulators of Igμ processing (27). Moreover, because the transgene was more widely expressed than the Igμ gene, it could be seen that neither of the splicing patterns in resting or lipopolysaccharide-stimulated B cells represented a default or regulated state. Analysis of transgene tissue-specific splicing has been carried out for exon 9 of the F1γ gene (38). However, only two mutant minigenes were tested in transgenic animals so that a direct comparison with a wild type construct could not be made. Alternative splicing of Smn exon 7 has also been analyzed in transgenic mice, although no tissue specificity of splicing was observed (39). Our investigation is one of the most detailed quantitative analyses to date of a mammalian model alternative splicing system in vivo. It has served to validate findings from the partially differentiated PAC1 cell line, highlighted some potential complications with transgenic models, and also given a number of important insights that would not have been available using previous methods.

Acknowledgments—We thank Nick McGlincy, Mat Wollerton, and Jim Patton for critical comments on the manuscript.

REFERENCES

1. Black, D. L. (2000) Cell 103, 367–370
2. Manczko, T., and Tazelaar, E. (2002) Nature 418, 236–243
3. Roberts, G. C., and Smith, C. W. (2002) Curr. Opin. Chem. Biol. 6, 375–383
4. Modrek, B., and Lee, C. (2002) Nat. Genet. 30, 15–19
5. Jorcano, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Lorch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. (2003) Science 302, 2141–2144
6. Lewis, B. P., Green, R. E., and Brenner, S. E. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 189–192
7. Black, D. L. (2003) Annu. Rev. Biochem. 72, 291–336
8. Owens, G. K. (1995) Physiol. Rev. 75, 457–471
9. Shanahan, C. M., Weissberg, P. L., and Metcalfe, J. C. (1993) Circ. Res. 73, 204–204
10. Mullen, M. P., Smith, C. W., Patton, J. G., and Nadal-Ginard, B. (1991) Genes Dev. 5, 642–655
11. Gooding, C., Roberts, G. C., Moreau, G., Nadal Ginard, B., and Smith, C. W. J. (1994) EMBO J. 13, 3861–3872
12. Gooding, C. G., Roberts, G. C., and Smith, C. W. J. (1998) RNA (N. Y.) 4, 10–100
13. Gromek, N., and Smith, C. W. (2002) Nucleic Acids Res. 30, 3548–3557
14. Gromek, N., Rideau, A., Souby, J., Bradden, A. D., Gooding, C., Hutelman, S., Singer, R. H., and Smith, C. W. (2003) EMBO J. 22, 6356–6364
15. Roberts, G. C., Gooding, C., and Smith, C. W. (1996) EMBO J. 15, 6301–6310
16. Wollerton, M. C., Gooding, C., Robinson, P., Brown, E. C., Jackson, R. J., and Smith, C. W. (2001) RNA (N. Y.) 7, 819–832
17. Smith, C. W. J., and Nadal Ginard, B. (1989) Cell 56, 749–758
18. Wieczorek, D. F., Smith, C. W. J., and Nadal Ginard, B. (1988) Mol. Cell. Biol. 8, 679–694
19. Lees Miller, J. P., Goodwin, L. O., and Helfman, D. M. (1990) Mol. Cell. Biol. 10, 1729–1742
20. Gooding, C., Kemp, P., and Smith, C. W. (2003) J. Biol. Chem. 278, 15201–15207
21. Rothman, A., Kulik, T. J., Taubman, M. B., Berk, B. C., Smith, C. W. J., and Nadal Ginard, B. (1992) Circulation 86, 1977–1986
22. Perez, I., Lin, C.-H., McAfee, J. G., and Patton, J. G. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 7785–7790
23. Okabe, M., Ikawa, M., Kominami, K., Nakaniishi, T., and Nishimune, Y. (1997) FEBS Lett. 407, 313–319
24. Dye, B. T., Buvoli, M., Mayer, S. A., Lin, C. H., and Patton, J. G. (1998) RNA (N. Y.) 4, 1523–1536
25. Muthuchamy, M., Pajak, L., Howles, P., Doetschman, T., and Wieczorek, D. F. (1993) Mol. Cell. Biol. 13, 3311–3323
26. Charlet, B. N., Logan, P., Singh, G., and Cooper, T. A. (2002) Mol. Cell 9, 649–658
27. Seipelt, R. L., Spear, B. T., Snow, E. C., and Peterson, M. L. (1998) Mol. Cell. Biol. 18, 1042–1048
28. Caceres, J. F., and Kornblihtt, A. R. (2002) Trends Genet. 18, 186–189
29. Bronson, S. K., Plaehn, E. G., Kluckman, K. D., Hagaman, J. R., Maeda, N., and Smithies, O. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 9067–9072
30. Good, P. J., Chen, Q., Warner, S. J., and Herring, D. C. (2000) J. Biol. Chem. 275, 28583–28592
31. Ladd, A. N., Charlet-B, N., and Cooper, T. A. (2001) Mol. Cell. Biol. 21, 1285–1296
32. Blanchard, E. M., Iizuka, K., Christe, M., Conner, D. A., Geisterfer-Lowrance, A., Schoen, F. J., Maughan, D. W., Seidman, C. E., and Seidman, J. G. (1997) Circ. Res. 81, 1005–1010
33. Rethinasamy, P., Muthuchamy, M., Hewett, T., Boivin, G., Wolska, B. M., Evans, C., Solaro, R. J., and Wieczorek, D. F. (1998) Circ. Res. 82, 116–123
34. Maquat, L. E., and Carmichael, G. G. (2001) Cell 104, 173–176
35. Wagner, E., and Lykke-Andersen, J. (2002) J. Cell Sci. 115, 3033–3038
36. Zhang, J., Sun, X., Qian, Y., LaDuca, J. P., and Maquat, L. E. (1998) Mol. Cell. Biol. 18, 5272–5283
37. Crenshaw, E. B., III, Russo, A. F., Swanson, L. W., and Rosenfeld, M. G. (1987) Cell 49, 389–398
38. Ichida, M., Hakamata, Y., Hayakawa, M., Ueno, E., Ikeda, U., Shimada, K., Hamamoto, T., Kagawa, Y., and Endo, H. (2000) J. Biol. Chem. 275, 15992–16001
39. DiDonato, C. J., Lorson, C. L., De Repentigny, Y., Simard, L., Chartrand, C., Androphy, E. J., and Kothary, R. (2001) Hum. Mol. Genet. 10, 2727–2736
Regulated Tissue-specific Alternative Splicing of Enhanced Green Fluorescent Protein Transgenes Conferred by α-Tropomyosin Regulatory Elements in Transgenic Mice

Peter D. Ellis, Christopher W. J. Smith and Paul Kemp

J. Biol. Chem. 2004, 279:36660-36669.
doi: 10.1074/jbc.M405380200 originally published online June 11, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M405380200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 19 of which can be accessed free at
http://www.jbc.org/content/279/35/36660.full.html#ref-list-1