doi:10.1093/jmcb/mjr030
Published online November 1, 2011

Article

Exon-centric regulation of pyruvate kinase M alternative splicing via mutually exclusive exons

Zhenxun Wang[1,2], Deblina Chatterjee[1,3], Hyun Yong Jeon[1,3], Martin Akerman[1], Matthew G. Vander Heiden[4], Lewis C. Cantley[5], and Adrian R. Krainer[1,*]

[1] Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA
[2] Watson School of Biological Sciences, Cold Spring Harbor, NY 11724, USA
[3] Graduate Program in Molecular and Cellular Biology, Stony Brook University, Stony Brook, NY 11794, USA
[4] Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA
[5] Division of Signal Transduction, Beth Israel Deaconess Medical Center and Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
* Correspondence to: Adrian R. Krainer, E-mail: krainer@cshl.edu

Alternative splicing of the pyruvate kinase M gene (*PK-M*) can generate the M2 isoform and promote aerobic glycolysis and tumor growth. However, the cancer-specific alternative splicing regulation of *PK-M* is not completely understood. Here, we demonstrate that *PK-M* is regulated by reciprocal effects on the mutually exclusive exons 9 and 10, such that exon 9 is repressed and exon 10 is activated in cancer cells. Strikingly, exonic, rather than intronic, *cis*-elements are key determinants of *PK-M* splicing isoform ratios. Using a systematic sub-exonic duplication approach, we identify a potent exonic splicing enhancer in exon 10, which differs from its homologous counterpart in exon 9 by only two nucleotides. We identify SRSF3 as one of the cognate factors, and show that this serine/arginine-rich protein activates exon 10 and mediates changes in glucose metabolism. These findings provide mechanistic insights into the complex regulation of alternative splicing of a key regulator of the Warburg effect, and also have implications for other genes with a similar pattern of alternative splicing.

Keywords: alternative splicing, cancer metabolism, pyruvate kinase, SRSF3

Introduction

Cancer cells exhibit a metabolic phenotype termed aerobic glycolysis, or the Warburg effect, characterized by increased glycolysis with lactate generation, regardless of oxygen availability (Vander Heiden et al., 2009). Expression of the type II isoform of the pyruvate kinase M gene (*PKM2*, referred to here as *PK-M*) mediates this metabolic phenotype, and confers a proliferative advantage to tumor cells *in vivo* (Christofk et al., 2008a).

Pyruvate kinase (PK) catalyzes the final step in glycolysis, generating pyruvate and ATP from phosphoenolpyruvate and ADP (Dombrauckas et al., 2005). The exons 9 and 10 of the *PK-M* gene can each encode a 56-amino-acid segment, and be alternatively spliced in a mutually exclusive (ME) fashion to give rise to M1 and M2 isoforms, respectively (Noguchi et al., 1986). PK-M1 is constitutively active and predominantly expressed in terminally differentiated tissues (Christofk et al., 2008a; Clower et al., 2010). PK-M2 is expressed in cancer cells, as well as in fetal and undifferentiated adult tissues, and is allosterically regulated by fructose-1,6-bisphosphate (FBP) and can interact with tyrosine-phosphorylated signaling proteins (Christofk et al., 2008a, b). The growth signal-mediated inhibition of PK-M2 activity contributes to cancer cell growth by decreasing carbon flux through the catabolic glycolytic pathway, allowing accumulated upstream intermediates to be shunted to anabolic pathways to facilitate cell proliferation (Hitosugi et al., 2009).

ME exons are responsible for $<2\%$ of alternative splicing (Chacko and Ranganathan, 2009). Several mechanisms involved in ME exon selection have been described, but how they are coordinately regulated in a ME fashion is not well understood (Smith, 2005). ME exons are often homologous, indicating an exon-duplication origin (Letunic et al., 2002). The exons 9 and 10 of *PK-M* are ME exons whose ME splicing mechanism might be novel, as the length (401 bp) and sequence of intron 9 rule out steric interference that could prevent double splicing due to the spacing of the branch site and the 5′ splice site (5′ss) (Smith and Nadal-Ginard, 1989).

Recently, we and others demonstrated that exon 10 is preferred in cancer and proliferating cells, and also implicated two pairs of splicing-repressor paralogs—PTB/nPTB and hnRNPA1/A2—in exon 9 repression (Clower et al., 2010; David et al., 2010). It remains unclear whether additional repressors block exon 9, and exon 10 is the 'default' exon in proliferating cells, i.e. is its selection actively promoted, or is it included simply as a consequence of exon 9 repression? Moreover, though hnRNPA1/A2 and PTB appear to bind in the intronic regions flanking exon 9 (David et al., 2010), it remains unclear where the critical *cis*-elements responsible for the

Received July 11, 2011. Revised August 9, 2011. Accepted August 21, 2011.

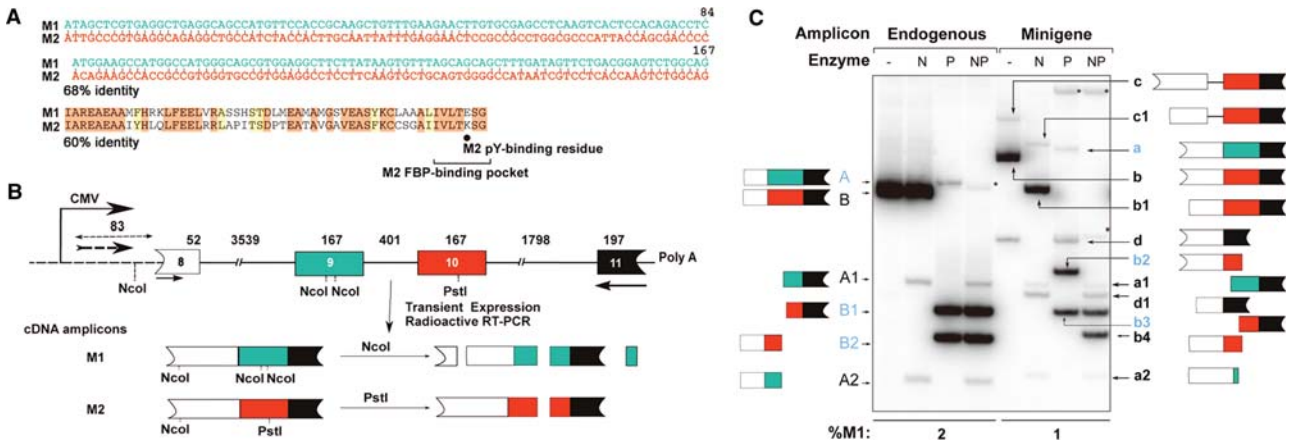

Figure 1 Detection of endogenous and minigene-specific *PK-M* spliced isoforms. (**A**) Nucleotide (top) and amino acid (bottom) sequence alignments of ME exons 9 (M1) and 10 (M2). Identical nucleotides are shown by vertical dashes. Identical and similar amino acids are highlighted in red and yellow, respectively. The distinctive phosphotyrosine-binding residue and the FBP-binding pocket of PK-M2 are indicated. The percentages of nucleotide and amino acid identity are shown. (**B**) Diagram of the human *PK-M* minigene. The minigene comprises the intact introns 8, 9, and 10, the intact alternative exons 9 and 10, and portions of the flanking constitutive exons 8 and 11. The numbers above each exon and intron show the length in nucleotides. A vector-specific forward primer (dashed arrow) and a reverse primer annealing to exon 11 were used to amplify minigene-derived transcripts; to amplify endogenous transcripts in untransfected cells, a forward primer annealing to exon 8 (solid arrow) was used instead. To distinguish between exon 9-included (M1 isoform) and exon 10-included (M2 isoform) transcripts, cDNA amplicons were cleaved with *Nco*I (N), *Pst*I (P) or both (NP). Note the additional *Nco*I site at the 5′ end of the *PK-M* minigene, which is absent from the endogenous gene. (**C**) Radioactive RT–PCR and restriction digest of endogenous and minigene-derived *PK-M* transcripts in HEK-293 cells. RNA was isolated from untransfected cells or 48 h after transfection of the plasmid minigene. Asterisks (*) indicate restriction fragments corresponding to unspliced *PK-M* pre-mRNA. cDNAs and fragments from endogenous mRNAs are indicated on the left in uppercase font; those derived from minigene-specific transcripts are shown on the right in lowercase bold font. The most important bands are indicated in blue font. The bands correspond to: uncut M1 fragment (A, 398 nt; a, 481 nt); uncut M2 fragment (B, 398 nt; b, 481 nt); *Nco*I-cleaved M1 3′ fragment (A1, 248 nt; a1, 248 nt); *Nco*I-cleaved M1 5′ fragment (A2, 144 nt; a2, 150 nt); *Pst*I-cleaved M2 5′ fragment (B2, 185 nt; b2, 268 nt); *Pst*I-cleaved M2 3′ fragment (B1, 213 nt; b3, 213 nt). Minigene M2 cDNA is cleaved by *Nco*I, giving rise to b1 (404 nt). Two additional species are observed from minigene-specific transcripts: a spliced mRNA that skips both exons 9 and 10 (d, 314 nt) and is cleaved by *Nco*I (d1, 237 nt), and an M2 transcript that uses a cryptic 3′ splice site 115 nt upstream of the authentic 3′ss of exon 10 (c, 596 nt), and whose 3′ fragment after *Pst*I digestion is identical to b3. Due to its low abundance, the corresponding 5′ end after *Pst*I digestion can only be detected with certain mutant minigenes (see Figure **2A**). The 5′ end of the *Pst*I-cleaved M2 fragment, b2, is additionally cleaved by *Nco*I to a shorter fragment (b4, 194 nt) plus a 74-nt fragment that runs off the gel. See Supplementary Figure S1 for a detailed description of these M2-variant species. The numbers below the gel indicate the percentage of exon 9-included transcripts (%M1); standard deviations (SD) are $\leq$0.2% ($n \geq 3$).

PK-M2 splicing pattern in proliferating cells are distributed, i.e. are they present in the exons, the introns, or both?

To address these questions, we constructed a minigene that recapitulates the splicing-regulatory features of endogenous *PK-M*. We demonstrate that exon 10 is activated in cancer cells independently of exon 9 repression. We further show that exonic, but not intronic, *cis*-elements, are the key determinants of *PK-M* alternative splicing. Using a sub-exonic duplication strategy, we further mapped an exonic splicing enhancer (ESE) in exon 10, and found that SRSF3 (formerly SRp20), an oncogenic member of the serine/arginine-rich (SR) protein family of splicing activators, is its cognate binding factor. SRSF3 knockdown in cancer cells rescues PK-M1 expression and decreases lactate production and cellular proliferation.

Results

PK-M minigene recapitulates alternative splicing of the endogenous gene

The ME exons 9 and 10 of *PK-M* are identical in length, and highly homologous at the nucleotide and amino acid sequence levels (Figure 1A). To analyze the mechanism of ME splicing of *PK-M* pre-mRNA, and identify splicing *cis*-elements, we generated a minigene transcribed from a CMV promoter (Figure 1B). The minigene consists of the genomic region encompassing exons 9 and 10, introns 8, 9, and 10, and proximal portions of the flanking exons 8 and 11. Detection and analysis of minigene-derived transcripts utilized an RT–PCR and restriction-digestion strategy, as described for endogenous *PK-M* transcripts (Figure 1B) (Clower et al., 2010). To characterize all possible minigene-derived species after cell transfection, we selectively amplified them from total cDNA using a forward primer specific for upstream vector sequence, and a reverse primer annealing to constitutive exon 11 (Figure 1B).

Figure 1C shows the profiles of endogenous and minigene-derived transcripts in HEK-293 cells (similar data not shown for HeLa, A172, and SKNBE cells). The minigene predominantly expressed the exon 10-included *PK-M2* isoform, paralleling endogenous *PK-M2*/*M1* splicing ratios. Two additional, minor minigene RNAs not found in endogenous *PK-M* transcripts were also reproducibly detected: an RNA lacking both exons 9 and 10,

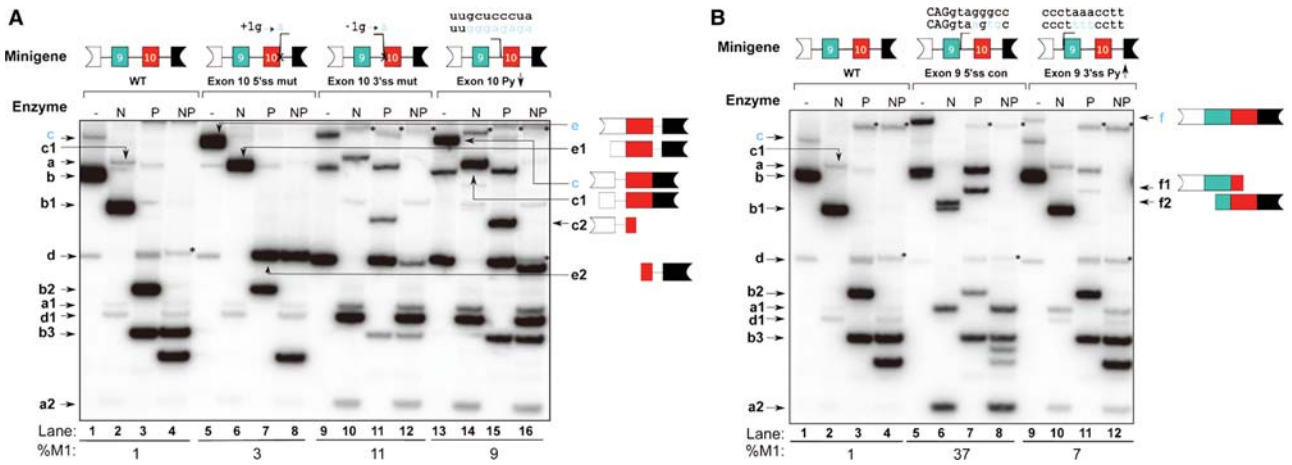

Figure 2 Effects of splice-site relative strengths on inclusion of exon 9. Mutant minigenes were analyzed by transient transfection into HEK-293 cells, followed by radioactive RT–PCR and restriction digests, as in Figure **1C**. (**A**) Mutations that inactivate or weaken the splice sites (ss) of exon 10. The mutated 5′ss, 3′ss, or PPT nucleotides are indicated in light blue. Numbers indicate the position of the mutated nucleotide, either upstream (+), or downstream (−) of the exon. Bands marked with asterisks (*) are fragments of *PK-M* pre-mRNA, as in Figure **1C**. Bands already described in Figure **1C** are indicated on the left. New and/or important bands are indicated on the right, and the key bands are labeled in blue font (see also Supplementary Figure S1). The 5′ss mutation gives rise to a new M2 variant (e, 586 nt) derived from use of a cryptic 5′ss 105 nt downstream of the authentic 5′ss; this band is cleaved by *Nco*I (e1, 509 nt), and *Pst*I digestion generates a longer 3′ fragment (e2, 318 nt). The increase in the M2 3′ss cryptic variant (c, as described in Figure **1C**) upon mutation of the PPT or 3′ss allows detection of the 5′ *Pst*I fragment (c3, 383 nt). The %M1 inclusion is indicated at the bottom; SD: 0.2% (WT); 0.3% (exon 10 5′ss mut); 2% (exon 10 3′ss mut); and 2% (exon 10 Py↓) ($n \geq 3$). (**B**) Mutations that strengthen the splice sites of exon 9. The mutated 5′ss and PPT nucleotides are indicated in light blue. Uppercase and lowercase letters indicate exonic and intronic sequences, respectively. Bands are labeled as in **A**. A new exon 9–exon 10 doubly included mRNA is indicated on the right (f, 648 nt). This band is sensitive to both *Pst*I (f1, 435 nt) and *Nco*I (f2, 415 nt) digestion. See also Supplementary Figure S1. The %M1 inclusion is indicated at the bottom; SD: 0.2% (WT), 4% (exon 9 5′ss con), and 3% (exon 9 3′ss Py↑) ($n \geq 3$).

and comprising only the flanking exons (Figure 1C; double-skipped species d) and a variant *PK-M2* RNA spliced via a cryptic 3′ss upstream of the authentic 3′ss of exon 10 (Figure 1C and Supplementary Figure S1; exon 10 3′ss cryptic species c).

Blocking exon 10 inclusion does not fully rescue exon 9 inclusion

Given the predominant *PK-M2* splicing of endogenous and minigene transcripts in proliferating cells, we tested whether blocking exon 10 inclusion by inactivating or weakening its splice sites might force a corresponding increase in exon 9 inclusion. The 5′ss or 3′ss of exon 10 were inactivated by mutating the invariant G residues at the intron borders, and the 3′ss was separately weakened by transversions within its upstream polypyrimidine tract (PPT) (Figure 2A).

Mutating the exon 10 5′ss resulted in a larger mRNA with an extended exon 10 (Figure 2A and Supplementary Figure S1, lanes 5–8; exon 10 5′ss cryptic species e). This mRNA resulted from splicing via a cryptic 5′ss 105 nt downstream of the authentic 5′ss of exon 10. There was no significant increase in the double-skipped RNA species d. Thus, exon 10 definition was essentially preserved, even though the normal 5′ss was inactivated.

In contrast, mutating the exon 10 3′ss largely abrogated exon 10 definition, giving a large increase in double-skipped transcripts (Figure 2A and Supplementary Figure S1, lanes 9–12; species d). A small increase in the exon 10 cryptic 3′ss species c was also observed. These results suggest that the exon 10 3′ss is essential for exon 10 definition. Similarly, weakening the exon 10 PPT led to some loss of exon 10 definition, with large increases in double-skipped RNA and in the use of the upstream cryptic 3′ss (Figure 2A, lanes 13–16; species c and d). This result implies that exon 10 inclusion is an active process, such that the loss of exon 10 activation results in non-productive splicing of the *PK-M* transcript.

With all three minigene mutants, there was only marginal recovery of exon 9 inclusion (Figure 2A, lanes 5–16). Therefore, exon 9 inclusion in proliferating cells is repressed independently of exon 10 splicing.

Strengthening the splice sites of exon 9 leads to aberrant PK-M splicing

Because blocking exon 10 splicing did not significantly rescue M1 splicing in transformed cells, we tested whether directly strengthening the exon 9 splice sites might do so. We therefore mutated the exon 9 5′ss to the consensus 5′ss sequence (Figure 2B, lanes 5–8), and strengthened the exon 9 3′ss by increasing the pyrimidine content of its upstream PPT (Figure 2B, lanes 9–12).

As expected, strengthening either the exon 9 5′ss or 3′ss led to an increase in M1 splicing. However, a new exon 9 plus exon 10 doubly-included mRNA species was also observed in both cases (Figure 2B and Supplementary Figure S1; double-spliced species f). Mutating the 5′ss to the consensus resulted in a large increase in the amount of M1 RNA and in high levels of the double-spliced mRNA (Figure 2B, lanes 5–7), whereas slightly increasing the 3′ss strength gave somewhat lower levels of these two mRNAs (lanes 9–11). The concurrent appearance of the double-spliced mRNA suggests that exon 10 definition—and by extension, exon 10 activation—is largely independent of exon 9 splicing.

Exonic splicing silencers and enhancers are key determinants of PK-M splicing ratios

To prioritize *PK-M* regions for *cis*-element analysis, we next asked whether intronic or exonic *cis*-elements play critical roles in activating exon 10 and/or repressing exon 9. First, we duplicated exon 10 in place of exon 9 in the minigene (Figure 3, lanes 9–12). If exon 9 repression depends on *cis*-elements present in intron 8 or 9, there should be inefficient use of the upstream copy of exon 10, because it would be under the influence of these repressive elements, and the pattern should be similar to that of the wild-type minigene. Instead, we observed the striking appearance of a doubly-included exon 10 mRNA species (Figure 3, lane 9; exon 10 doubly-included species g) indicating that the upstream exon 10 was still activated, regardless of its position. This finding strongly suggests that enhancer elements involved in exon 10 definition are present in the exon itself.

Similarly, we duplicated exon 9 in place of exon 10 in the minigene (Figure 3, lanes 5–8). If exon 10 splicing is normally activated through flanking *cis*-elements in intron 9 or 10, there should be a strong increase in exon-9-included transcripts, compared with the wild-type minigene, because the downstream copy of exon 9 would now be under the influence of such elements. However, there was no such increase from the exon-9-duplicated minigene transcripts (Figure 3, lanes 5–8). This finding suggests that splicing-silencing elements involved in repressing exon 9 are located in the exon itself.

When the positions of exons 9 and 10 were swapped, leaving all the introns unchanged (Figure 3, lanes 13–16), the M1 and M2 isoform ratio was similar to that of the wild-type minigene, although there was a decrease in M1 abundance, and an expected decrease in the use of the cryptic 3′ss upstream of the original exon 10 (i.e. in intron 9), because this 3′ss is now juxtaposed with the repressed exon 9 3′ss. This finding indicates that exon 10, when moved to exon 9, is spliced as efficiently as in its original location. This striking result suggests that exonic *cis*-elements involved in *PK-M* splicing are sufficient to activate exon 10 and repress exon 9, independently of their respective positions along the gene and the influence of their flanking introns.

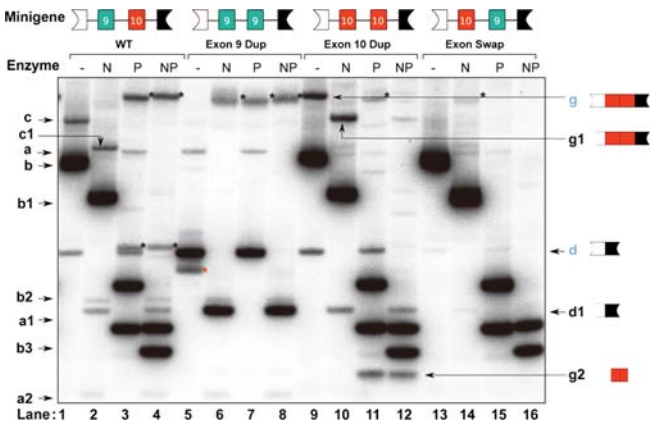

Figure 3 Effects of exonic *cis*-elements on *PK-M* alternative splicing. Mutant minigenes were analyzed by transient transfection into HEK-293 cells, followed by radioactive RT–PCR and restriction digests, as in Figure **1C**. Minigenes were constructed with clean duplications or swaps of exons 9 and 10, as shown at the top. Bands marked with black asterisks (*) or labeled on the left are as described in Figure **1C**. Key bands are labeled in blue font. The band marked with a red asterisk indicates a PCR artifact. A new exon 10–exon 10 doubly-included mRNA expressed from the exon 10 Dup minigene is indicated on the right (g, 648 nt). It generates unique fragments upon *Nco*I (g1, 571 nt) and *Pst*I (g2, 167 nt) digestion. See also Supplementary Figure S1.

To confirm these results, and to determine the role of the splice sites in ME exon use, we swapped the 3′ss or 5′ss of both exons. Computational analysis (Yeo and Burge, 2004) suggested that the 3′ss of exon 9 is weaker than those of exons 10 and 11 (Supplementary Figure S2). Exon 9 inclusion was therefore expected to increase when its 3′ss was replaced by the stronger one from exon 10. Instead, the M1 isoform abundance decreased, suggesting that proper definition of exon 9 requires its native, albeit weaker, 3′ss, for contextual reasons (Supplementary Figure S2, lanes 3 and 11). There were no significant changes when the 5′ss were swapped (lanes 3 and 7). These observations again indicate that exonic, rather intronic, *cis*-elements are the key determinants of *PK-M* splicing ratios.

A strong ESE in exon 10 is necessary and sufficient for activation of the exon

We systematically searched for critical ESEs in exon 10. Taking advantage of the high nucleotide-sequence identity between exons 9 and 10, and their identical lengths (Figure 1A), we duplicated 15–30 nt stretches of exon 10 into the corresponding location in exon 9, to find sub-exonic regions that are sufficient to activate exon 9 inclusion (Figure 4A).

The last 30 nt, but not the last 15 nt of exon 10 strongly increased exon 9 inclusion, suggesting that a strong ESE is present within, or overlaps with, the penultimate 15 nt of exon 10 (Figure 4B, compare lanes 4–6 with lanes 7–9). We then analyzed the entire 30-nt stretch using SFmap (Akerman et al., 2009; Paz et al., 2010), a method to predict splicing-regulatory motifs. This analysis yielded a near-consensus, conserved SRSF3 motif (Schaal and Maniatis, 1999) within the penultimate 15-nt segment (Supplementary Figure S3). To determine whether this SRSF3 motif alone can account for the observed M1 splicing activation, we duplicated the 7-nt SRSF3 motif from exon 10 into exon 9, by mutating the only two nucleotides that differ between exons 9 and 10 within this heptamer. Remarkably, duplicating only the SRSF3 motif, but not the flanking 8-nt or the final 15-nt region, activated exon 9 to a similar extent as that achieved by duplicating the entire 30-nt region (Figure 4B, compare lanes 10–12 with lanes 7–9 and 13–15). We conclude that the SRSF3 motif is an actual exon 10 ESE.

We then reciprocally abrogated the motif in exon 10 by mutating it into the corresponding exon 9 sequence (Figure 4C). This resulted in decreased exon 10 splicing, and a large increase in double-skipped RNA (species d), together with a concomitant increase in the M1 isoform (species a). Taken together, these data suggest that the SRSF3 motif is a *bona fide* ESE that is both necessary and sufficient for exon 10 activation.

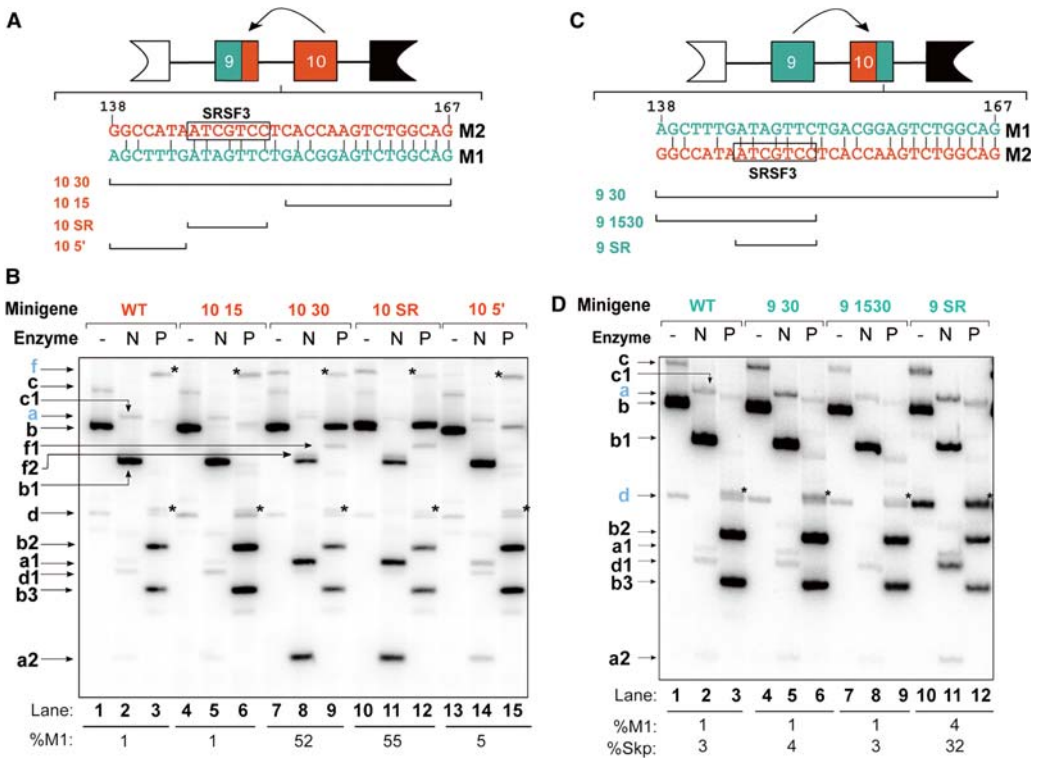

Figure 4 Mapping an ESE in exon 10. (**A**) Method used to map an ESE within the last 30 nt of exon 10. The indicated exon 9 (green) nucleotides were mutated to corresponding exon 10 (red) sequences. The last 30 nt of exon 10, when moved to exon 9, activated inclusion of exon 9. An SRSF3 SELEX motif (Schaal and Maniatis, 1999) identified by SFmap (Akerman et al., 2009) is shown by a black rectangle. Exon 10 candidate regions that were duplicated into exon 9 are indicated at the bottom, and the construct names (red) are given on the left. (**B**) The SRSF3 motif is the ESE in the last 30 nt of exon 10. The constructs from (**A**) are indicated at the top. Labeled bands are as in Figure **1C**. The %M1 inclusion is shown at the bottom ($n \geq 3$). Duplication of the exon 10 SRSF3 motif into exon 9 was sufficient to rescue exon 9 inclusion. The %M1 inclusion is indicated at the bottom; SD: 0.2% (WT), 0.1% (10 15), 5% (10 30), 4% (10 SR), 1% (10 5′) ($n = 3$). (**C**) Replacing the SRSF3 motif in exon 10 by duplicating exon 9 sequences. The indicated exon 10 (red) nucleotides were mutated to the corresponding exon 9 (green) sequences. Exon 9 regions that were duplicated into exon 10 are indicated below, and the construct names (green) are given on the left. (**D**) Inactivating the SRSF3 motif in exon 10 causes skipping of exon 10. Candidate exon 9 regions duplicated into exon 10 are indicated at the top. Labeled bands are as in Figure **1C**. %M1-included and %double-skipped (%Skp) transcripts are indicated below. Inactivating the SRSF3 motif in exon 10 is sufficient to increase the abundance of the double-skipped RNA species. Transcript-level changes are indicated at the bottom; SD: 0.3% (WT), 0.3% (9 30), 1% (9 1530), 2% (9 SR) ($n = 3$).

SRSF3 binds specifically to the exon 10 ESE

The mapped ESE resembles a consensus SRSF3 motif, so we used RNA-affinity pulldowns to ask whether SRSF3 indeed binds to this sequence. Synthetic 24-nt RNAs were covalently linked via their 3′ ends to agarose beads, incubated with HeLa nuclear extract under splicing conditions, and washed at two different salt concentrations. Bound proteins were then eluted and analyzed for SRSF3 binding by immunoblotting (Figure 5). We initially compared RNA sequences 6–29-nt upstream from the last 3′ nucleotide in both exons (Figure 5B). As a control, we mutated the exon 10 SRSF3 motif to the consensus (AUCGUCC to CUCGUCC). Coomassie-blue staining and mass spectrometry analysis revealed a different composition of bound proteins between exon 9 and exon 10 RNAs (Figure 5A, lane 2 vs lane 3). We observed similar patterns under more stringent washing conditions (300 mM instead of 150 mM KCl; data not shown). More importantly, SRSF3 bound strongly to the exon 10 RNA. Because hnRNPA1 bound to all RNAs tested, we used it as an internal loading control. Immunoblotting unambiguously revealed that SRSF3 bound only to the exon 10 and consensus SRSF3 RNAs, but not to the exon 9 RNA, at both salt concentrations (Figure 5B).

To pinpoint the precise location of SRSF3 binding in exon 10, we mutated the SRSF3 motif to the corresponding exon 9 sequence (Figure 5C). This gave a large decrease in SRSF3 binding (Figure 5C), indicating that the motif is necessary for strong SRSF3 binding to these short exon 10 RNA fragments.

SRSF3 is necessary for exon 10 inclusion

Although abrogating the SRSF3 motif in exon 10 led to greater exon 9 inclusion (Figure 4D), we could not rule out the presence of functional *cis*-elements in the corresponding exon 9 sequence that might influence the observed *PK-M1* inclusion ratio. To better assess the effect of SRSF3 on exon 10 inclusion, we tested whether knocking down SRSF3 in the context of both the wild-type minigene and the duplicated SRSF3 motif minigene mutant (10 SR minigene; Figure 4B) would enhance exon 9 inclusion, as exon 10 inclusion would be expected to decrease. Indeed, knocking down SRSF3 using two different siRNAs (Figure 5D) increased M1

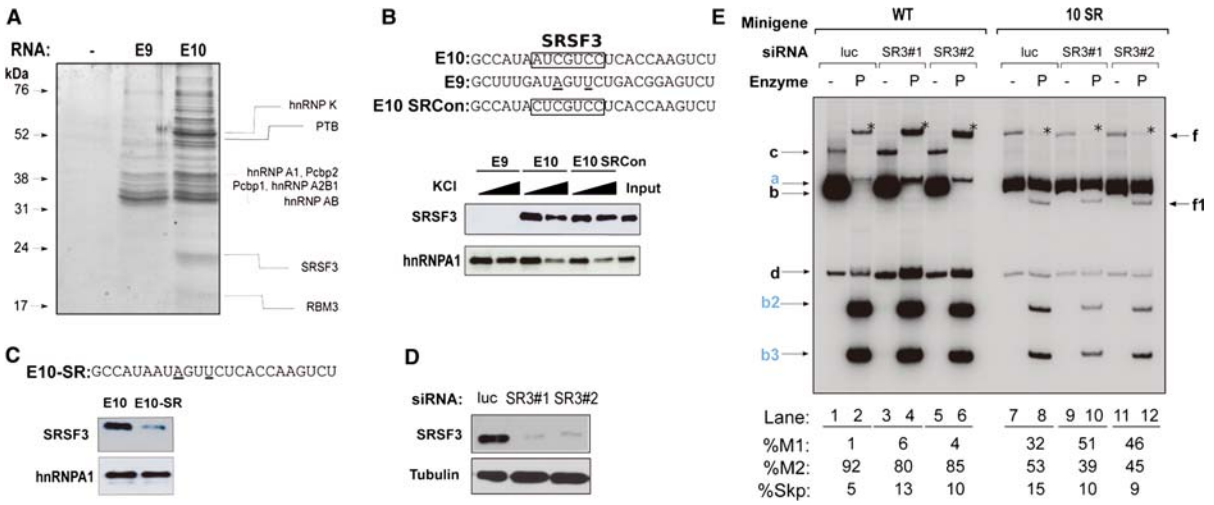

Figure 5 SRSF3 binds to the motif in exon 10 and is necessary for exon 10 inclusion. (**A**) The indicated RNAs covalently linked to agarose beads were incubated with HeLa cell nuclear extract under splicing conditions, and the beads were washed three times with buffer containing 150 mM KCl. Bound proteins were eluted with SDS and analyzed by SDS–PAGE and Coomassie-blue staining. The mobilities of size markers (M) are indicated. Prominent bands were excised and analyzed by mass spectrometry; the identified proteins are indicated on the right. (**B** and **C**) Sequences of synthetic RNA oligonucleotides used for affinity pulldowns and western-blotting analysis. Descriptive names of RNAs are indicated on the left. The SRSF3 motifs are enclosed by rectangles. Underlined nucleotides indicate the differences between the SRSF3 motif and the corresponding sequences in exon 9. Western blots of eluted proteins are shown on the right of each set of sequences. The RNAs and wash conditions are indicated at the top. Antibodies against SRSF3 and hnRNPA1 were used. HeLa nuclear extract was used as a positive control; hnRNPA1 was used as a loading control, as it binds to these short RNAs to similar extents. (**D** and **E**) The indicated SRSF3 siRNA or control luciferase siRNA was co-transfected into HEK-293 cells with the wild-type or 10 SR minigene. Knockdown of SRSF3 was verified by immunoblotting, as shown in **D**. Minigene-specific transcript-level changes for *PK-M1* and *PK-M2* are shown in **E**. siRNAs used are indicated at the top. Labeled bands are as in Figure **1C**. Asterisks indicate *Pst*I-cleaved pre-mRNA. %M1-included, %M2-included, and %double-skipped (%Skp) transcripts are indicated below, with the following SD: 0.3% (WT/luc), 1.2% (WT/SR3#1), 1.5% (WT/SR3#2), 5.7% (10 SR/luc), 3.5% (10 SR/SR3#1), and 2.7% (10 SR/SR3#2) ($n = 4$). *P*-values (Student's *t*-test) comparing %M2 from luc siRNA to that from SR3#1 and SR3#2 siRNAs co-transfected with the 10 SR minigene were 0.01 and 0.02, respectively.

inclusion and the double-skipped RNA with the wild-type minigene (Figure 5E, lanes 1–6), similar to when we mutated the SRSF3 ESE (*9* SR minigene; Figure 4D). We also observed a strong decrease in exon 10 inclusion, and a corresponding increase in exon *9* inclusion with the 10 SR minigene (Figure 5E, lanes 7–12). Thus, SRSF3 is necessary for exon 10 inclusion.

SRSF3 affects endogenous PK-M splicing

We next tested whether SRSF3 could affect endogenous *PK-M* alternative splicing; as expected, knockdown and overexpression gave reciprocal effects. SRSF3 knockdown in HEK-293 cells increased *PK-M1* mRNA to 12 folds, which was also reflected at the protein level (Figure 6A and B). In a previous study, SRSF3 knockdown led to a smaller increase in *PK-M1*, which was deemed not significant (David et al., 2010), but this could have reflected weaker knockdown and/or negative selection resulting from the use of a stably transfected shRNA. As a reciprocal experiment, we overexpressed SRSF3 in the glioblastoma cell line A172 (Supplementary Figure S4B), which expresses relatively high levels of PK-M1 (Clower et al., 2010) and low levels of SRSF3 (Supplementary Figure S4D). As expected, SRSF3 overexpression promoted an increase in M2, and a 5-fold decrease in M1 (Figure 6C).

SRSF3 activates endogenous PK-M exon 10

To determine whether the effect of SRSF3 overexpression (Figure 6C) reflects increased exon 10 activation and/or exon *9* repression, we first rescued M1 inclusion by knocking down hnRNPA1/A2 and PTB in HEK-293 cells, and then we overexpressed SRSF3 and assessed its effects on endogenous *PK-M* transcripts (Figure 6D and Supplementary Figure S4A). SRSF3 overexpression partially restored the level of *PK-M2* mRNA (Figure 6D), suggesting that SRSF3 only activates exon 10. Indeed, SRSF3 overexpression did not affect splicing of the duplicated exon *9* minigene, and SRSF3 knockdown in the context of the *9* SR minigene had no significant effect on *PK-M1/M2* splicing ratios, further suggesting that SRSF3 does not directly affect exon *9* splicing (Supplementary Figure S5).

SRSF3 is necessary for aerobic glycolysis and cellular proliferation

PK-M isoform ratios influence aerobic glycolysis (Christofk et al., 2008a), and therefore, we determined the effect of SRSF3 knockdown on this process, as assayed by the extent of cellular lactate production. SRSF3 knockdown in HEK-293 cells resulted in a significant 2-fold decrease in lactate production (Figure 6E). Because the Warburg effect also strongly influences cellular proliferation (Christofk et al., 2008a), we next assayed the effect of SRSF3 knockdown on cell growth and viability. A 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay was performed over the course of 7 days in HEK-293 cells. Strikingly, SRSF3 knockdown significantly decreased the proliferation rate. Thus, SRSF3 promotes cellular proliferation and aerobic glycolysis at least in part by influencing *PK-M* isoform ratios. Moreover, because SRSF3 knockdown

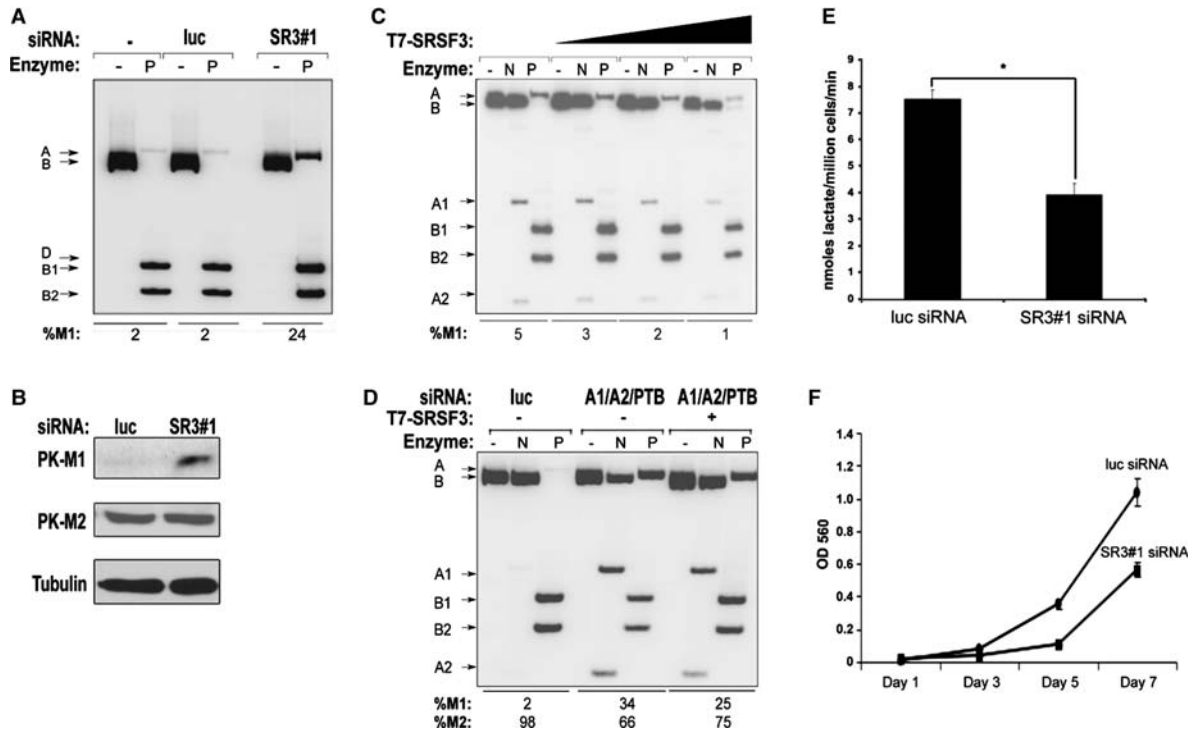

Figure 6 SRSF3 affects endogenous levels of PK-M1/M2, aerobic glycolysis, and cellular proliferation. (**A** and **B**) The indicated SRSF3 siRNA or control luciferase siRNA was transfected into HEK-293 cells. (**A**) Transcript-level changes for *PK-M1* and *PK-M2* are shown. %M1 is indicated at the bottom, with the following SD: $\leq 0.2\%$ (luc); 0.3 % (SRSF3) ($n = 3$). (**B**) Changes at the protein level, as seen in a representative blot. (**C**) T7-tagged SRSF3 cDNA was transfected in increasing amounts into A172 cells. Cells were harvested after 48 h. The *PK-M* mRNA level was determined by RT–PCR. SD $\leq 1\%$ in all cases ($n = 5$). (**D**) hnRNPA1/A2/PTB siRNAs were co-transfected into HEK-293 cells, followed by transfection of SRSF3 cDNA 24 h later. Cells were harvested 36 h after the second transfection. Transcript-level changes are indicated at the bottom, with the following SD: 1% (A1/A2/PTB siRNA); 3% (A1/A2/PTB siRNA + SRSF3) ($n = 5$). *P*-value (Student's *t*-test) comparing A1/A2/PTB siRNA with A1/A2/PTB + SRSF3 is 0.002. (**E**) The indicated SRSF3 or luciferase siRNA was transfected into HEK-293 cells. Lactate production was measured 48 h after transfection. Error bars represent SD ($n = 3$). $^*P < 0.05$ (Student's *t*-test). (**F**) Analysis of cell proliferation by the MTT assay. HEK-293 cells were transfected with luciferase or SRSF3 siRNA and grown on 96-well plates. Error bars represent the standard error of the mean. Day 1 is the next day after seeding.

mainly increased PK-M1 protein (Figure 6B), this suggests that the increase in PK-M1, rather than a decrease in PK-M2, is responsible for the effects on metabolism and proliferation.

Discussion

We have demonstrated that *PK-M* ME splicing involves a two-component circuit: exon 9 is repressed and exon 10 is activated in proliferating cells, and these two effects are essentially independent (Supplementary Figure S6).

By duplicating and swapping ME exons in a *PK-M* minigene, we showed that the key *cis*-elements controlling *PK-M* alternative splicing are located within the ME exons themselves. This is the first demonstration of exon-centric alternative splicing regulation via ME exons. As a first proof of this principle, we mapped a *bona fide* SRSF3 ESE in exon 10 that proved sufficient to activate exon 9 splicing in cancer cells when placed in this exon. Although double inclusion of exon 10 was not the major product from the exon 10-duplication minigene (Figure 3), this cannot be due to repression via the flanking intron 8 and 9 regions, because exon 10 was spliced efficiently when it was cleanly swapped with exon 9. Instead, we speculate that intronic elements are likely involved in the ME exon selection properties of *PK-M*.

Remarkably, the SRSF3 ESE motif in exon 10 differs from the corresponding exon 9 sequence by only two nucleotides, which are wobble bases in the respective codons (Supplementary Figure S3A and B). The first wobble base is highly conserved, perhaps reflecting the importance of this nucleotide in mediating SRSF3 recruitment and functionality. The corresponding two nucleotides in exon 9 are also conserved, suggesting selection against the creation of an exon 9 SRSF3 activation motif (Supplementary Figure S3B). However, we cannot rule out the existence of a corresponding exonic splicing silencer element in exon 9. The use of two wobble nucleotides to code for a key splicing signal illustrates the impact of sequence changes that would be translationally neutral, except that they drastically affect the structure of the resulting protein by changing alternative splicing of the entire exon (Cartegni et al., 2002).

Although mutating just two nucleotides in exon 10 to the corresponding exon 9 sequence (9 SR minigene) gave rise to more double-skipped and *PK-M1* mRNAs, this was not the case when the surrounding exon 9 sequence was also duplicated (9 1530 and 9 30 minigenes). This probably reflects the presence of additional *cis*-elements in exon 9, such as putative ESEs that

may compensate for the loss of exon definition upon mutation of the SRSF3 motif.

Whereas the 5′ss of exons 9 and 10 do not play a dominant role in exon selection—considering that the *M1* and *M2* mRNA levels do not change upon swapping these 5′ss (Supplementary Figure S2B)—mutational analysis indicates that the 3′ss are necessary for definition of their respective exons, and they compete with each other. Exon definition, and ultimately, proper ME exon selection in the *PK-M* gene, is apparently dependent on the outcome of competition between the alternative 3′ss. We speculate that the recovery of endogenous *PK-M1* transcripts upon SRSF3 knockdown may reflect this competition mechanism, resulting in a loss of exon 10 definition, and allowing the basal splicing machinery to be recruited to the exon 9 3′ss. Loss of exon 10 definition is also supported by the increase in double-skipped RNA from the wild-type minigene upon SRSF3 knockdown. However, it is unclear why the extent of *PK-M1* inclusion is weaker in minigene than in endogenous RNA, leading to the accumulation of unproductively spliced double-skipped RNA.

We were surprised to find no rescue of exon 9 inclusion when its 5′ss was swapped with that of exon 10—despite a report that hnRNPA1 represses exon 9 inclusion by binding to the exon 9 5′ss (David et al., 2010)—as this swap presumably removed the repressive hnRNPA1 binding site; perhaps this lack of rescue reflects contextual effects. However, our results confirm and extend an earlier study that duplicated the exon 10 5′ss in a heterologous minigene reporter, and found no change in *PK-M* splicing (Takenaka et al., 1996). Moreover, in the context of intact pre-mRNAs, it is not known how well hnRNPA1 binding to a motif that is part of a 5′ss can compete with binding of spliceosomal components, such as U1 and U6 snRNPs. Given that hnRNPA1 does affect exon 9 inclusion *in vivo* (Clower et al., 2010; David et al., 2010), hnRNPA1-induced exon 9 repression could occur either indirectly—through regulation of a splicing factor that in turn regulates *PK-M* alternative splicing—or through additional *cis*-elements located elsewhere on the *PK-M* pre-mRNA.

The change in endogenous *PK-M1* upon SRSF3 knockdown was roughly comparable to the effects of knocking down the known repressors of exon 9, hnRNPA1/A2 and PTB (Clower et al., 2010; David et al., 2010). However, knocking down these factors did not completely rescue exon 9 inclusion, and as in other systems, we anticipate the existence of additional activators of exon 10 and/or repressors of exon 9 that contribute to maintaining exon 10 definition in proliferating cells. With respect to the enhancer region we identified in exon 10, knockdown of other candidate splicing factors identified by RNA-affinity chromatography—hnRNPK and RBM3—did not change *PK-M* splicing ratios (data not shown).

Consistent with its expected ability to facilitate cellular proliferation by altering glycolytic metabolism, SRSF3 is overexpressed in ovarian cancers (He et al., 2004) and cervical cancer cell lines, whereas in normal cervical tissue, its expression is restricted to the basal proliferating layers (Jia et al., 2009). SRSF3 is also a target of the oncogenic β-catenin/TCF-4 pathway in colorectal cancer cells (Goncalves et al., 2008). Moreover, SRSF3 overexpression is sufficient for transformation of immortal mouse fibroblasts (Jia et al., 2010), indicating that similar to its paralog, SRSF1 (Karni et al., 2008), SRSF3 is an oncoprotein. Given the multitude of potential SRSF3 downstream target genes, we expect that SRSF3-mediated tumorigenesis reflects splicing changes in multiple effector genes, including *PK-M*.

Important unanswered questions include: what are the additional factors that govern exon 9 and 10 use in tumor cells? Can the PK-M2 isoform be completely switched to the PK-M1 isoform by manipulating splicing factor levels in tumor cells? How is exon 9 preferentially selected in quiescent cells? Are there tissue-specific differences in the mechanisms of exon 9 selection in differentiated cells? Answers to these questions could provide the basis to develop splicing-targeted cancer therapeutics.

Materials and methods

Cells and transfections

HeLa and HEK-293 cells were grown and transiently transfected as described (Clower et al., 2010). Total RNA was harvested after 36 h.

RNA interference

Two siRNAs targeting human *SFRS3* (Sigma Genosys) have the sense-strand sequences 5′-CGAUCUAGGUCAAAUGAAA-3′ (SR3#1) and 5′-CGUAGUCGAUCUAGGUCAA-3′ (SR3#2). siRNAs against hnRNPA1, hnRNPA2, and PTB were used as described (Cartegni et al., 2006; Clower et al., 2010). HEK-293 cells (1×10^6) in six-well plates were transfected with 200 pmol of siRNA duplex using Lipofectamine 2000 (Invitrogen). Cells were harvested 48 h later.

Immunoblotting

Immunoblotting was carried out as described (Clower et al., 2010). Primary antibodies were: β-tubulin (Genscript rAb, 1:5000), hnRNPA1 (mAb UP1-55, culture supernatant) (Hua et al., 2008), SRSF3 (Zymed mAb, 1:1000), PK-M2 (rabbit, 1:2000), and PK-M1 (rabbit, 1:2000) (Christofk et al., 2008a, b). Secondary antibodies were goat anti-mouse or anti-rabbit HRP conjugates (1:20000; Bio-Rad).

Plasmids

DNA oligonucleotides were obtained from Sigma Genosys. The *PK-M2* minigene was constructed by amplifying a 6.4-kb *PK-M* exon 8–11 fragment from human genomic DNA (Promega) using Phusion High-Fidelity DNA Polymerase (Finnzymes) and primers PKMminigeneF (5′-GGGGAAAGATCTGCCACCATGGGAGAAACAGCCAAAGGGGAC-3′) and PKMminigeneR (5′-GGGGAACTCGAGCTAGACATTCATGGCAAAGTTCACC-3′). The product was cloned between the *Bam*HI and *Xho*I sites of pcDNA3.1(+) (Invitrogen). See Supplementary experimental procedures for the generation of mutant, exonic and sub-exonic duplication, exon-swap, and splice-swap constructs.

RT–PCR

Radioactive RT–PCR was carried out as described (Clower et al., 2010). The human-specific primer sets used to amplify endogenous transcripts anneal to *PK-M* exons 8 and 11, and their sequences are: hPKMF: 5′-AGAAACAGCCAAAGGGGACT-3′; hPKMR: 5′-CATTCATGGCAAAGTTCACC-3′. To amplify minigene-specific transcripts, the forward primer was replaced with a primer annealing to the pcDNA3.1(+) vector, pcDNAF: 5′-TAATACGACTCACTATAGGG-3′. After 26 amplification cycles for

minigene-derived transcripts, and 24 cycles for endogenous transcripts, the reactions were divided into four aliquots for digestion with *Nco*I, *Pst*I (New England Biolabs), both, or neither. The products were analyzed on a 5% native polyacrylamide gel, visualized by autoradiography, and quantified on a FLA-5100 phosphoimager (Fuji Medical Systems) using Multi Gauge software Version 2.3. All the PCR products were gel-purified, cloned, and sequenced to verify their identities.

RNA-affinity chromatography and mass spectrometry

RNA-affinity chromatography was performed as described (Caputi et al., 1999; Hua et al., 2008). After the final wash, the beads were resuspended in 4× Laemmli buffer and boiled for 5 min to elute bound proteins. After SDS–PAGE and Coomassie-blue staining, prominent bands were excised for in-gel trypsin digestion and analysis by MALDI-TOF mass spectrometry. Spectra were analyzed with the MASCOT search engine.

Lactate assay

For measurements of lactate secretion, cells were transfected with siRNA in six-well plates. Twenty-four hours later, the cells were replated (three replicates per condition) at subconfluent density (25000 cells/well) in 12-well dishes, and after 24 h, the cells were switched to serum-free medium without phenol red for 20 min, and lactate secreted into the medium was measured, in triplicate for each sample, using a Lactate Assay Kit II (Biovision Inc.). Optical density readings at 450 nm were averaged for each sample replicate set, then averaged for each condition replicate set, and finally normalized to the cell number measured from parallel wells.

MTT assay

HEK-293 cells (3×10^6) were transfected with 400 pmol of luciferase control or SRSF3 siRNA in a 6-cm dish. After 24 h, cells were seeded into 96-well plates (2500 cells/well). The next day, and every 2 days thereafter, MTT (Sigma) was added to fresh medium at a final concentration of 0.5 mg/ml. Cells were incubated at 37°C for 4 h and then solubilized with DMSO. Optical density at 560 nm was then determined for each well and averaged.

Supplementary material

Supplementary material is available at *Journal of Molecular Cell Biology* online.

Acknowledgements

We thank Xavier Roca and Yimin Hua (Cold Spring Harbor Laboratory), and Cynthia Clower (Harvard Medical School) for helpful discussions.

Funding

Z.W. is supported by a fellowship from the Agency for Science, Technology and Research, Singapore. This work was supported by grant I1-A34 from the Starr Cancer Consortium to A.R.K., L.C.C., and M.G.V.H.

Conflict of interest: none declared.

References

Akerman, M., David-Eden, H., Pinter, R.Y., et al. (2009). A computational approach for genome-wide mapping of splicing factor binding sites. Genome Biol. *10*, R30.

Caputi, M., Mayeda, A., Krainer, A.R., et al. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. EMBO J. *18*, 4060–4067.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat. Rev. Genet. *3*, 285–298.

Cartegni, L., Hastings, M., Calarco, J.A., et al. (2006). Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. Am. J. Hum. Genet. *78*, 63–77.

Chacko, E., and Ranganathan, S. (2009). Comprehensive splicing graph analysis of alternative splicing patterns in chicken, compared to human and mouse. BMC Genomics *10(Suppl 1)*, S5.

Christofk, H., Vander Heiden, M., Harris, M.H., et al. (2008a). The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. Nature *452*, 230–233.

Christofk, H., Vander Heiden, M., Wu, N., et al. (2008b). Pyruvate kinase M2 is a phosphotyrosine-binding protein. Nature *452*, 181–186.

Clower, C.V., Chatterjee, D., Wang, Z., et al. (2010). The alternative splicing repressors hnRNP A1/A2 and PTB influence pyruvate kinase isoform expression and cell metabolism. Proc. Natl Acad. Sci. USA *107*, 1894–1899.

David, C.J., Chen, M., Assanah, M., et al. (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. Nature *463*, 364–368.

Dombrauckas, J.D., Santarsiero, B.D., and Mesecar, A.D. (2005). Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. Biochemistry *44*, 9417–9429.

Goncalves, V., Matos, P., and Jordan, P. (2008). The beta-catenin/TCF4 pathway modifies alternative splicing through modulation of SRp20 expression. RNA *14*, 2538–2549.

He, X., Ee, P.L., Coon, J.S., et al. (2004). Alternative splicing of the multidrug resistance protein 1/ATP binding cassette transporter subfamily gene in ovarian cancer creates functional splice variants and is associated with increased expression of the splicing factors PTB and SRp20. Clin. Cancer Res. *10*, 4652–4660.

Hitosugi, T., Kang, S., vander Heiden, M.G., et al. (2009). Tyrosine phosphorylation inhibits PKM2 to promote the Warburg effect and tumor growth. Sci. Signal. *2*, ra73.

Hua, Y., Vickers, T.A., Okunola, H.L., et al. (2008). Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am. J. Hum. Genet. *82*, 834–848.

Jia, R., Liu, X., Tao, M., et al. (2009). Control of the papillomavirus early-to-late switch by differentially expressed SRp20. J. Virol. *83*, 167–180.

Jia, R., Li, C., McCoy, J.P., et al. (2010). SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. Int. J. Biol. Sci. *6*, 806–826.

Karni, R., Hippo, Y., Lowe, S.W., et al. (2008). The splicing-factor oncoprotein SF2/ASF activates mTORC1. Proc. Natl Acad. Sci. USA *105*, 15323–15327.

Letunic, I., Copley, R.R., and Bork, P. (2002). Common exon duplication in animals and its role in alternative splicing. Hum. Mol. Genet. *11*, 1561–1567.

Noguchi, T., Inoue, H., and Tanaka, T. (1986). The M1- and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. J. Biol. Chem. *261*, 13807–13812.

Paz, I., Akerman, M., Dror, I., et al. (2010). SFmap: a web server for motif analysis and prediction of splicing factor binding sites. Nucleic Acids Res. *38(Suppl)*, W281–W285.

Schaal, T.D., and Maniatis, T. (1999). Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. Mol. Cell. Biol. *19*, 1705–1719.

Smith, C.W. (2005). Alternative splicing—when two's a crowd. Cell *123*, 1–3.

Smith, C.W., and Nadal-Ginard, B. (1989). Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. Cell *56*, 749–758.

Takenaka, M., Yamada, K., Lu, T., et al. (1996). Alternative splicing of the pyruvate kinase M gene in a minigene system. Eur. J. Biochem. *235*, 366–371.

Vander Heiden, M., Cantley, L., and Thompson, C.B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science *324*, 1029–1033.

Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J. Comput. Biol. *11*, 377–394.