MOLECULAR AND CELLULAR BIOLOGY, Sept. 1995, p. 4825–4834 Vol. 15, No. 9
0270-7306/95/$04.00+0

Exon and Intron Sequences, Respectively, Repress and Activate Splicing of a Fibroblast Growth Factor Receptor 2 Alternative Exon

FABIENNE DEL GATTO AND RICHARD BREATHNACH*

Institut National de la Santé et de la Recherche Médicale U211, Institut de Biologie-Centre Hospitalier Régional, 44035 Nantes Cedex 01, France

Received 17 February 1995/Returned for modification 13 April 1995/Accepted 20 June 1995

Two alternative exons, BEK and K-SAM, code for part of the ligand binding site of fibroblast growth factor receptor 2. Splicing of these exons is mutually exclusive, and the choice between them is made in a tissue-specific manner. We identify here pre-mRNA sequences involved in controlling splicing of the K-SAM exon. The short K-SAM exon sequence 5′-TAGGGCAGGC-3′ inhibits splicing of the exon. This inhibition can be overcome by mutating either the exon's 5′ or 3′ splice site to make it correspond more closely to the relevant consensus sequence. Two separate sequence elements in the intron immediately downstream of the K-SAM exon, one of which is a sequence rich in pyrimidines, are both needed for efficient K-SAM exon splicing. This is no longer the case if either the exon's 5′ or 3′ splice site is reinforced. Furthermore, if the exon inhibitory sequence is removed, the intron sequences are not required for splicing of the K-SAM exon in a cell line which normally splices this exon. At least three elements are thus involved in controlling splicing of the K-SAM exon: suboptimal 5′ and 3′ splice sites, an exon inhibitory sequence, and intron activating sequences.

Alternative splicing is a mechanism frequently used by cells for the production of several proteins from one gene (29, 33). When subject to control, this mechanism allows the production of similar but distinct versions of a protein in a cell-type-specific manner. Alternative splicing of the fibroblast growth factor receptor 2 (FGFR-2) gene pre-mRNA is a particularly attractive example of this phenomenon (9, 21, 30, 50). The FGFR-2 extracellular domain is composed of three immunoglobulin-like domains, the third of which is particularly important for ligand binding (30, 50). The carboxy-terminal moiety of the third immunoglobulin-like domain is encoded by two alternative exons (9, 30, 50), K-SAM (or IIIb) and BEK (or IIIc). Splicing of these exons is mutually exclusive and is controlled both in cell lines and during development (9, 30, 31). Epithelial cells use mainly the K-SAM exon and synthesize a receptor with high affinity for keratinocyte growth factor. Other cell types use mainly the BEK exon and synthesize a high-affinity receptor for basic fibroblast growth factor. That the splicing choice is linked to the cell type is illustrated by the observation that oncoprotein fos activation in epithelial cells induces an epithelio-mesenchymal conversion and changes the exon spliced from K-SAM to BEK (36). Controlling the splicing of the FGFR-2 pre-mRNA may be very important for the cell, as a change in the splicing pattern has been correlated with malignant progression of rat prostate cells (48).

This control is presumably exerted at the level of construction of the spliceosome, which is composed of four small nuclear ribonucleoproteins (snRNPs) as well as many associated splicing factors (19, 37). Early steps in spliceosome formation are the recognition of the 5′ splice site sequence by U1 snRNP and the recognition of the branch point sequence, associated with the 3′ splice site, by U2 snRNP. These steps are assisted by other splicing factors, several of which contain a domain rich in Arg-Ser dipeptides and are members of the SR protein family (28, 51). Thus the splicing factor U2AF, whose 35-kDa U2AF subunit contains an SR domain (53), binds to the polypyrimidine tract usually found immediately upstream of 3′ splice sites and helps binding of U2 snRNP to the branch point sequence (52). On the other hand it has been shown recently (23) that the SR protein ASF/SF2 directly recognizes the 5′ splice site and cooperates with U1 snRNP in binding to the pre-mRNA.

SR proteins thus help to define individual 5′ and 3′ splice sites. The exon definition model for splicing (34) proposes that the exon is recognized as an entity, and this requires coupled definition of the 5′ and 3′ splice sites. The SR proteins SC35 and ASF/SF2 can interact specifically with both U1 snRNP and U2AF and could thus be involved in coordinating the binding of appropriate snRNPs to 3′ and 5′ splice sites of an exon (46). Such coordinated binding occurs in the splicing of a rat preprotachykinin gene exon (17), in which U1 snRNP bound to the 5′ splice site facilitates the binding of U2AF to the 3′ splice site.

Splicing thus appears to involve the organization of a network of protein-protein and protein-RNA interactions spanning exons. Splicing could clearly be controlled by blocking or accelerating the setup of these complex interactions. In *Drosophila melanogaster*, sex-lethal (sxl) protein represses a male-specific 3′ splice site on the transformer (tra) pre-mRNA by binding to its associated polypyrimidine sequence and blocking binding of U2AF (44). The tra and tra-2 proteins activate a female-specific 3′ splice site on the doublesex (dsx) pre-mRNA by binding of a multiprotein complex containing SR proteins to a repeated sequence within the downstream exon. The resulting complex of tra, tra-2, and SR proteins stabilizes the binding of U2AF to the 3′ splice site polypyrimidine sequence (2, 42, 46). In mammals, purine-rich exonic splicing enhancer sequences may function similarly (25, 40).

The heterogeneous nuclear RNP (hnRNP) A1 protein counteracts the effects of the SR protein ASF/SF2 in a concentration-dependent manner both in vivo and in vitro (7, 11, 24, 27, 49). ASF/SF2 promotes the use of proximal 5′ splice sites and

* Corresponding author. Mailing address: INSERM U211, Institut de Biologie-CHR, 9 Quai Moncousu, 44035 Nantes Cedex 01, France. Phone: 40 08 47 50. Fax: 40 35 66 97.

exon inclusion, while hnRNP A1 promotes the use of distal 5′ splice sites and exon exclusion. The relative levels of the two proteins may vary in different cell types and be partly responsible for controlling alternative splicing of some pre-mRNAs. Exactly how hnRNP A1 counteracts the effects of ASF/SF2 is unknown; however, hnRNP A1 is a sequence-specific RNA-binding protein (6) and could disrupt or prevent interactions between SR proteins such as ASF/SF2 and other spliceosome components.

For the FGFR-2 pre-mRNA, we have shown previously that control is exerted on both the alternative exons (12). Deletion of the K-SAM exon does not lead to efficient use of the BEK exon in its place and vice versa. The K-SAM exon 3′ splice site is associated with a polypyrimidine sequence containing several purines, and thus it is unlikely to bind U2AF efficiently. Changing these purines to pyrimidines, or replacing the K-SAM exon polypyrimidine tract with that of the BEK exon, leads to efficient splicing of the K-SAM exon in cells which normally use preferentially the BEK exon. In this case the K-SAM exon is spliced to the BEK exon. These observations led us to hypothesize that efficient splicing of the K-SAM exon requires the action of an activator, which could be a sequence-specific RNA-binding protein capable of recruiting U2AF to the 3′ splice site, for example. We set out to test this hypothesis by searching for *cis*-acting RNA sequences whose mutations block splicing of the K-SAM exon. While *cis*-acting sequences necessary for K-SAM exon splicing do indeed exist, we discovered in the course of this work an additional level of complexity: another *cis*-acting sequence inhibits splicing of the K-SAM exon.

MATERIALS AND METHODS

Minigenes. A *Pvu*II fragment containing the neomycin resistance gene was deleted from pRcRSV (Invitrogen) to make pRK1. The 5.6-kb *Bam*HI-*Eco*RI fragment from pBK3 (12), which contains the C1, K-SAM, BEK, and C2 exons, was subcloned into the corresponding sites of pBluescript (Stratagene). A *Bam*HI-*Sal*I fragment of the resulting plasmid was cloned between the *Bam*HI and *Xho*I sites of pRK1 to make pRK3 (see Fig. 1). To generate pRK2, an *Ava*I-*Xba*I fragment from pRK3, containing most of the K-SAM exon and some downstream intron sequences (see Fig. 1), was cloned between the *Ava*I and *Xba*I sites of a pBluescript derivative in which the *Kpn*I-*Bam*HI polylinker fragment had been replaced by a new polylinker containing a compatible *Ava*I site. To make pRK4, a *Sal*I site was introduced into the K-SAM exon of pRK2 with the Transformer site-directed mutagenesis kit from Clontech and the mutation oligonucleotide 5′-ACCTTGCTGTTTTGGGTCGACAGTGAGCCAG GC-3′.

A variety of other minigenes were made from pRK2 or pRK4. The *Ava*I-*Xba*I fragments of all mutated versions of these plasmids were verified by sequencing and then transferred back into pRK3 to make mutated minigenes. For pRK8 and pRK11, the *Ava*I-*Sal*I fragment of pRK4 was replaced, respectively, by *Ava*I-*Sal*I fragments obtained by PCR amplification of (i) a rabbit beta-globin gene (32) (amplified with the primers 5′-GGCTCGGGGCTGGTTGTCTACCCAT-3′ and 5′-GGGTCGACGAAGGCAGCCAGCACCT-3′) and (ii) a bacterial chloramphenicol acetyltransferase (CAT) gene (pCAT-Basic plasmid; Promega) (amplified with the primers 5′-GGCTCGGGCTTTATTCACATTCTTG-3′ and 5′-GG GTCGACTTGCTCATGGAAAACGG-3′) followed by *Ava*I and *Sal*I digestion. pRK8 thus contains bp 4 to 125 of the second exon of the globin gene between the *Ava*I and *Sal*I sites of the K-SAM exon, while pRK11 contains codons 59 to 98 of the CAT gene similarly placed. pRK10 was made by site-directed mutagenesis of the K-SAM exon-intron junction in pRK2 with the mutation oligonucleotide 5′-GACAAAAATGAAAGCTTACTTACCTTGCTGT-3′. pRK12 was made from pRK11 by site-directed mutagenesis with the mutation oligonucleotide 5′-GAGCTGGTGATATCGGATAGTGTTCA-3′ to introduce an *Eco*RV site in the CAT sequences. pRK13, pRK14, and pRK15 (and derivatives thereof) were made by replacing, respectively, the *Ava*I-*Eco*RI, *Eco*RI-*Eco*RV, and *Eco*RV-*Sal*I fragments of pRK12 with double-stranded oligonucleotides with compatible extremities derived from the K-SAM exon sequences as described in the legend to Fig. 2. pRKS10 was made with appropriate oligonucleotides by replacing 10 bp of the CAT sequences carried by the *Eco*RV-*Sal*I fragment of pRK12 with the K-SAM exon sequence 5′-TAGGGCAGGC-3′ (in double-stranded form). In pRKS10, the K-SAM sequence is situated relative to the exon's 5′ and 3′ splice sites as it is in pRK3, but in pRKS10 it is surrounded by CAT sequences. pRK18 was made by replacing the *Ava*I-*Xba*I fragment of pRK2 with a fragment obtained by PCR amplification of pRK3 with the primers 5′-CTGTTCTAGCACTCGGGG-3′ and 5′-CCTCTAGACATTGTTACCTTG CTGTT-3′ followed by *Ava*I and *Xba*I digestion. pRK20 was made by annealing 5′-CCATGCATTCGCGTGCCATGCCATTGTTACCTTGCTGTTTTGGC-3′ and 5′-TCGAGCCAAAACAGCAAGGTAACAATGGCATGGCACGCGAA TGCATGG-3′ and ligating them to a *Sal*I-*Dra*I digest of pRK4. pRK22 was made similarly, with 5′-TAATAATGACAATAATGATAGCATTGTTACCTT GCTGTTTTGGC-3′ and 5′-TCGAGCCAAAACAGCAAGGTAACAATGC TATCATTATTGTCATTATTA-3′. pRK21 was made by replacing a *Dra*I-*Pst*I fragment of pRK2 with a polylinker fragment sequence of the same size. pRK23 was made by *Pst*I and *Xba*I digestion of a fragment obtained by PCR amplification of an elongation factor 1 alpha (43) gene intron (bp 417 to 827) with the oligonucleotides 5′-GGTCTAGACGCACGCGGCGGCCCCAG-3′ and 5′-AC AGTCCCCGAGAAGTT-3′ followed by ligation of the digest to a *Pst*I-*Xba*I digest of pRK2.

pRK9 was made by replacing the *Bam*HI-*Ava*I fragment of pRK3 with the corresponding fragment of s(t)SbB (12). pRK16 was made from RK3 by deleting a *Kpn*I-*Bgl*II fragment (bp 24 to 686) of the 1,149-bp intron between exons C1 and K-SAM. pRK17 was made from s(g)SbB (12) by deleting a *Bgl*II-*Sal*I fragment (bp 690 to 1105) of the 1,149-bp intron between exons C1 and K-SAM followed by cloning of the *Bam*HI-*Ava*I fragment containing exon C1, the deleted intron, and the start of the K-SAM exon, between the corresponding sites of pRK3. A fragment obtained by PCR amplification of pRK3 with the oligonucleotide 5′-TGTCTAGATTTTTGTCTTTTTTTAAAAAG-3′ and T3 primer was digested with *Xba*I and ligated to *Xba*I-digested pRK18 to make pRK19. To make pRK25, the *Ava*I-*Hpa*I fragment of pRK18 replaced the corresponding *Ava*I-*Hpa*I fragment of pRK9 containing a reinforced K-SAM exon polypyrimidine sequence. A *Hin*dIII-*Xba*I fragment of pRK10 was deleted to make pRK26 (in pRK10 there is a *Hin*dIII site just downstream of the K-SAM exon 5′ splice site [see the legend to Fig. 4]). pRK27 and pRK28 were made by ligation of oligonucleotides 5′-CTAGACATTGTTACCTTGCTGTTTTGGC-3′ and 5′-TCGAGCCAAAACAGCAAGGTAACAATGT-3′ to *Sal*I-*Xba*I digests of pRK8 and pRK11, respectively.

RNA analysis. SVK14 or HeLa cells (as described in reference 12) were transfected by the calcium phosphate coprecipitation technique (3). For stable transfections, minigenes (20 μg) and pSGneo (2 μg) were cotransfected into cells. Geneticin (800 μg/ml) was added to culture medium at 18 h posttransfection. RNA was harvested from bulk cultures of geneticin-resistant cells by a rapid cytoplasmic RNA extraction technique (15). For transient transfections of HeLa cells, cells were transfected with 20 μg of minigenes, and RNA was harvested at 60 h posttransfection.

The reverse transcription PCR (RT-PCR) protocol used was as described previously, except that primers P1 (5′-GGAATTCACGAGCGATCGCCTCAC CGG-3′, from the C1 sequence) and P2 (5′-GGCAACTAGAAGGCACAG-3′, from the bovine growth hormone [BGH] sequence) were used and 20 cycles of amplification were routinely used, so as to remain within the exponential range of amplification. The results obtained remained, however, essentially unchanged if 30 cycles of amplification were carried out. This was done to obtain amounts of DNA visible by ethidium bromide staining to facilitate the subcloning of PCR products into pBluescript for sequencing. Otherwise, Southern blotting was used to visualize the PCR products as described previously (12). For some experiments, suitably exposed autoradiograms were scanned with an LKB UltroScanXL laser densitometer to determine the percentage of the hybridization signal represented by each PCR product. For key experiments, RT-PCR results were confirmed by Northern (RNA) blotting. Analysis of RNA samples separated by electrophoresis on a denaturing agarose gel was carried out according to established procedures (3).

RESULTS

We have previously described the minigenes BK1 and BK3, containing the alternative K-SAM and BEK exons together with flanking exon sequences (12). We used these minigenes to investigate the splicing of the FGFR-2 pre-mRNA following the transient transfection of BK1 and BK3 into HeLa cells (which splice preferentially the BEK exon) or SVK14 cells (which splice preferentially the K-SAM exon). Correct splicing of minigene-derived pre-mRNA in SVK14 cells required the presence of an open reading frame. Subsequent experiments showed that this requirement was linked to the transfection technique used: minigenes stably transfected into SVK14 cells showed no such requirement (data not shown). We can propose no satisfactory explanation for this difference. Consequently, we decided in our further work to stably transfect SVK14 cells. To facilitate these experiments, we prepared a new minigene, RK3 (Fig. 1A). In RK3, the simian virus 40 early promoter and beta-globin gene polyadenylation site of

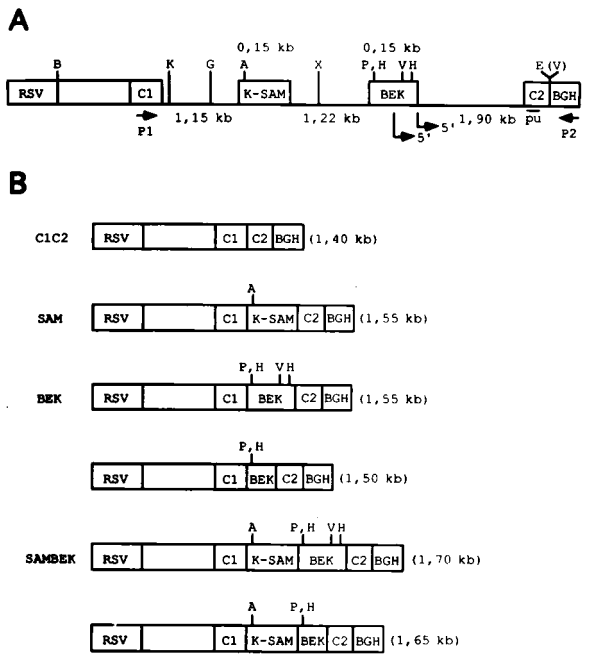

FIG. 1. Minigene RK3 and possible spliced transcripts derived from it. (A) Minigene RK3. RSV is the Rous sarcoma virus long terminal repeat promoter. BGH is the BGH polyadenylation site. K-SAM, BEK, C1, and C2 exons as described in reference 12 are shown. The sizes of introns between the C1 and K-SAM exons, between the K-SAM and BEK exons, and between the BEK and C2 exons are marked below the map. The sizes of the BEK and K-SAM exons are marked above the map. FGFR-2 cDNA sequences (12) lie between the RSV and C1 sequences. The origins of primers P1 and P2 used for RT-PCR analysis of minigene transcripts are shown. pu indicates a long purine-rich sequence in the C2 exon discussed in the text. For the BEK exon, two alternative 5′ splice sites are shown. The downstream site (5′-CAGgtatat-3′) is used most frequently; the upstream site (5′-CGGgtaatt-3′) is a minor alternative site whose use has been described occasionally (35). A, *Ava*I; B, *Bam*HI; G, *Bgl*II; H, *Hin*cII; K, *Kpn*I; P, *Hpa*I; E, *Eco*RI; V, *Eco*RV; X, *Xba*I. The *Eco*RV site between the C2 exon and the BGH sequence is shown in parentheses as it does not exist in all minigenes (Table 1). (B) Structures and sizes of various possible spliced transcripts of minigene RK3 and its derivatives. The transcripts are represented as double-stranded cDNAs. Note the existence of two BEK transcripts and two SAMBEK transcripts, linked to the use of the two alternative BEK exon 5′ splice sites. Restriction enzyme abbreviations are as listed above for panel A.

BK3 were replaced with the stronger Rous sarcoma virus long terminal repeat promoter and the BGH polyadenylation site, respectively.

In our search for *cis*-acting sequences implicated in the splicing of the K-SAM exon, various versions of RK3, carrying different mutations in the K-SAM exon and flanking intron sequences, were made. These minigenes were transfected into SVK14 or HeLa cells. RNA was harvested and subjected to RT-PCR. Primers used were from the upstream flanking exon (P1) or the BGH sequences (P2) and were designed to amplify specifically cDNA derived from minigene transcripts. A number of such transcripts are possible, depending on the splicing choices (Fig. 1B). Note that for the BEK exon, two alternative 5′ splice sites are available (Fig. 1A; see also reference 35). The downstream 5′ splice site is, however, used preferentially.

The calculated sizes of the PCR product obtained from each transcript and those of the fragments obtained on digestion of the PCR products with the restriction enzymes *Ava*I, *Hin*cII, *Hpa*I, and *Eco*RV are listed in Table 1. All four enzymes recognize sites in either the K-SAM or the BEK exon (Fig. 1). This information was used for the identification of PCR products actually obtained in the experiments described below. The

TABLE 1. PCR products obtained from minigene transcripts[a]

PCR product	Size (kb) before digestion	Size (kb) after digestion with:		
		*Ava*I	*Eco*RV	*Hin*cII-*Hpa*I
C1C2	0.32	0.32	0.32[b]	0.32
SAM	0.46	0.26 + 0.20	0.46[b]	0.46[c]
BEK	0.46	0.46	0.31 + 0.15[b]	0.21 + 0.13[d] + 0.12[d]
	0.41	0.41	0.41[b]	0.21 + 0.20
SAMBEK	0.61	0.41 + 0.20	0.46 + 0.15[b]	0.36[e] + 0.13[d] + 0.12[d]
	0.56	0.36 + 0.20	0.56[b]	0.36[e] + 0.20

[a] Minigene transcripts are shown in Fig. 1.
[b] This fragment is further cleaved to liberate a 0.07-kb fragment, representing the BGH sequence, for minigenes RK8, RK10, RK17, and RK18.
[c] This fragment is further cleaved by *Hin*cII to liberate a 0.13-kb fragment for minigenes with a *Sal*I site in the K-SAM exon (RK4, RK5, RK7, RK8, RK11, RK12, RK13, RK14, RK15, RK20, RK22, RK27, and RK28).
[d] This fragment is replaced with a 0.25-kb fragment in *Hpa*I digests.
[e] This fragment is further cleaved by *Hin*cII to liberate a 0.03-kb fragment for minigenes with a *Sal*I site in the K-SAM exon (see footnote *c*).

*Ava*I and *Hpa*I sites are particularly useful, as only PCR products containing the K-SAM exon will be digested by *Ava*I (yielding a characteristic 0.20-kb fragment), while only PCR products containing the BEK exon will be digested by *Hpa*I. On the other hand, PCR products containing neither exon are not cut by *Ava*I, *Hpa*I, or *Hin*cII.

No evidence that internal K-SAM exon sequences are needed for splicing. We began by searching for exon sequences which might be involved in splicing of the K-SAM exon, as some purine-rich sequences within exons have been shown to act as splicing enhancers. One such sequence (AAAACAGC AAG) can be found juxtaposed to the K-SAM exon 5′ splice site (Fig. 2). Another such sequence (GGAAGAGAAAAGG AGA) lies in the downstream C2 exon, 5 nucleotides from the 3′ splice site (Fig. 1A). While each sequence could work on its own, splicing of the K-SAM exon's 5′ splice site to the C2 exon's 3′ splice site creates an even longer purine-rich stretch, which might be involved in the activation of the K-SAM exon's 3′ splice site. These purine-rich sequences were deleted from RK3 (Fig. 3A) either individually or together. RK3 and these mutated minigenes were stably transfected into SVK14 cells and the RT-PCR analysis described above was carried out. Results (Fig. 3B) were interpreted (Fig. 3C) by using the data provided in Table 1. Levels of PCR products were expressed as the percentage of the total hybridization signal (see Materials and Methods).

With the RK3 minigene, 0.46-kb PCR products containing an *Ava*I site and thus representing SAM transcripts predominated, although some 0.32-kb products representing C1C2 transcripts (less than 5% of the hybridization signal) were also detected. The deletion of either the K-SAM or the C2 exon's purine-rich sequence or both of these sequences had no very marked effect on K-SAM exon splicing, as PCR products of 0.46-kb with an *Ava*I site continued to predominate (data not shown).

To search for possible *cis*-acting activating sequences elsewhere in the K-SAM exon, a *Sal*I site (and thus also a *Hin*cII site) was first introduced into the K-SAM exon (Fig. 2) to generate the minigene RK4. We replaced the *Ava*I-*Sal*I fragment of RK4 with a globin exon sequence of the same size (Fig. 3A, minigene RK8). The reading frames of the K-SAM exon and the globin exon were conserved. Results of the RT-PCR analysis carried out on RNA from SVK14 cells stably

```
  1  cac


     AvaI                                                  EcoRI
  4  tcggggataaatagttccaatgcagaagtgctggctctgttcaatg           Seg. 1


     EcoRI                                EcoRV
     AATT                                 A
 50  tgaccgaggcggatgctggggaatatatatgtaaggt                    Seg. 2


     EcoRV
     A                                           SalI
 87  ctccaattatatagggcaggccaaccagtctgcctggctcactg             Seg. 3
      1---------         3------      5---------
                2-------         4--------


       GAC
135  tcctgccaaaacagcaag
         -------
         Purines
```

FIG. 2. Sequence of K-SAM exon. This presentation is designed to facilitate the identification of K-SAM exon features and segments discussed in the text. The true exon sequence is shown in lowercase letters. The position of the pRK4 *Sal*I site, as well as that of the mutation introduced to create it (GAC), is shown above the sequence. The purine-rich sequence deleted in pRK5 is indicated with dashes and the word "Purines" below. The K-SAM exon segments (Seg.) 1, 2, and 3, present in pRK13, pRK14, and pRK15, respectively, are identified to the right. Segments 2 and 3 do not have exactly the same sequence as the K-SAM exon. To obtain the actual sequences of segments 2 and 3, replace the bases in lowercase letters with the bases in uppercase letters immediately above them. These replacements create the *Eco*RI, *Sal*I, and *Eco*RV sites shown. In pRK15-1, -2, -3, -4, and -5, the sequences of segment 3 (correspondingly numbered), indicated with dashes below, have been replaced with 5′-gcttaggcc-3′, 5′-caaggat-3′, 5′-agctgg-3′, 5′-aagcttga-3′, and 5′-atcgataac-3′, respectively.

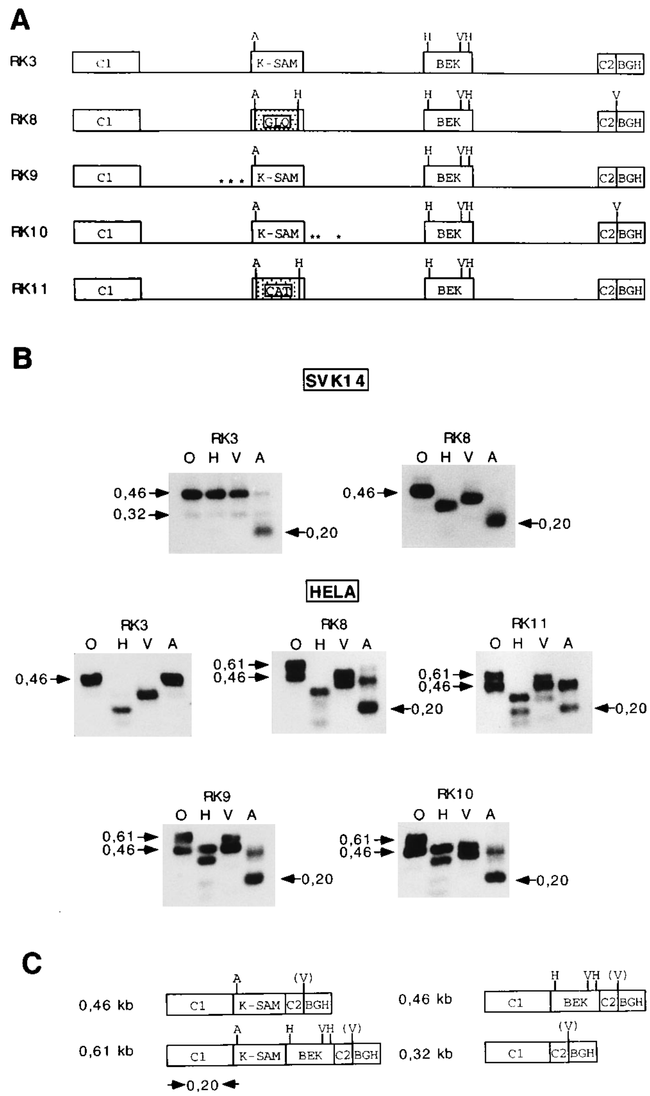

FIG. 3. Effects of K-SAM exon sequences on splicing of K-SAM exon in SVK14 and HeLa cells. (A) Schematic representations of the mutations in K-SAM exon and flanking intron sequences in different minigenes. A, *Ava*I; H, *Hinc*II; V, *Eco*RV. GLO in RK8 represents a globin exon sequence introduced between the *Ava*I and *Sal*I sites of pRK4. For RK9, the asterisks represent three pyrimidines which replace three purines in the K-SAM polypyrimidine sequence. In RK10, the 5′ splice site has been changed (three asterisks) from 5′-AAGg taacaatgctt-3′ to 5′-AAGgtaagtaagctt-3′ (exon sequence in uppercase letters, intron sequence in lowercase letters). Note that this mutation creates a *Hin*dIII site (aagctt) just downstream of the K-SAM exon. CAT in RK11 represents a CAT sequence introduced between the *Ava*I and *Sal*I sites of pRK4. (B) RNA from minigenes stably transfected into SVK14 or HeLa cells was subjected to RT-PCR as described in the text. PCR products were left undigested (O) or were digested with *Hinc*II (H), *Eco*RV (V), or *Ava*I (A) before Southern blotting. Autoradiograms of blots obtained are shown. The sizes of certain fragments are given in kb (see Table 1). The 0.20-kb *Ava*I fragment characteristic of K-SAM exon use is identified to the right of each autoradiogram. (C) Structures of PCR products. The *Eco*RV site (V) after the C2 sequence is present only in RK8- and RK10-derived PCR products. The origin of the 0.20-kb *Ava*I fragment is indicated.

transfected with RK8 are shown in Fig. 3B. Once again, the K-SAM exon was efficiently spliced (Fig. 3C, 0.46-kb PCR products with an *Ava*I site). Taken together, the above experiments suggest that no internal K-SAM exon sequence is required absolutely for splicing in SVK14 cells.

A K-SAM exon sequence represses splicing. We then stably transfected the RK3 and RK8 minigenes into HeLa cells. Results of the RT-PCR analysis carried out on RNA from the resulting cells are shown in Fig. 3B. With the RK3 minigene, 0.46-kb PCR products which contain *Hinc*II and *Eco*RV sites, but not *Ava*I sites, predominate and represent BEK transcripts (Fig. 3C). PCR products of 0.32 kb, representing C1C2 transcripts, were also detected at a low level (less than 5% of the hybridization signal). Different results were obtained for the RK8 minigene (Fig. 3B), in which globin sequences replace the K-SAM exon internal sequences (Fig. 3A). PCR products of 0.61 and 0.56 kb were detected and together made up about 60% of the hybridization signal, the remainder of the signal being made up by 0.46-kb fragments. In marked contrast to the results obtained with RK3, most of these PCR products contained an *Ava*I site. The results of digestion with *Ava*I, *Eco*RV, and *Hinc*II suggested that the 0.61- and 0.56-kb products derived from the two SAMBEK transcripts (Fig. 1B and 3C). This was confirmed by sequencing of subcloned PCR products (data not shown). Purification of the 0.46-kb PCR product and digestion with *Ava*I or *Hpa*I showed that it derived mainly from K-SAM transcripts, with only a minor contribution from BEK transcripts (data not shown).

Replacement of internal K-SAM exon sequences by globin exon sequences thus activates splicing of the K-SAM exon in HeLa cells. Indeed, this replacement has an effect very similar to that seen with RK3 minigenes in which the K-SAM exon polypyrimidine sequence (RK9) or 5′ splice site (RK10) has been mutated to make these sequences conform more closely to the corresponding consensus sequences (Fig. 3A). Both of these mutations activate K-SAM exon splicing in HeLa cells and lead to efficient production of SAMBEK and SAM transcripts. This is seen most clearly by comparing lanes O and A of Fig. 3B for the minigenes RK8, RK9, and RK10 and contrasting the corresponding lanes for RK3. Comparison of the

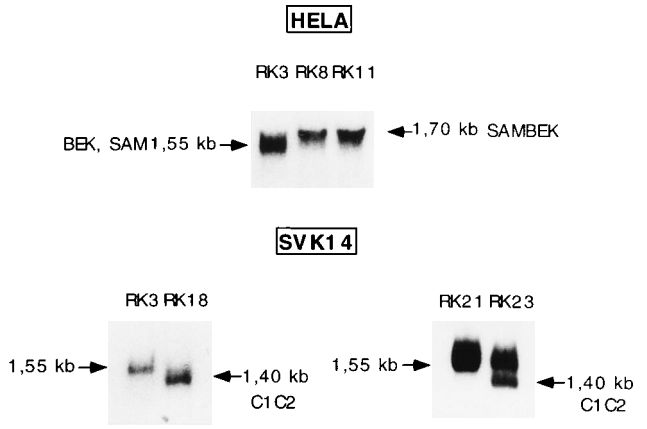

FIG. 4. Northern blot analysis of effects of minigene mutations on splicing of K-SAM exon. For HeLa cells, analysis of RNA from cells transiently transfected with RK3, RK8, or RK11 is shown. The 1.70-kb mRNA is the SAMBEK transcript (Fig. 1B), while the 1.55 kb mRNAs are mixtures of SAM and BEK mRNAs. For SVK14 cells, analysis of RNA from cells stably transfected with RK3, RK18, RK21 or RK23 is shown. The 1.40-kb mRNA is the C1C2 transcript (Fig. 1B), while the 1.55-kb mRNAs are SAM or BEK mRNAs.

other lanes is complicated by the presence of variable numbers of *Hin*cII and *Eco*RV sites in the different minigenes (Fig. 3A).

Is there a splicing enhancer in the globin sequence or a splicing repressor in the K-SAM sequence? To distinguish between these two possibilities, we constructed a further derivative of RK4, RK11 (Fig. 3A), in which the *Ava*I-*Sal*I sequence was replaced by a sequence of the same size from the protein coding region of a bacterial CAT gene. The reading frames of the K-SAM exon and the CAT sequences were conserved. The RT-PCR analysis carried out with RNA from HeLa cells stably transfected with RK11 yielded results similar to those obtained with RK8 (Fig. 3B): the 0.61- and 0.56-kb PCR products corresponding to SAMBEK transcripts were readily detectable. For both RK8 and RK11, results for both transient transfections and stable transfections in HeLa cells were identical (data not shown). Northern blotting of RNA from cells transiently transfected with RK3, RK8, and RK11 confirmed the RT-PCR analysis: a majority of the RK8- and RK11-derived transcripts are larger than the RK3-derived transcripts (Fig. 4). It seemed unlikely to us that both a globin gene exon and a bacterial gene would contain a splicing enhancer, and we concluded that the K-SAM exon probably contains a sequence needed for inhibition of its splicing in HeLa cells.

The K-SAM exon sequence 5′-TAGGGCAGGC-3′ inhibits splicing. We set out to localize the K-SAM exon inhibitory sequence. A derivative of RK11 (RK12) was made in which an *Eco*RV site was introduced into the CAT sequences (Fig. 5A). In RK12, the CAT sequences are divided into three segments by *Ava*I, *Eco*RI, *Eco*RV, and *Sal*I sites. Each of these segments was individually replaced with corresponding segments of the K-SAM exon (Fig. 2) to generate minigenes RK13, RK14, and RK15 (Fig. 5A). These three minigenes were transiently transfected into HeLa cells, and RT-PCR analysis was carried out on RNA from the transfected cells. PCR products were digested by *Ava*I (specific for the K-SAM exon) or by *Hpa*I (specific for the BEK exon; the *Hpa*I site corresponds to the upstream *Hin*cII site [Fig. 1A]). As shown in Fig. 5B, results obtained with RK13 or RK14 were similar to those obtained with RK11 (Fig. 3B). PCR products of 0.61 and 0.56 kb made up 55% of the hybridization signal, while products of 0.46 kb

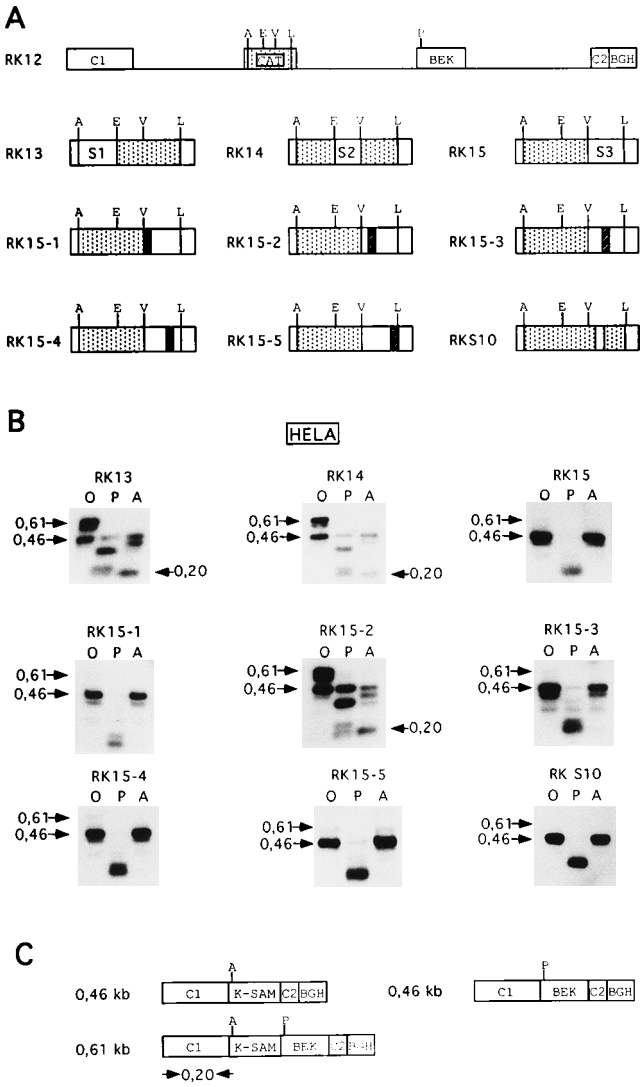

FIG. 5. The K-SAM exon inhibitory element is localized to a short exon sequence. (A) Schematic representations of mutated K-SAM exons in different minigenes. A, *Ava*I; E, *Eco*RI; L, *Sal*I; P, *Hpa*I; V, *Eco*RV. In RK12, K-SAM internal exon sequences have been replaced with CAT sequences (identified by vertical dotted lines). RK13, RK14, and RK15 were made from RK12 by replacing parts of these CAT sequences with the K-SAM sequences S1, S2, and S3, which correspond to segments 1, 2, and 3, respectively, as shown in Fig. 2. RK15-1 to -5 are versions of RK15 in which different blocks of segment 3 have been mutated (black boxes with diagonal white lines) as indicated in Fig. 2. RKS10 is identical to RK12 except that it contains the sequence 5′-TAGG GCAGGC-3′ from the K-SAM exon within the CAT sequences. (B) RNA from minigenes transiently transfected into HeLa cells was subjected to RT-PCR as described in the text. Abbreviations are as in the legend to Fig. 3B, except that P stands for *Hpa*I. (C) Structures of PCR products. The origin of the 0.20-kb *Ava*I fragment is indicated.

made up 45% of the signal. The 0.61- and 0.56-kb fragments were digested by both *Ava*I and *Hpa*I. Sequencing of subcloned 0.61- and 0.56-kb fragments (data not shown) confirmed that they represented SAMBEK transcripts (Fig. 1B and 5C). Different results were obtained with RK15: no 0.61- or 0.56-kb PCR products were detected, while the 0.46-kb products that were seen behaved as if they derived from BEK transcripts (digestion by *Hpa*I but not by *Ava*I). We conclude that segment 3 of the K-SAM exon sequence (Fig. 2) contains most of the activity responsible for repressing K-SAM exon splicing in HeLa cells.

The systematic mutation of blocks of nucleotides (Fig. 2) in segment 3 of the K-SAM exon sequence in RK15 was undertaken to further localize the repressing activity. Mutated minigenes RK15-1 to -5 (Fig. 5A) were transiently transfected into HeLa cells, RNA was harvested, and RT-PCR analysis was carried out. Minigenes in which block 1, 3, 4, or 5 (Fig. 2) had been mutated behaved like RK15 (Fig. 5B, RK15-1, -3, -4, and -5): few or no PCR products of 0.61 and 0.56 kb representing SAMBEK transcripts (Fig. 5C) were detected, while PCR products of 0.46 kb represented mainly BEK transcripts. However, the minigene in which block 2 was mutated (RK15-2, mutation of the sequence 5′-GGGCAGG-3′) behaved like RK13 and RK14: PCR products of 0.61 and 0.56 kb, representing SAMBEK transcripts, made up 60% of the hybridization signal. We conclude that the sequence 5′-GGGCAGG-3′ plays an important role in the repression of K-SAM exon splicing in HeLa cells. None of the mutations introduced into minigenes of the RK15 series had any detectable effect on splicing of the K-SAM exon in SVK14 cells (data not shown).

None of the block mutations described above mutate the TA base pairs immediately upstream of the 5′-GGGCAGG-3′ sequence (Fig. 2) or the C immediately downstream from it. We conclude that part or all of the sequence 5′-TAGGGCAGG C-3′ appears to be the only element indispensable for inhibition of splicing of the K-SAM exon in HeLa cells. To verify this, we constructed pRKS10 (Fig. 5A), which contains the sequence 5′-TAGGGCAGGC-3′ positioned relative to splice sites as it is normally positioned in the K-SAM exon but in which the other internal exon sequences are CAT sequences. pRKS10 was transiently transfected into HeLa cells, RNA was harvested, and RT-PCR analysis was carried out. PCR products of 0.46 kb (Fig. 5B), representing BEK transcripts (Fig. 5C), were obtained. The sequence 5′-TAGGGCAGGC-3′ is thus the only internal exon sequence necessary to repress K-SAM exon splicing.

Flanking intron sequences activate splicing. Having failed to find a *cis*-acting activating sequence in the K-SAM exon, we began to search for one in intron sequences. There are a 1.15-kb intron between the C1 and K-SAM exons (intron 1) and a 1.22-kb intron between the K-SAM and BEK exons (intron 2). We prepared minigenes (Fig. 6A) lacking bp 24 to 686 (RK16) or bp 690 to 1105 (RK17) of intron 1 or lacking bp 10 to 213 of intron 2 (RK18). RT-PCR analysis was carried out on RNA from SVK14 cells stably transfected with these minigenes. Deletion of bp 24 to 686 of intron 1 did not block splicing of the K-SAM exon (Fig. 6B, RK16; compare with Fig. 3B, RK3). PCR products of 0.46 kb with an *Ava*I site, representing SAM transcripts (Fig. 6C), predominated. Only very small amounts (less than 5% of the hybridization signal) of 0.32-kb PCR products, representing C1C2 transcripts, were detected. When bp 690 to 1105 of intron 1 were deleted (RK17), these 0.32-kb PCR products increased slightly in abundance (to reach 8% of the hybridization signal) relative to the 0.46-kb products, suggesting that this deletion removes some sequences which play minor roles in splicing of the K-SAM exon. This conclusion was confirmed by Northern blotting of RNA from cells transfected with RK17: a small amount of the 1.40-kb C1C2 transcript was detected (data not shown).

In marked contrast, deletion of bp 10 to 213 of intron 2 (Fig. 6A, RK18) blocked splicing of the K-SAM exon completely (Fig. 6B). The major 0.32-kb PCR product corresponded to C1C2 transcripts (Fig. 6C) (this was confirmed by sequencing of subcloned 0.32-kb fragments), while the minor 0.46-kb PCR product contained sites for *Hin*cII and *Eco*RV but not *Ava*I, and thus corresponded to BEK transcripts. No 0.46-kb PCR products containing an *Ava*I site were detected. Similar results

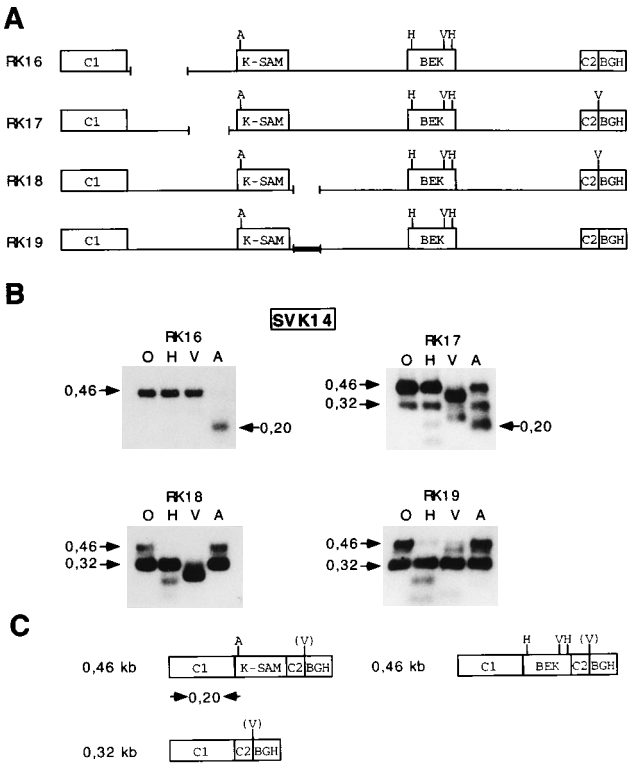

FIG. 6. An intron sequence is required for K-SAM exon splicing. (A) Schematic representations of mutated introns in different minigenes. A, *Ava*I; H, *Hin*cII; V, *Eco*RV. For RK16 to RK18, some intron sequences have been deleted. The boundaries of the intron deletions are described in the text. In RK19, the orientation of an intron sequence has been reversed (black box with white dots). (B) RNA from minigenes stably transfected into SVK14 cells was subjected to RT-PCR as described in the text. Abbreviations are as in the legend to Fig. 3B. (C) Structures of PCR products. The origin of the 0.20-kb *Ava*I fragment is indicated. The *Eco*RV site (V) after the C2 exon sequence exists only in RK17- and RK18-derived PCR products.

were obtained when bp 10 to 213 of intron 2 were inverted (Fig. 6A, RK19). The removal of an important sequence rather than the shortening of intron 2 is thus responsible for blocking K-SAM exon splicing. This effect is specific for the K-SAM exon, as none of the mutations present in minigenes RK16 to RK19 had any marked effect on splicing in HeLa cells, in which the BEK exon is spliced correctly (data not shown). Northern blotting of RNA from SVK14 cells stably transfected with RK3 or RK18 (Fig. 4) confirmed the PCR results, which show that the deletion of bp 10 to 213 of intron 2 leads to skipping of the K-SAM exon. The major FGFR-2 RNA species detected with RK18 RNA is 0.15-kb shorter than the RK3 RNA.

The intron 2 sequence up to bp 219 (Fig. 7) can be divided into three blocks: a zone rich in pyrimidines (from the 5′ splice site to a *Dra*I site), a zone rich in purines (from the *Dra*I site to a *Pst*I site), and downstream sequences (from the *Pst*I site to the *Xba*I site). Minigenes were made in which each of these blocks was mutated (Fig. 8A). In minigenes RK20 and RK21, the pyrimidine- and purine-rich sequences were replaced with random sequences. In minigene RK22, several pyrimidines in the pyrimidine-rich sequence were mutated to purines. In minigene RK23, the *Pst*I-*Xba*I fragment was replaced with an intron sequence of the same size from another gene. RT-PCR analysis was carried out on RNA from SVK14 cells stably transfected with these minigenes.

Randomization of the purine-rich sequence has no detect-

```
              Pyrimidines
         ---------------------
         TCTAGA              DraI
gtaacaatgctttcatttttgtctttttttaaaaagaaagctggatatagaagctgaaaaga     62
           A     A     A  A  A---------------------------------
                                           Purines

                  PstI
cttggtgctttgggagactgcaggcagcttataggataactcttgtggccttggtatattta     124

                                            NcoI
taataatctttcttcggtgatgcagctggtatgatgccagtagccatggaaaaatgcccaca     186

                            XbaI
acgttcaaagtgcttgctccaatttcttctaga                                  219
```

FIG. 7. Sequence of the first 219 bp of the intron separating the K-SAM and BEK exons. The *Xba*I site shown corresponds to that shown in Fig. 1. A pyrimidine-rich sequence subjected to mutation is shown (indicated with dashes and the word "Pyrimidines" above). This sequence is replaced by a random sequence in RK20; in RK22, indicated t bases have been mutated to A. In RK18, bp 10 to 213 have been deleted. This brings the *Xba*I site up to bp 10 (represented by TCTAGA). A purine-rich sequence subjected to mutation in RK21 is shown (indicated with dashes and the word "Purines" below).

able effect on splicing of the K-SAM exon (Fig. 8B, RK21). PCR products of 0.46 kb, representing SAM transcripts (Fig. 8C), predominate, and only minor amounts (less than 5% of the hybridization signal) of the 0.32-kb products representing C1C2 transcripts are detected. However, when the pyrimidine-rich sequence was replaced with a random sequence, splicing of the K-SAM exon became less efficient (Fig. 8B, RK20). The 0.32-kb PCR products corresponding to the C1C2 transcript were now more abundant (58% of the hybridization signal) than the 0.46 kb products (42% of the signal). When some pyrimidines of the pyrimidine-rich sequence were replaced by purines (RK22), or when the *Pst*I-*Xba*I fragment was replaced by a fragment of intron sequence from another gene (RK23), a similar though less marked effect was seen (Fig. 8B). Northern blotting of RNA from SVK14 cells stably transfected with RK21 or RK23 confirmed the PCR results (Fig. 4). The RK21 transcripts corresponded in size to SAM transcripts (1.55 kb), while the RK23 transcripts were composed of a mixture of 1.55- and 1.40-kb transcripts, the latter corresponding to C1C2 transcripts. None of the mutations present in minigenes RK20 to RK23 had any effect on splicing in HeLa cells, with the BEK exon continuing to be spliced correctly (data not shown).

Two separate sequence elements in the intron downstream of the K-SAM exon are thus needed for efficient exon splicing. One of these is a pyrimidine-rich sequence. We observed that deletion of the pyrimidine-rich sequence alone (RK24) blocked splicing of the K-SAM exon completely (data not shown). The explanation for this is probably that this deletion has two effects: loss of an important sequence element and repositioning of a second, downstream element such that it no longer functions correctly.

The intron activating sequence is needed only in the presence of the exon repressing sequence. If the intron sequences identified above really function to activate splicing of the K-SAM exon, we would expect that they would no longer be needed if K-SAM exon splicing was activated in other ways. We have shown above (Fig. 3) and elsewhere (12) that splicing of the K-SAM exon can be activated in HeLa cells by mutations which reinforce the polypyrimidine sequence or the 5′ splice site sequence of the K-SAM exon. The minigenes carrying these mutations (RK9 and RK10, respectively) (Fig. 3A) were made into minigenes RK25 and RK26, respectively (Fig. 9A), by the deletion of bp 10 to 213 of intron 2. These minigenes were stably transfected into SVK14 cells, and RNA was

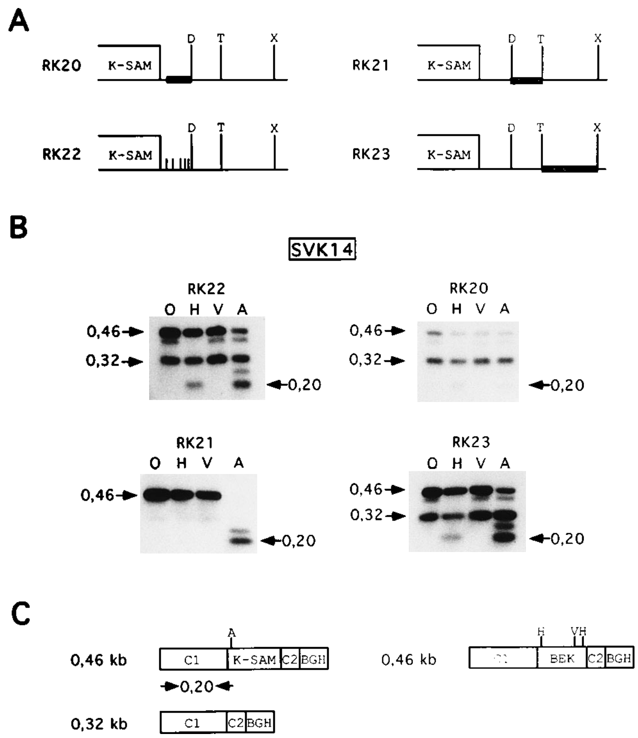

FIG. 8. The K-SAM intron activating sequence is bipartite. (A) Schematic representations of mutated introns in different minigenes. D, *Dra*I; T, *Pst*I; X, *Xba*I. The block mutations in RK20, RK21, and RK23 are shown as black boxes with diagonal white lines. In RK20, a pyrimidine rich sequence has been mutated. In RK21 a purine-rich sequence has been mutated. In RK22, individual pyrimidines in the pyrimidine-rich sequence have been replaced by purines (short vertical lines). Further details are given in the text and in the legend to Fig. 7. (B) RNA from minigenes stably transfected into SVK14 cells was subjected to RT-PCR as described in the text. Abbreviations are as in the legend to Fig. 3B. (C) Structures of PCR products. The origin of the 0.20-kb *Ava*I fragment is indicated.

harvested. The results of RT-PCR (Fig. 9B) demonstrate that, in marked contrast to RK18 (Fig. 6A), whose K-SAM exon carries the natural splice sites, for RK25 and RK26 the efficient splicing of the K-SAM exon is no longer dependent on the intron 2 sequence: 0.46-kb PCR products with an *Ava*I site (Fig. 9C) predominate. The intron 2 sequence thus functions as a true activating sequence.

Finally, we asked whether the intron 2 activating sequence is still needed for K-SAM exon splicing in SVK14 cells when the K-SAM internal exon sequence which represses K-SAM exon splicing in HeLa cells has been removed. To answer this question we prepared RK3 derivatives in which this exon sequence had been replaced with globin gene exon sequences or with CAT sequences and which also had deletions of the intron 2 activating sequences (RK27 and RK28, respectively) (Fig. 9A). These minigenes were stably transfected into SVK14 cells. When RT-PCR analysis was carried out on RNA from the resulting cells, 0.46-kb PCR products with an *Ava*I site (Fig. 9B), representing SAM transcripts (Fig. 9C), predominated. These results demonstrate that if the intron 2 activating sequence is normally needed for K-SAM exon splicing in SVK14 cells (Fig. 6A, RK18), it is no longer needed if the K-SAM internal exon sequences are replaced with other sequences.

The intron activating sequences and the exon inhibitory sequence are active in both SVK14 and HeLa cells. When the minigene RK18, which contains the normal K-SAM exon but

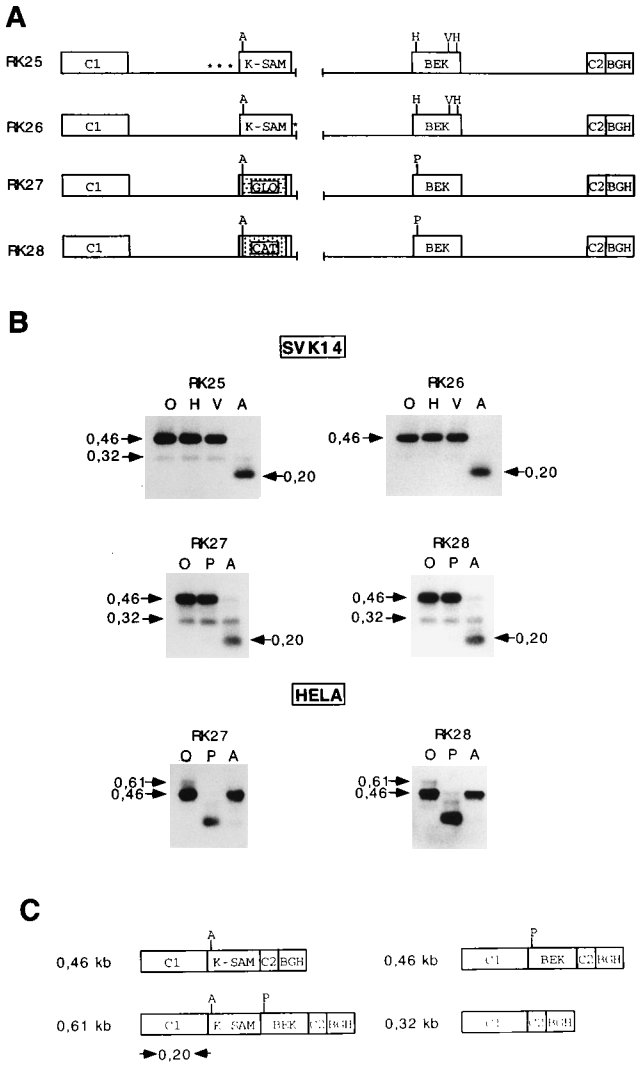

FIG. 9. Reinforcing the K-SAM exon splicing signals removes the intron sequence requirement for splicing in SVK14 cells. (A) Schematic representation of minigenes used. A, *Ava*I; H, *Hin*cII; V, *Eco*RV. For RK25, the asterisks represent three pyrimidines which replace three purines in the K-SAM exon polypyrimidine sequence. In RK26, the 5′ splice site has been changed from 5′-AAGgtaacaatgctt-3′ to 5′-AAGgtaagtaagctt-3′, and this mutation is indicated by an asterisk (exon sequence in uppercase letters, intron sequence in lowercase letters). In RK27 and RK28, K-SAM internal exon sequences containing the exon inhibitory sequence have been replaced by globin and CAT sequences, respectively. For all the minigenes, the K-SAM intron activating sequences have been deleted. (B) RNA from minigenes stably transfected into SVK14 cells or transiently transfected into HeLa cells was subjected to RT-PCR as described in the text. Abbreviations are as in the legend to Fig. 3B. (C) Structures of PCR products. The origin of the 0.20-kb *Ava*I fragment is indicated.

not the intron activating sequence (Fig. 6A), is transfected into SVK14 cells, the K-SAM exon is not spliced (Fig. 6B). The results obtained with RK27 and RK28 (Fig. 9B) show that when the K-SAM exon internal sequences containing the inhibitory sequence are removed from RK18, the K-SAM exon is spliced. This demonstrates that exon inhibitory sequences function in SVK14 cells as well as in HeLa cells.

Is the same true for the intron activating sequences? Minigenes RK27 and RK28, which contain neither the exon inhibitory sequence nor the intron activating sequences (Fig. 9A), were transiently transfected into HeLa cells, and RT-PCR analysis was carried out on RNA from the transfected cells. PCR products were digested by *Ava*I (specific for the K-SAM exon) or by *Hpa*I (specific for the BEK exon). As shown in Fig. 9B, 0.46-kb PCR products with an *Hpa*I site, representing BEK transcripts (Fig. 9C), predominated. The minigenes RK8 and RK11 (Fig. 3), which do not contain the exon inhibitory sequence but do contain the intron activating sequences, use the K-SAM exon efficiently in HeLa cells (and splice it to the BEK exon). We thus conclude that the deletion of the intron activating sequence reduces the efficiency of K-SAM exon splicing and therefore that this sequence is active in HeLa cells as well as in SVK14 cells.

Note that we cannot compare the behaviors of minigenes RK27 and RK28 in the two different cell lines used for the transfections. In HeLa cells, the BEK exon is available for splicing, and the K-SAM exon is in competition with the BEK exon for splicing to the flanking C1 and C2 exons. In SVK14 cells, the BEK exon is not available for splicing and this competition does not take place. The skipping of the K-SAM exon of RK27 and RK28 in HeLa cells but not in SVK14 cells may reflect this difference.

DISCUSSION

We show in this report that multiple sequence elements influence splicing of the K-SAM exon. We have identified a short K-SAM exon sequence (5′-TAGGGCAGGC-3′) which represses splicing of this exon and two downstream intron sequences which activate splicing of the exon. These sequences act in conjunction with the K-SAM exon's suboptimal 5′ and 3′ splice site sequences. Thus, when the K-SAM exon's polypyrimidine sequence or 5′ splice site is reinforced, the exon inhibitory element can no longer repress splicing in HeLa cells, while the intron activating sequence is no longer needed for splicing in SVK14 cells. Furthermore, in the absence of the exon inhibitory sequence the intron activating sequence is not necessary for splicing of the K-SAM exon in SVK14 cells (which normally splice this exon). Positive and negative signals are thus involved in the splicing of the K-SAM exon.

The description of a short, defined exon inhibitory sequence seems particularly interesting to us, as while much recent work has described purine-rich sequences which act as splicing enhancers in exons (8, 25, 40, 41, 45, 47), relatively few examples of splicing-inhibitory sequences are known. Perhaps the best-characterized example of the latter is an exon inhibitory sequence close to the 3′ splice site of a β-tropomyosin alternative exon, which functions by forming a secondary structure that "hides" an important splicing signal at the 3′ splice site (10, 26). For K-SAM, we have been unable to detect any obvious similar secondary structure which includes the exon inhibitory sequence. Furthermore, the K-SAM exon inhibitory sequence is more than 90 bp downstream from the 3′ splice site. The formation of secondary structures with other K-SAM exon internal sequences is not necessary either, as the inhibitory sequence alone is sufficient to block splicing of a K-SAM exon composed of bacterial CAT gene sequences.

Other examples of exon sequences for which mutation derepresses splicing of the exon are known. These include a short sequence (5′-CAAGG-3′) in the human fibronectin EDA exon (8) and a 28-bp sequence in a human immunodeficiency virus type 1 *tat* exon (1). The human α-tropomyosin gene SK exon may also contain an inhibitory element (16). There is, however, no obvious similarity between these sequences (the important parts of which have often not been defined) and the K-SAM exon inhibitory sequence. The mode of action of these sequences remains, furthermore, unknown.

The third intron of the *Drosophila* P-element transposase

gene is spliced efficiently in germ line cells but not in the soma (38). This control involves binding of multiprotein complexes, including the hnRNP A1-like protein hrp48 (which may in fact be the *Drosophila* homolog of mammalian hnRNP A1), to specific nucleotides in a pseudo-5′ splice site within an exon inhibitory element (39). U1 snRNP subsequently binds to the pseudo-5′ splice site and no longer binds to the true 5′ splice site. Implication of binding of an hnRNP A1-like protein to specific exon sequences in the repression of splicing is particularly interesting, as hnRNP A1 and SR proteins such as ASF/SF2 are known to be antagonistic splicing factors (7, 11, 24, 27, 49). Thus, for example, while ASF/SF2 stimulates splicing of a bovine growth hormone exon by binding to a purine-rich exon sequence, hnRNP A1 counteracts this stimulation (40). It appears likely that some exon inhibitory elements will block splicing by binding hnRNP A1 or similar proteins. They would then represent formal counterparts to the exon splicing enhancer sequences which have been described, at least some of which function by binding SR proteins (25, 40).

We have localized the K-SAM exon inhibitory element to part or all of the sequence 5′-TAGGGCAGGC-3′. This sequence is similar to a 3′ splice site (two copies of the motif YAGG) and could act as a decoy, like the P-element pseudo-5′ splice site. On the other hand, the K-SAM exon sequence motif is also related to the consensus sequence proposed for high-affinity hnRNP A1 binding sites (5′-UAGGGA-3′ and 5′-UAGGGU-3′), as defined by the systematic evolution of ligands by exponential enrichment (SELEX) approach (6). It is thus tempting to speculate that binding of hnRNP A1 or a related protein to the K-SAM exon inhibitory element is involved in inhibiting splicing of the K-SAM exon. This could in turn hinder communication between the exon's 5′ and 3′ splice sites and block splicing.

We have also shown that sequences required for K-SAM exon splicing exist in the intron downstream of this exon. The deletion of sequences in the downstream intron blocks K-SAM exon splicing completely. There appear to be at least two elements involved in this intron. One is a pyrimidine-rich sequence lying immediately adjacent to the 5′ splice site. This is separated from a second, downstream element by a purine-rich sequence. The purine-rich sequence does not seem to be involved in splicing control. The two elements may have partially overlapping functions, as, while mutation of either element alone reduces the efficiency of K-SAM exon splicing, mutation of both elements is necessary to block K-SAM exon splicing completely.

These intron sequences could in principle function by causing RNA polymerase II to pause during transcription, allowing time for the K-SAM exon's 3′ splice site to be spliced to the C1 exon's 5′ splice site before transcription of the BEK exon takes place. We believe this to be unlikely, as it would imply that the upstream intron is always removed before the downstream intron. Using a PCR-based technique developed to study the order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNAs (22), we have found no evidence for a defined order of intron removal for the K-SAM introns (our unpublished data). We prefer to hypothesize that the intron sequences represent binding sites for proteins which interact with spliceosome components. Several proteins have been described which bind to pyrimidine-rich RNA sequences (for a review, see reference 29), and any one of these could be implicated in activation of splicing, although their mechanisms of action remain to be defined.

Several other examples of intron sequences which control splicing are known. Thus several α- and β-tropomyosin genes contain intron sequences exerting a negative effect on splicing. For example, the rat α-tropomyosin gene contains two alternative exons (2 and 3) which are spliced in a strict mutually exclusive manner (14). Two conserved sequence elements, one in each of the introns flanking exon 3, are responsible for the repression of splicing of this exon in smooth muscle cells (14). Similarly, sex-lethal protein needs to bind to multiple *cis*-acting sequences both upstream and downstream of the male exon to autoregulate splicing of its pre-mRNA (18). A negative regulator of splicing within the *gag* gene of Rous sarcoma virus binds snRNPs U11 and U12 and spliceosomal snRNPs and controls the relative levels of spliced and unspliced viral mRNAs (13).

On the other hand, several activating sequences in introns have been described. A pyrimidine-rich sequence just downstream of one of a pair of chicken β-tropomyosin alternative exons is involved in activation of splicing of this exon (4). This sequence may play a role for the tropomyosin gene similar to that played by the K-SAM intron pyrimidine-rich sequence, although we have shown that it cannot replace it functionally, at least not in the minigene that we have used here (our unpublished observations). Another example of intron activating sequences is furnished by the neural-specific splicing of the c-*src* gene N1 exon, which requires at least two elements in the downstream intron (5). Splicing of exon IIIB of the rat fibronectin gene depends on repeats of the motif 5′-UGC AUG-3′ in the downstream intron (20). Interestingly, this motif can be found also in one of the c-*src* gene intron elements, and its mutation reduces inclusion of the N1 exon (5, 20). Once again, the mode of action of this sequence is not known, although it could clearly represent a binding site for a protein involved in splicing activation. However, the intron elements required for K-SAM exon splicing contain neither this sequence nor any other similar repeated sequence.

Are the intron activating sequences and the exon inhibitory sequence involved in the tissue-specific splicing of the K-SAM exon? It is tempting to speculate that this is the case, because tampering with these sequences leads to inappropriate splicing choices. It may be that the combination of weak splice sites, an exon inhibitory sequence, and intron activating sequences is needed to allow tissue-specific signals received elsewhere on the pre-mRNA to exert their effects on splicing of the K-SAM exon. It is also possible that tissue-specific differences in the levels of *trans*-acting factors which bind to the inhibitory and activating sequences control splicing of the K-SAM exon. Insofar as both these sequences are active in HeLa cells (which normally splice the BEK exon) and in SVK14 cells (which normally splice the K-SAM exon), we can rule out the extreme model for K-SAM exon splicing, which invokes the binding of a repressor active only in HeLa cells to the exon inhibitory sequence and the binding of an activator active only in SVK14 cells to the intron sequences. It remains possible that K-SAM exon splicing is controlled by the relative levels of these putative *trans*-acting factors. HeLa cells might express relatively higher levels of the repressor, and SVK14 cells might express higher levels of the activator, for example. Now that pre-mRNA target sequences have been identified, it will be possible to search for such regulatory proteins.

ACKNOWLEDGMENTS

We thank Marie-Claude Gesnel for excellent technical assistance.

This work was supported by grants from the Association pour la Recherche sur le Cancer and the Ligue Nationale contre le Cancer, Comité Départemental de Loire-Atlantique.

REFERENCES

1. **Amendt, B. A., D. Hesslein, L.-J. Chang, and C. M. Stoltzfus.** 1994. Presence of negative and positive *cis*-acting RNA splicing elements within and flanking the first *tat* coding exon of human immunodeficiency virus type 1. Mol. Cell. Biol. **14:**3960–3970.
2. **Amrein, H., M. L. Hedley, and T. Maniatis.** 1994. The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by transformer 2. Cell **76:**735–746.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1991. Current protocols in molecular biology. John Wiley and Sons, New York.
4. **Balvay, L., D. Libri, M. Gallego, and M. Fiszman.** 1992. Intronic sequences with both positive and negative effects on the regulation of alternative splicing of the chicken beta-tropomyosin transcripts. Nucleic Acids Res. **20:**3987–3992.
5. **Black, D. L.** 1992. Activation of c-src neuron-specific splicing by an unusual RNA element in vivo and in vitro. Cell **69:**795–807.
6. **Burd, C. G., and G. Dreyfuss.** 1994. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. EMBO J. **13:**1197–1204.
7. **Caceres, J. F., S. Stamm, D. M. Helfman, and A. R. Krainer.** 1994. Regulation of alternative splicing by overexpression of antagonistic splicing factors. Science **265:**1706–1709.
8. **Caputi, M., G. Casari, S. Guenzi, R. Tagliabue, A. Sidoli, C. A. Melo, and F. E. Baralle.** 1994. A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. Nucleic Acids Res. **22:**1018–1022.
9. **Champion-Arnaud, P., C. Ronsin, E. Gilbert, M. C. Gesnel, E. Houssaint, and R. Breathnach.** 1991. Multiple mRNAs code for proteins related to the BEK fibroblast growth factor receptor. Oncogene **6:**979–987.
10. **Clouet d'Orval, B., Y. d'Aubenton Carafa, P. Sirand-Pugnet, M. Gallego, E. Brody, and J. Marie.** 1991. RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. Science **252:**1823–1828.
11. **Ge, H., and J. L. Manley.** 1990. A protein factor, ASF, controls cell specific alternative splicing of SV40 early pre-mRNA in vitro. Cell **62:**25–34.
12. **Gilbert, E., F. Del Gatto, P. Champion-Arnaud, M.-C. Gesnel, and R. Breathnach.** 1993. Control of BEK and K-SAM splice sites in alternative splicing of the fibroblast growth factor receptor 2 pre-mRNA. Mol. Cell. Biol. **13:**5461–5468.
13. **Gontarek, R. R., M. T. McNally, and K. Beemon.** 1993. Mutation of an RSV intronic element abolishes both U11/U12 snRNP binding and negative regulation of splicing. Genes Dev. **7:**1926–1936.
14. **Gooding, C., G. C. Roberts, G. Moreau, B. Nadal-Ginard, and C. W. J. Smith.** 1994. Smooth muscle-specific switching of alpha-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon. EMBO J. **13:**3861–3872.
15. **Gough, N.** 1988. Rapid and quantitative preparation of cytoplasmic RNA from small numbers of cells. Anal. Biochem. **173:**93–95.
16. **Graham, I. R., M. Hamshere, and I. C. Eperon.** 1992. Alternative splicing of a human α-tropomyosin muscle-specific exon: identification of determining sequences. Mol. Cell. Biol. **12:**3872–3882.
17. **Hoffman, B. E., and P. J. Grabowski.** 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. Genes Dev. **6:**2554–2568.
18. **Horabin, J. I., and P. Schedl.** 1993. *Sex-lethal* autoregulation requires multiple *cis*-acting elements upstream and downstream of the male exon and appears to depend largely on controlling the use of the male exon 5′ splice site. Mol. Cell. Biol. **13:**7734–7746.
19. **Horowitz, D. S., and A. R. Krainer.** 1994. Mechanisms for selecting 5′ splice sites in mammalian pre-mRNA splicing. Trends Genet. **10:**100–106.
20. **Huh, G. S., and R. O. Hynes.** 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. Genes Dev. **8:**1561–1574.
21. **Johnson, D. E., and L. T. Williams.** 1993. Structural and functional diversity in the FGF receptor multigene family. Adv. Cancer Res. **60:**1–41.
22. **Kessler, O., Y. Chiang, and L. A. Chasin.** 1993. Order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNA. Mol. Cell. Biol. **13:**6211–6222.
23. **Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. Garcia-Blanco, and J. L. Manley.** 1994. Protein-protein interactions and 5′-splice site recognition in mammalian mRNA precursors. Nature (London) **368:**119–124.
24. **Krainer, A. R., G. C. Conway, and D. Kozak.** 1990. The essential pre-mRNA splicing factor SF2 influences 5′ splice site selection by activating proximal sites. Cell **62:**35–42.
25. **Lavigueur, A., H. La Branche, A. R. Kornblihtt, and B. Chabot.** 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes Dev. **7:**2405–2417.
26. **Libri, D., A. Piseri, and M. Y. Fiszman.** 1991. Tissue-specific splicing in vivo of the beta-tropomyosin gene: dependence on an RNA secondary structure. Science **252:**1842–1845.
27. **Mayeda, A., D. M. Helfman, and A. R. Krainer.** 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol. Cell. Biol. **13:**2993–3001.
28. **Mayeda, A., A. M. Zahler, A. R. Krainer, and M. B. Roth.** 1992. Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing. Proc. Natl. Acad. Sci. USA **89:**1301–1304.
29. **McKeown, M.** 1992. Alternative mRNA splicing. Annu. Rev. Cell Biol. **8:**133–155.
30. **Miki, T., D. P. Bottaro, T. P. Fleming, C. L. Smith, W. H. Burgess, A. M. L. Chan, and S. A. Aaronson.** 1992. Determination of ligand-binding specificity by alternative splicing: two distinct growth factor receptors encoded by a single gene. Proc. Natl. Acad. Sci. USA **89:**246–250.
31. **Orr-Urtreger, A., M. T. Bedford, T. Burakova, E. Arman, Y. Zimmer, A. Yayon, D. Givol, and P. Lonai.** 1993. Developmental localization of the splicing alternatives of fibroblast growth factor receptor-2 (FGFR-2). Dev. Biol. **158:**475–486.
32. **Rautmann, G., H. W. D. Matthes, M. J. Gait, and R. Breathnach.** 1984. Synthetic donor and acceptor splice sites function in an RNA polymerase B(II) transcription unit. EMBO J. **3:**2021–2028.
33. **Rio, D. C.** 1993. Splicing of pre-mRNA: mechanism, regulation and role in development. Curr. Opin. Genet. Dev. **3:**574–584.
34. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNA with multiple introns. Mol. Cell. Biol. **10:**84–94.
35. **Sato, M., T. Kitazawa, A. Katsumata, M. Mukamoto, T. Okada, and T. Takeya.** 1992. Tissue-specific expression of two isoforms of chicken fibroblast growth factor receptor, bek and Cek3. Cell Growth Differ. **3:**355–361.
36. **Scotet, E., E. Reichmann, R. Breathnach, and E. Houssaint.** 1995. Oncoprotein fos activation in epithelial cells induces an epithelio-mesenchymal conversion and changes the receptor encoded by the FGFR-2 mRNA from K-SAM to BEK. Oncol. Rep. **2:**203–207.
37. **Sharp, P. A.** 1994. Split genes and RNA splicing. Cell **77:**805–815.
38. **Siebel, C. W., L. D. Fresco, and D. C. Rio.** 1992. The mechanism of somatic inhibition of drosophila P-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5′ splice site control U1 snRNP binding. Genes Dev. **6:**1386–1401.
39. **Siebel, C. W., R. Kanaar, and D. C. Rio.** 1994. Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI. Genes Dev. **8:**1713–1725.
40. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. **7:**2598–2608.
41. **Tanaka, K., A. Watakabe, and Y. Shimura.** 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol. Cell. Biol. **14:**1347–1354.
42. **Tian, M., and T. Maniatis.** 1994. A splicing enhancer exhibits both constitutive and regulated activities. Genes Dev. **8:**1703–1712.
43. **Uetsuki, T., A. Naito, S. Nagata, and Y. Kaziro.** 1989. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha. J. Biol. Chem. **264:**5791–5798.
44. **Valcarel, J., S. Ravinder, P. D. Zamore, and M. R. Green.** 1993. The protein sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. Nature (London) **362:**171–175.
45. **Watakabe, A., K. Tanaka, and Y. Shimura.** 1993. The role of exon sequences in splice site selection. Genes Dev. **7:**407–418.
46. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75:**1061–1070.
47. **Xu, R., J. Teng, and T. A. Cooper.** 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. Mol. Cell. Biol. **13:**3660–3674.
48. **Yan, G., Y. Fukabori, G. McBride, S. Nikolaropolous, and W. L. McKeehan.** 1993. Exon switching and activation of stromal and embryonic fibroblast growth factor (FGF)-FGF receptor genes in prostate epithelial cells accompany stromal independence and malignancy. Mol. Cell. Biol. **13:**4513–4522.
49. **Yang, X., M. R. Bani, S. J. Lu, S. Rowan, Y. Ben-David, and B. Chabot.** 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5′ splice site selection in vivo. Proc. Natl. Acad. Sci. USA **91:**6924–6928.
50. **Yayon, A., Y. Zimmer, S. Guo-Hong, A. Avivi, Y. Yarden, and D. Givol.** 1992. A confined variable region confers ligand specificity on fibroblast growth factor receptors: implications for the origin of the immunoglobulin fold. EMBO J. **11:**1885–1890.
51. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: conserved family of pre-mRNA splicing factors. Genes Dev. **6:**837–847.
52. **Zamore, P. D., and M. R. Green.** 1991. Biochemical characterization of U2 snRNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intranuclear distribution. EMBO J. **10:**207–214.
53. **Zamore, P. D., J. Patton, and M. R. Green.** 1992. Cloning and domain structure of the mammalian splicing factor U2AF. Nature (London) **355:**609–614.