JOURNAL OF VIROLOGY, Dec. 1997, p. 9096–9107 Vol. 71, No. 12
0022-538X/97/$04.00+0

Structural, Functional, and Protein Binding Analyses of Bovine Papillomavirus Type 1 Exonic Splicing Enhancers

ZHI-MING ZHENG,* PEI-JUN HE,† AND CARL C. BAKER

Basic Research Laboratory, Division of Basic Sciences, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892-5055

Received 8 May 1997/Accepted 16 August 1997

Alternative splicing plays an important role in regulation of bovine papillomavirus type 1 (BPV-1) gene expression. We have recently identified in BPV-1 late pre-mRNAs two purine-rich exonic splicing enhancers (SE1 and SE2) which also stimulate splicing of a *Drosophila doublesex (dsx)* pre-mRNA containing a suboptimal 3′ splice site. In vivo studies now demonstrate that both SE1 and SE2 are required for preferential use of the BPV-1 nucleotide (nt) 3225 3′ splice site in nonpermissive cells. Deletion or mutation of either element in a BPV-1 late pre-mRNA switches splicing to the late-specific alternative 3′ splice site at nt 3605. To investigate the sequence specificity of these exonic splicing enhancers, various mutant SE1 or SE2 elements were connected to *dsx* pre-mRNAs and tested for their stimulatory effects on *dsx* pre-mRNA splicing in vitro. Substitution of U residues for either A or G residues in and around potential ASF/SF2 binding sites in SE1 or SE2 resulted in a significant reduction of splicing enhancer activity. However, the G-to-U substitutions in both enhancers had the largest effect, reducing splicing to near control levels. Further in vitro analyses showed that splicing enhancement by SE2 could be competed with excess unlabeled SE2 RNA, indicating that SE2 activity in HeLa nuclear extracts is mediated by *trans*-acting factors. UV cross-linking plus immunoprecipitation assays showed that both wild-type SE1 and SE2 RNAs could bind directly to purified HeLa SR proteins SRp30a (ASF/SF2), SRp55, and SRp75. UV cross-linking experiments also identified a 23-kDa protein which binds to SE2 but not SE1. This protein is present in both HeLa nuclear extracts and S100 extracts but absent from SR protein preparations, suggesting that it is not a classical SR protein. Mutant SE elements (containing G- to U-mutations) which had minimal splicing enhancer activity also had very weak binding capacity for these proteins, strongly suggesting that the binding of these proteins is required for splicing enhancer function.

Processing of a pre-mRNA requires precise intron excision and ligation of the exons in a process known as splicing (reviewed in reference 29). Four critical sequence elements, the 5′ splice site, the 3′ splice site, the branch point, and the polypyrimidine tract, are essential for constitutive splicing in mammalian systems. The machinery which carries out this process is known as the spliceosome and is composed of multiple snRNPs and splicing factors. The earliest stages in spliceosome assembly involve the formation of the E (early) or commitment complex in which the 5′ splice site is recognized by U1 snRNP and the 3′ splice site by binding of $U2AF^{65}$ at the polypyrimidine tract (20, 28, 37, 40, 55). Binding of U1 snRNP and $U2AF^{65}$ is facilitated by the arginine/serine-rich (SR) family of splicing factors which interact with both U1-70K and $U2AF^{35}$ (13, 22, 51, 57). The SR proteins also mediate splice site pairing through a network of protein-protein interactions which may bridge both the exon and the intron (6, 51). One recent study suggests that subsequent to commitment complex formation, the basic amino acid side chains of the RS domain of $U2AF^{65}$ promote the base pairing of U2 snRNP at the branch point by neutralization of the negatively charged phosphates of the RNA (47). Finally, addition of U4, U5, and U6 snRNPs to the spliceosome ultimately catalyzes the splicing reactions (reviewed in references 29, 32, and 39).

Alternative pre-mRNA splicing is an important mechanism for the regulation of gene expression. Alternatively spliced mRNAs can generate variant proteins from a single gene, and these proteins may have different functions or may be expressed in a developmental stage- or tissue-specific manner (19). The mechanisms by which alternative splice site selection is regulated in eukaryotic cells and viruses are now partially understood. Many studies have demonstrated that alternatively spliced genes usually have either a suboptimal 5′ splice site or a suboptimal 3′ splice site characterized by a nonconsensus branch point and/or a suboptimal polypyrimidine tract (41; reviewed in reference 4). Use of suboptimal splice sites can be facilitated by purine-rich exonic sequences known as exonic splicing enhancers (ESEs) (reviewed in references 4, 5, and 26). This facilitation is mediated by the SR proteins which bind to the ESE and then recruit U2AF to the suboptimal 3′ splice sites (23, 33, 40, 42, 43).

The SR proteins are a family of splicing factors characterized by N-terminal RNA binding domains and C-terminal RS domains. The RS domains can be phosphorylated in vitro on serine residues by cellular SRPK1 (17, 18) or Clk/Sty (8) protein kinase or by DNA topoisomerase I (34). The phosphorylated RS domain serine residues are part of the phosphoepitope recognized by monoclonal antibody (MAb) 104 (54). The N-terminal RNA recognition motifs (also known as RNA binding domains) cooperate to define the overall RNA binding specificity (7, 43; reviewed in references 11 and 26). The SR protein family now has at least nine members: SRp20 (X16 or RBP1), ASF/SF2 (SRp30a), SC35 (SRp30b or PR264), SRp30c, 9G8, SRp40 (HRS), p54, SRp55 (B52), and SRp75 (reviewed in references 11, 26, and 48). SR proteins are essential splicing factors that are required for constitutive splicing. They are also alternative splicing factors and can regulate use

* Corresponding author. Mailing address: Basic Research Laboratory, National Cancer Institute, Building 41, Room C111, 41 Library Dr., MSC 5055, Bethesda, MD 20892-5055. Phone: (301) 496-9489. Fax: (301) 402-0055. E-mail: Zhengt@dce41.nci.nih.gov.

† Present address: Department of Microbiology, University of Maryland, College Park, MD 20742-2021.

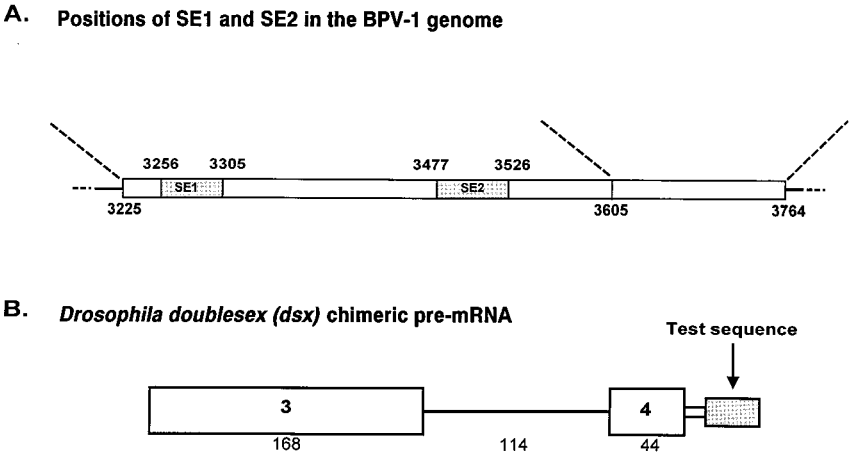

FIG. 1. Locations of BPV-1 SE1 and SE2 in BPV-1 genome and experimental strategy. (A) Positions of the SE sequences in BPV-1 are indicated by shaded boxes. Numbers are the nucleotide positions in the BPV-1 genome. The nt 3225 and 3605 3′ splice sites and the nt 3764 5′ splice site are indicated. (B) Schematic representation of *Drosophila dsx* chimeric pre-mRNA. Exon sequences (boxes) and intron sequences (lines) are shown. The lengths (in nucleotides) of the exons and introns are indicated below the corresponding regions of the construct. The 3′ exon contains 30 nt of the *dsx* fourth exon and 14 nt of polylinker sequence derived from pSP72. The shaded box represents exonic splicing enhancer sequences which are inserted into exon 4. The plasmid from which the chimeric *dsx* pre-mRNA is transcribed contains a *Bgl*II site upstream of the *dsx* sequences and *Xba*I (5′) and *Hin*dIII (3′) sites into which putative ESE sequences are cloned.

of alternative 5′ splice sites by binding to sequences at or near the 5′ splice site. As mentioned above, regulation of alternative 3′ splice site selection and internal exon inclusion is mediated by the binding of SR proteins to ESE elements.

Viruses frequently use alternative splicing as a way to pack maximum coding capacity into a small genome. The expression of the late genes of bovine papillomavirus type 1 (BPV-1) is regulated in part through alternative splicing, which itself is regulated as a function of differentiation of the host cell, the keratinocyte (2, 3). Early and intermediate stages of the viral life cycle take place in the basal and spinous layers of the epidermis, respectively. The majority of viral pre-mRNAs at these stages of the viral life cycle are spliced by utilizing a common 3′ splice site at nucleotide (nt) 3225. However, in the fully differentiated keratinocytes of the granular cell layer, late viral pre-mRNAs are also spliced by using a late-specific alternative 3′ splice site at nt 3605. This shift in alternative splicing is required to produce the L1 mRNA which codes for the major capsid protein. We have previously identified two ESEs (SE1 and SE2) between the nt 3225 and nt 3605 3′ splice sites in BPV-1 late pre-mRNAs and demonstrated that SE1 is required for efficient utilization of the suboptimal nt 3225 3′ splice site in vitro (Fig. 1) (56). This study also identified an exonic splicing suppressor (ESS) adjacent to SE1 that can repress utilization of this suboptimal 3′ splice site.

In this study, we have mutated SE1 and SE2 and shown that mutations in potential binding sites for the SR protein ASF/SF2 dramatically reduced their splicing enhancer activity in vitro and dramatically affected alternative splicing in vivo. Mutation or deletion of either SE1 or SE2 in a BPV-1 late pre-mRNA switched alternative splicing from preferential use of the common 3′ splice site at nt 3225 to predominant use of the late-specific nt 3605 3′ splice site. In addition, multiple SR proteins including ASF/SF2 (SRp30a), SRp55, and SRp75 bound to both SE1 and SE2. Importantly, mutations in SE1 and SE2 which abolished function also significantly reduced or eliminated SR protein binding, indicating that the binding of SR proteins to these ESEs is functionally relevant. Finally, we have identified a 23-kDa protein which bound to SE2 but not SE1.

MATERIALS AND METHODS

Pre-mRNA substrates. Template plasmids for chimeric *doublesex* (*dsx*) pre-mRNA preparation were constructed by cloning of annealed synthetic oligonucleotides containing wild-type (WT) and mutant BPV-1 SE sequences into p*dsx*-Sa (44, 49) at *Xba*I and *Hin*dIII sites as described previously (56). The same vector containing a strong synthetic splicing enhancer [$(AAG)_8$] or Py3 (with four CT interruptions in the polypurine sequence [Fig. 2A]) served as a positive or negative control, respectively (44). The locations of the SE sequences in the BPV-1 genome are shown in Fig. 1A. A map of the *dsx* chimeric mRNAs is shown in Fig. 1B. All constructs were made by using standard cloning procedures and confirmed by sequencing (38).

The template plasmids were linearized with *Hin*dIII and transcribed in vitro with T7 RNA polymerase in the presence of the cap analog (m^7GpppG) (31) and [α-^{32}P]GTP. Full-length transcripts were separated on 6% denaturing polyacrylamide gels and eluted from gel slices by overnight incubation in elution buffer as described previously (56). The eluted RNAs were then phenol extracted, ethanol precipitated, and resuspended in diethyl pyrocarbonate-treated water.

To prepare the pre-mRNA substrates containing BPV-1 SE and no *dsx* sequences, the *dsx* sequences were removed from the above-described constructs by *Xba*I and *Bgl*II digestion, and the plasmid was recircularized by T4 DNA ligase. The resulting plasmids were then linearized with *Hin*dIII and transcribed as unlabeled RNAs for competition experiments or as uniformly labeled RNA, using [α-^{32}P]UTP, [α-^{32}P]ATP, [α-^{32}P]GTP, and [α-^{32}P]CTP for UV cross-linking and immunoprecipitation experiments.

In vitro splicing reactions. In vitro splicing reactions were performed at 30°C for 2 h, using the Promega RNA splicing system with 4 ng of ^{32}P-labeled pre-mRNAs and HeLa cell nuclear extracts as recommended by the manufacturer. The reaction products were extracted with phenol, precipitated with ethanol, and analyzed by electrophoresis on denaturing polyacrylamide gels followed by quantitation with a Molecular Dynamics PhosphorImager. The splicing efficiency for each pre-mRNA was calculated as the percentage of total splicing products (intermediates and fully spliced) divided by the sum of total splicing products plus the remaining pre-mRNA.

Competition experiments were performed by preincubation of HeLa cell nuclear extracts with cold competitor RNAs for 10 min on ice (44, 49). Radiolabeled pre-mRNAs and 5× splicing buffer (3% polyvinyl alcohol, 2 mM ATP, 100 mM creatine phosphate, 25 mM HEPES [pH 7.9]; Promega) were then added to give the standard splicing conditions as described by the manufacturer, and the reaction mixtures were incubated at 30°C for an additional 2 h. Analysis of the spliced products was carried out as described above.

BPV-1 expression vector and transformation mutagenesis. The BPV-1 late gene expression vector CCB458-2 (p3231) contains two fragments of BPV-1 DNA (nt 7254 to 7563 and nt 3073 to 7356) that were joined by PCR overlap extension with an intervening *Xho*I site. The BPV-1 DNA was cloned downstream of the human cytomegalovirus (CMV) IE1 promoter in pUC18. This vector expresses a BPV-1 late pre-mRNA that has a large deletion in intron 1, removing all of the early promoters. In vitro site-directed mutagenesis was used to delete BPV-1 SE1 or SE2 or to introduce specific base changes into SE1 or SE2. A Transformer site-directed mutagenesis kit (Clontech Laboratories, Inc.,

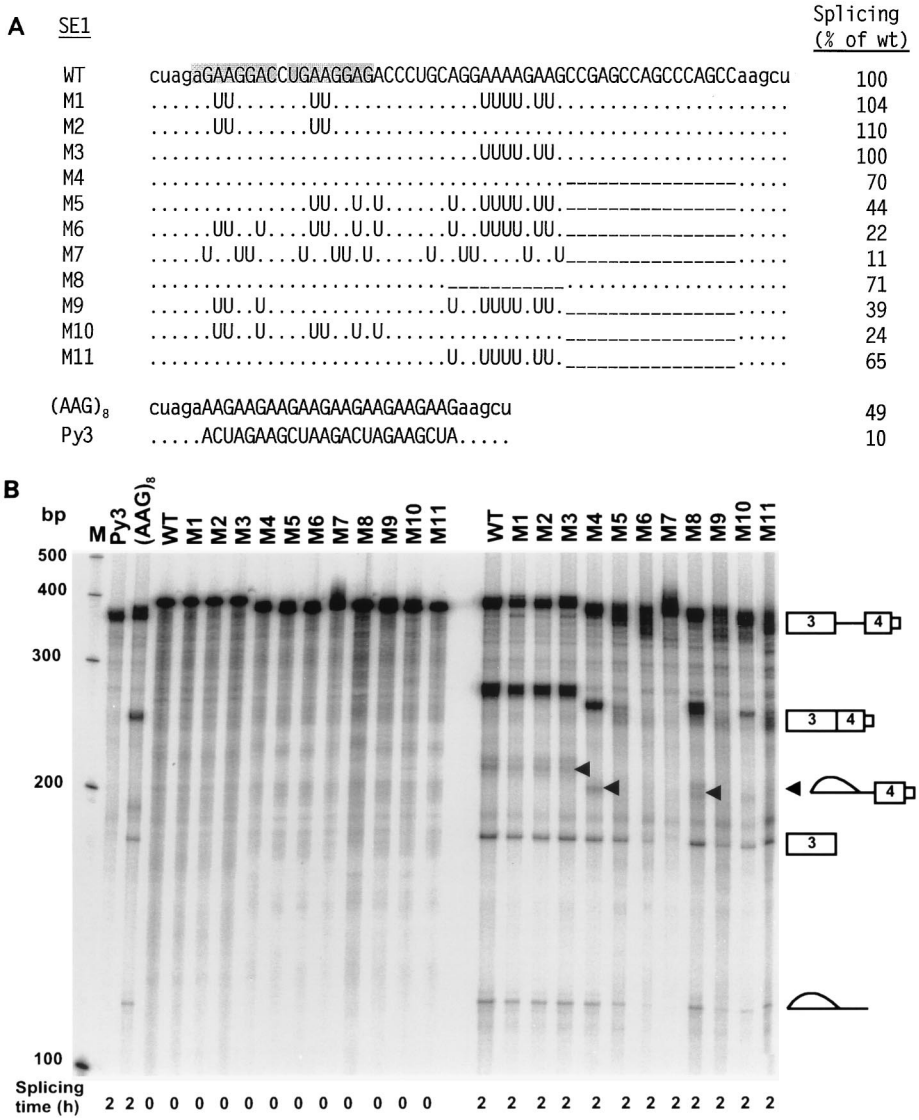

FIG. 2. Mutational analysis of BPV-1 SE1. Mutation (A or G to U) of BPV-1 SE1 reduces stimulation of *dsx* pre-mRNA splicing. (A) The names and sequences of WT and mutant SE1. The flanking sequences from the polylinker on both ends of the SE1 are indicated in lowercase. The shaded sequences on SE1 are potential ASF/SF2 binding sites which match most closely with the A13 sequence (AGAAGGAC) obtained through a SELEX procedure described in reference (43). In the mutant elements, only the mutated nucleotides are indicated. Unchanged nucleotides (.) and deletions (–) are indicated. The splicing efficiency of each pre-mRNA was calculated from the data in panel B as described in Materials and Methods and then expressed as a percentage of the WT value. (B) Effects of various SE1 mutations on *dsx* pre-mRNA splicing. The chimeric *dsx* pre-mRNAs containing WT and mutant BPV-1 SE1, $(AAG)_8$, or Py3 (44, 56) were incubated in HeLa cell nuclear extracts at 30°C for 2 h as indicated below the lanes. Electrophoresis was performed with a 5% polyacrylamide gel containing 8 M urea. Positions of DNA size markers (M) corresponding to a 100-bp ladder are shown at the left. The identities of the spliced products and intermediates are shown at the right. The mobilities of the fully spliced products and lariat-exon 4 intermediates differ among the various templates due to the different sizes of the ESE sequences downstream of exon 4. The top band in each lane is the unspliced pre-mRNA. One representative experiment of three is shown. The following plasmids were used to generate the chimeric *dsx* pre-mRNAs: WT, p3013; M1, p3027; M2, p3028; M3, p3029; M4, p3042; M5, p3043; M6, p3044; M7, p3045; M8, p3046; M9, p3047; M10, p3049; M11, p3048; $(AAG)_8$, p2701; and Py3, p3019.

Palo Alto, Calif.) was used for generation of the desired mutations or deletions. The primers used for mutagenesis and selection are summarized in Table 1. All mutations were verified by DNA sequencing.

Transfection and RT-PCR analysis of spliced BPV-1 late mRNAs. Transfectam reagent (Promega) was used to transfect 4 μg of WT or mutant expression vector DNA into human 293 cells or HeLa MK cells plated in 60-mm-diameter dishes. Total cellular RNA was prepared after 48 h, using TRIzol (GIBCO BRL) as instructed by the manufacturer. Following DNase I treatment, 500 ng of total cellular RNA was reverse transcribed at 42°C, using random hexamers as primers, and then amplified for 35 cycles as described previously (2). The primers used for reverse transcription (RT)-PCR analysis are summarized in Table 1. PCR products were then separated on a 3% agarose gel.

Proteins and antibodies. HeLa cell nuclear extracts used for in vitro splicing reactions and UV cross-linking assays were purchased from Promega. HeLa cell SR proteins were purified by two precipitation steps, using ammonium sulfate and magnesium (54). HeLa cell S100 was prepared according to the procedure of Dignam et al. (10). The mouse hybridoma cell line mAb104 (35, 36) (ATCC 2067-CRL) (referred to here as MAb 104) was obtained from R. Gontarek and D. Derse (National Cancer Institute-Frederick Cancer Research and Development Center). The mouse hybridoma cell line αSC35 (12) (ATCC 2031-CRL) was purchased from the American Type Culture Collection, and the hybridoma cell culture supernatant was immunoaffinity purified by using a Cappel goat anti-mouse immunoglobulin G (IgG) affinity column (Organon Teknika Corp.). MAb αSF2/ASF (42) and MAb 16H3 (30) were generous gifts from A. Krainer (Cold Spring Harbor Laboratory) and from K. Neugebauer and M. Roth (Fred Hutchinson Cancer Research Center, Seattle, Wash.), respectively.

UV cross-linking and immunoprecipitation. UV cross-linking procedures were modified from the literature (9, 25, 50). Briefly, purified HeLa SR proteins (20

TABLE 1. Primers used for transformation mutagenesis and RT-PCR analysis

Oligonucleotide	Sequence[a]	Position[b] (nt)
Mutagenic primers		
SE1m	GTCTGGGTCGCATCCtAAttACCTtAAttAtACCCTtCAttAAAAtAAtCCaAaCCAGCCCAGCC	3241–3305
SE1d	GTCTGGGTCGCATCC | CCAGCCCAGCCTG	3241–3255/3295–3307
SE2m	CCGGTGGACTTGGCATCAAttCAttAAtAAtAttAtCAtTCGCCCGACTCCACAtAttAAtAACC AGTGACTCTCCC	3460–3536
SE2d	CGGTACCGGTGGACTTGGC | CCAGTGACTCTCCCAAGG	3455–3473/3523–3540
Selection primer	GGGACGCcAATTgGTAATCATGG	—[c]
PCR primers		
Pr7250 (sense)	AATTATTGTGCTGGCTAGAC	7250–7269
Pr3746 (antisense)	AAGGTGATCAGTATTTGTGC	3746–3727

[a] Lowercase indicates nucleotide substitutions in BPV-1 wild-type sequence (mutagenic primers) or cloning sites (selection primers). Deletions are indicated by vertical lines.

[b] Positions of the first and last nucleotides of the primer or continuous segment of the primer. Deletions are indicated by "/". Numbers refer to nucleotide positions in the BPV-1 genome.

[c] —, the selection primer spans the *Eco*RI site at the junction between BPV-1 and pUC18 and destroys this restriction endonuclease cleavage site.

μg), nuclear extract (20 μg), or S100 extract (10 μg) were mixed with 5× splicing buffer (2 μl) and ^{32}P-labeled pre-mRNA (1.4×10^5 cpm/4 ng) in a total volume of 10 μl in each well of an ice-cold 96-well, flat-bottom plate. The plate was kept on ice for 15 min before addition of heparin to a final concentration of 2 mg/ml. After UV irradiation (480,000 μJ/cm^2) in a UV Stratalinker (Stratagene) at 4°C, each sample was digested for 30 min at 37°C with either RNase A (20 μg) or a combination of RNase A (20 μg), RNase T_1 (200 U), and RNase T_2 (40 U), heated at 100°C for 7 min in sodium dodecyl sulfate (SDS) sample buffer, and then analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) on a 12% gel. The dried gels were visualized with a Molecular Dynamics PhosphorImager and quantitated by using ImageQuant software.

For UV cross-linking and immunoprecipitation assays, 500 μl of MAb 16H3 or αSF2/ASF culture supernatant or 30 μg of affinity-purified αSC35 in 500 μl of IP (immunoprecipitation) buffer (10 mM Na_2PO_4 [pH 7.0], 150 mM NaCl, 20 mM EDTA, 5 mM potassium fluoride, 5 mM β-glycerophosphate, 2 μg of aprotinin per ml) was prebound to 50 μl of 50% GammaBind Plus Sepharose bead slurry (Pharmacia) for 1 h at 4°C. In the case of MAb 104, an anti-mouse IgM (1:500) was prebound to the beads. After three washes with IP buffer, 1 ml of MAb 104 (IgM type) culture supernatant was added to the beads precoated with anti-mouse IgM and incubated for another 1 h at 4°C. After three washes with IP buffer, 20 μl of the UV cross-linked protein-RNA mixtures was added to the antibody-prebound GammaBind beads, mixed by addition of 500 μl of IP buffer, and then rotated at 4°C overnight. After three washes with IP buffer, bound proteins were removed from the beads by addition of 25 μl of 2× SDS sample buffer and heating to 100°C for 7 min and then resolved by SDS-PAGE on a 12% gel.

Image analysis. Autoradiographic data were captured with a Molecular Dynamics PhosphorImager. The exposure of PhosphorImager data was adjusted with ImageQuant and exported as a TIFF file. Bitmapped images were cropped in Aldus Photostyler and imported into Micrografx Designer, where they were annotated and then printed on a Tektronix Phaser IIsdx dye sublimation printer.

RESULTS

Mutational analysis of BPV-1 ESEs SE1 and SE2. Our previous study demonstrated that either BPV-1 SE1 or SE2 can stimulate the splicing of a *dsx* pre-mRNA in vitro when located 44 nt downstream of the weak 3′ splice site in exon 4 (56) (lanes WT in Fig. 2B and 3B). We have used this assay to further characterize the sequence specificity of these purine-rich enhancer elements. In this study, A or G residues in SE1 or SE2 were substituted with U residues (Fig. 2A and 3A). The rationale for the U substitutions was based on a mutational analysis of an IgM purine-rich ESE which suggested that U residues are inhibitory while C residues are neutral (44).

The sequences of SE1 and SE2 were also scanned for potential binding sites for proteins known to bind to ESE elements. SR proteins have been shown to directly interact with purine-rich ESEs and mediate their function (23, 42, 43, 53). The optimal binding sites for ASF/SF2 and SC35 have been determined by an in vitro selection procedure (43). There are two adjacent potential ASF/SF2 binding sites at the 5′ end of SE1 (Fig. 2A) and two potential ASF/SF2 binding sites in SE2 that are separated by a central C-rich region (Fig. 3A). No good matches to the optimal SC35 binding site could be identified in either SE1 or SE2.

The results of the mutational analysis of SE1 are shown in Fig. 2. In a 2-h splicing reaction, *dsx* pre-mRNA containing either WT SE1 or the strong synthetic splicing enhancer $(AAG)_8$ as a positive control efficiently produced fully spliced products as well as splicing intermediates. Under the same conditions, a *dsx* pre-mRNA containing a mutant synthetic splicing enhancer (Py3) was only minimally spliced and served as a negative control. In these assays, the zero time point has been included to allow the splicing products to be easily distinguished from RNA degradation products. Substitutions of U's for A's in the two potential ASF/SF2 binding sites and/or the central A-rich stretch did not change the splicing enhancer activity of SE1 (M1, M2, and M3 [Fig. 2]). Deletion of either 16 nt at the 3′ end (SE1 M4) or deletion or mutation of 11 nt in the central A-rich part (SE1 M8 and M11) of SE1 reduced activity only about 30%. More extensive substitution of U's for A's in the two potential ASF/SF2 binding sites in combination with the 16-nt 3′-end deletion (SE1 M6 and M10) decreased SE1 activity another 50 to 60%. The most effective mutation was the substitution of U's for G's in combination with the 16-nt 3′ end deletion (SE1 M7). This mutation almost completely abolished the splicing enhancer activity of SE1, suggesting that G residues may be more important than A residues in the purine-rich ESE SE1. In general, most point mutations and deletion mutations had some effect on splicing enhancer function. In addition, the more extensive the mutation the greater the effect, suggesting that SE1 may be composed of multiple possibly overlapping binding sites.

As observed for SE1, substitutions of U's for A's within and surrounding the two potential ASF/SF2 binding sites in SE2 (M1) or deletion of the potential ASF/SF2 binding site from the 5′ end of SE2 (M4) only partially reduced splicing enhancer activity (Fig. 3). However, substitutions of U's for G's in either the 5′ or 3′ potential ASF/SF2 binding sites (SE2 M2 and M3) almost completely eliminated splicing enhancer activity. These results again suggest that G residues may be more important than A residues for splicing enhancer activity. In addition, these results suggest that both potential ASF/SF2 binding sites may be required for maximum SE2 splicing enhancer activity. Interestingly, deletion of the first potential

binding site had less of an effect than G-to-U substitutions in and around this site (compare M4 with M2 in Fig. 3A and B). This may indicate a role for sequences flanking this site. Alternatively, the G-to-U substitutions may create a sequence that binds other proteins and acts as a suppressor of splicing.

BPV-1 SE1 and SE2 are involved in alternative splice site selection in vivo. The results of the mutational analyses described above were used to design experiments to test the roles of SE1 and SE2 in splice site selection in vivo. In these experiments, we used a BPV-1 mini-late transcription unit expression vector driven by the CMV IE1 promoter (Fig. 4A). The pre-mRNA transcribed from this expression vector has a shortened first intron but otherwise contains all the splice sites and poly(A) sites required to generate the mRNAs normally processed from this pre-mRNA during a productive infection (1, 2). Point mutations and deletions which would be expected to eliminate splicing enhancer function were made in SE1 and SE2 (Fig. 4B). Mutation SE1m is similar to SE1 M7, while SE2m is similar to a combination of SE2 M2 and M3. The deletions entirely remove SE1 or SE2. WT and mutant expression vectors were assayed by transfection into 293 and HeLa cells. Total cellular RNA was isolated 48 h posttransfection, treated with RNase-free DNase I to eliminate contaminating DNA, and analyzed by RT-PCR using primers in exon 1 and exon 2 (Fig. 4A). The primers Pr7250 and Pr3746 would be expected to generate products of 658 and 278 bp from mRNAs spliced using the nt 3225 and 3605 3′ splice sites, respectively. An L1 cDNA and a 10:1 mixture of cloned cDNAs generated from L2 mRNAs spliced by using the nt 3225 and 3605 3′ splice sites were used as PCR controls (Fig. 4C, lanes 8 and 9, respectively). Steady-state levels of BPV-1 mRNAs in 293 cells transfected with the WT expression vector show a more than 10-fold preferentially use of the nt 3225 3′ splice site (Fig. 4C, lane 7). In contrast, late pre-mRNAs containing point mutations or deletions in either SE1 or SE2 were predominantly spliced using the nt 3605 3′ splice site (Fig. 4C, lanes 3 to 6). Similar results have been seen with a point mutation in the AG of the nt 3225 3′ splice site (1a). In these experiments, no band of 991 bp was ever seen, indicating that the pre-mRNA is either completely spliced or unstable and that essentially all input DNA was removed by DNase I treatment. Transfections of these plasmids were repeated in HeLa MK cells and gave similar splicing patterns (data not shown). These data confirm our previous in vitro studies (56) which suggested that utilization of the nt 3225 3′ splice site requires a functional SE1. In addition, SE2 appears to be required for preferential use of the nt 3225 3′ splice site.

***trans*-acting factors in HeLa cell nuclear extracts are required for the activity of the BPV-1 ESE SE2.** To determine if the splicing enhancement of *dsx* pre-mRNAs by BPV-1 ESEs is mediated by a specific interaction between SE sequences and *trans*-acting factors in HeLa cell nuclear extracts, competition experiments were conducted. HeLa nuclear extracts were preincubated on ice for 10 min with various cold competitor RNAs before addition of ^{32}P-labeled chimeric *dsx*–BPV-1 SE2 pre-mRNAs to the reaction mixture. Two sets of cold mRNAs were used for the competition. One set contained only ESE or control sequences, while the other set consisted of chimeric *dsx*–BPV-1 SE2, *dsx*-Py3, or *dsx*-$(AAG)_8$ pre-mRNA. As shown in Fig. 5, BPV-1 SE2 and the synthetic ESE $(AAG)_8$ inhibited in vitro splicing three- and fourfold, respectively, at a 165-fold (5 pM) molar excess. Under the same conditions, the nonfunctional Py3 RNA (44) inhibited splicing less than 1.6-fold. The *dsx*-containing competitors [*dsx*-SE2 and *dsx*-$(AAG)_8$] each suppressed splicing 3-fold at 165-fold (5 pM) molar excess, but competition again required the presence of a splicing enhancer

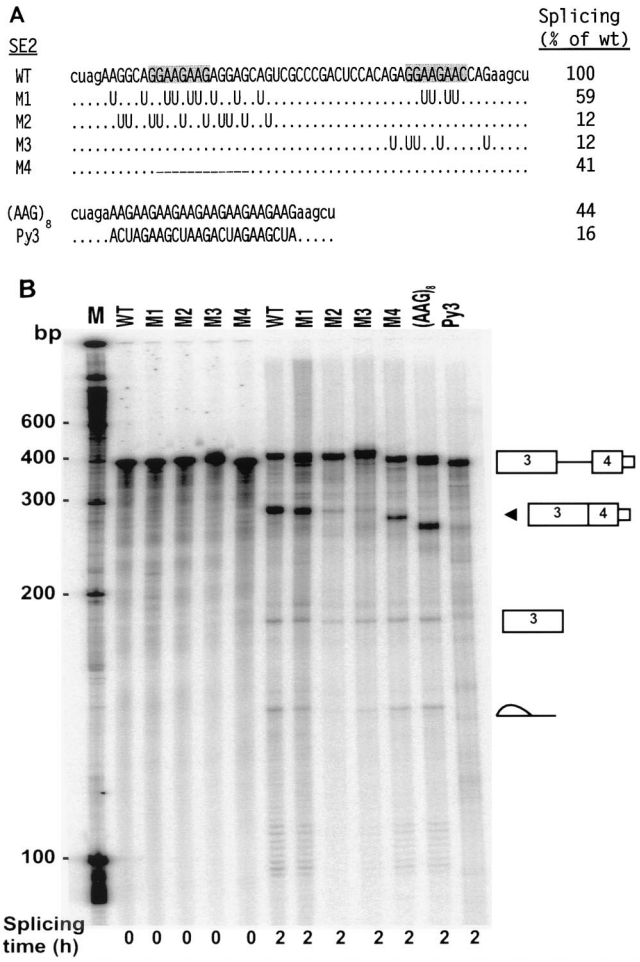

FIG. 3. Mutational analysis of SE2. Mutation (A or G to U) of BPV-1 SE2 reduces stimulation of *dsx* pre-mRNA splicing. (A) Names and sequences of the WT and mutant SE2 elements connected to *Drosophila dsx* exon 4. The flanking sequences from the polylinker on both ends of the SE2 are indicated in lowercase. The shaded sequences on SE2 are potential ASF/SF2 binding sites which closely match the consensus sequence (RGAAGAAC) obtained through a SELEX procedure described in reference (43). In the mutant elements, only the mutated nucleotides are indicated. Unchanged nucleotides (.) and deletions (–) are indicated. The splicing efficiency was calculated as described in the legend to Fig. 2. (B) Splicing gel after in vitro splicing reactions and gel electrophoresis. One representative experiment of two is shown. The following plasmids were used to generate the chimeric *dsx* pre-mRNAs: WT, p3014; M1, p3026; M2, p3050; M3, p3051; M4, p3052; $(AAG)_8$, p2701; and Py3, p3019. M, DNA size markers.

[Fig. 5; compare SE2 and $(AAG)_8$ with Py3]. These results indicate that *trans*-acting factors in HeLa nuclear extracts are required for SE activity.

UV cross-linking analysis of the proteins which bind to BPV-1 SE1 and SE2. UV cross-linking was used to identify proteins in HeLa nuclear extracts that bind to BPV-1 splicing enhancers. Since SR proteins have been shown to interact with purine-rich ESEs (23, 42, 43, 53), we also determined if purified SR proteins could bind to SE1 and SE2. In these assays, small in vitro-transcribed RNAs containing just SE1 or SE2 sequences and polylinker sequences were used for binding. In addition to WT SE sequences, the mutants with G-to-U substitutions (SE1 M7 and SE2 M2) were used to correlate protein binding with function since these mutants have minimal splicing enhancer activity. Both WT and mutant SE RNAs were labeled with [α-^{32}P]ATP, since WT and mutant RNAs

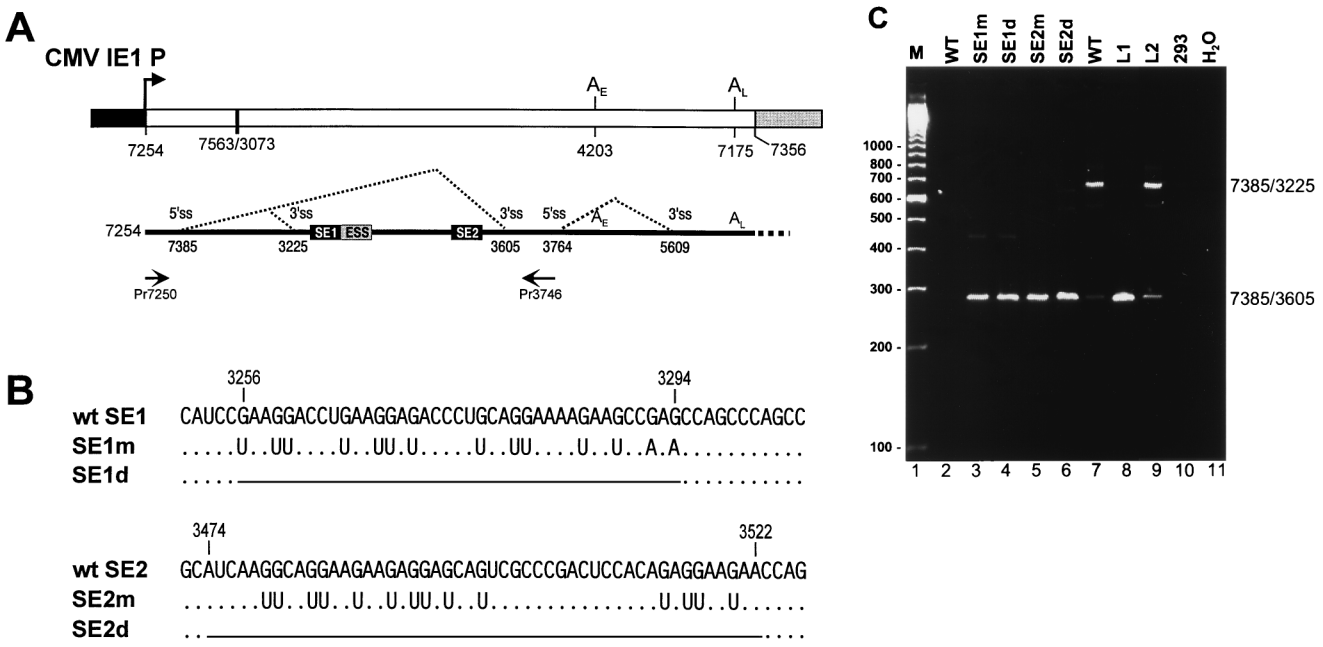

FIG. 4. Mutation or deletion of BPV-1 SE1 or SE2 switches alternative 3′ splice site selection in vivo. (A) Schematic diagrams of a BPV-1 mini-late transcription unit expression vector (top) and the splicing patterns of pre-mRNAs expressed from this vector (bottom). At each end of the mini-late transcription unit (open box) is the CMV IE1 promoter (solid box) and pUC18 (gray box). Numbers adjacent to or below the line indicate nucleotide positions in the BPV-1 genome. A large deletion in intron 1 of the L1 mRNA is shown by a heavy vertical line. Early and late poly(A) sites are indicated by vertical lines labeled A_E and A_L, respectively. Splice sites are shown as 5′ splice site (5′ss) and 3′ splice site (3′ss). The relative locations of SE1, SE2, and the ESS are shown as labeled boxes. Sense (Pr7250) and antisense (Pr3746) primers used for RT-PCR analysis reflect the position of the 5′ end of the primer. The drawing is not drawn to scale. (B) Nucleotide sequences of WT and mutant SE1- and SE2-containing pre-mRNAs. Mutant pre-mRNAs are identified at the left. Unchanged nucleotides (.) and deletions (horizontal line) are indicated. Deletion endpoints are labeled above each sequence and correspond to positions in the BPV-1 genome. The following plasmids were used to express these pre-mRNAs: WT, p3231; SE1m, p3031; SE1d, p3032; SE2m, p3033; and SE2d, p3034. (C) RT-PCR analysis of BPV-1 mRNAs transcribed and spliced in 293 cells from the expression vectors with SE1 or SE2 mutation or deletion. Total cell RNAs were extracted from the cells transfected by WT BPV-1 DNA (lanes 2 and 7) or the constructs containing SE1 mutation (SE1m; lane 3), SE1 deletion (SE1d; lane 4), SE2 mutation (SE2m; lane 5), or SE2 deletion (SE2d; lane 6), digested with RNase-free DNase I, and then subjected to conventional PCR amplification (lane 2) or RT-PCR analysis (lanes 3 to 7). Lane 8 and 9 are products amplified from a BPV-1 L1 cDNA and a 10:1 mixture of L2-L and L2-S cDNAs (2), respectively. Lane 10 is an untransfected 293 cell RNA control, and lane 11 is the water control (no RNA). PCR products are indicated as 5′ splice site/3′ splice site on the right. Marker (M) sizes are in base pairs.

contain equal numbers of adenosine residues and thus will be equally labeled. The labeled SE RNAs were then incubated with either HeLa nuclear extracts or purified SR proteins followed by UV cross-linking of bound proteins to the RNA and subsequent digestion of the RNA with RNase A. The proteins were then fractionated by SDS-PAGE and visualized by autoradiography. In this assay, only proteins which directly contact the RNA will be UV cross-linked to the RNA and will thus be labeled with covalently bound ^{32}P.

Both SE1 and SE2 RNAs were cross-linked to proteins in HeLa nuclear extracts with apparent molecular masses of 35 to 40, 55 to 60, and 70 to 80 kDa (Fig. 6A). SE2 RNA appeared to bind more of the 35- to 40-kDa protein(s) than did SE1, suggesting that there are more proteins in this molecular mass range that bind to SE2 or that the protein(s) has greater affinity for SE2 than for SE1. All three proteins appeared to be much less abundant or absent in S100 extracts. Since S100 extracts are specifically deficient in SR proteins (54), this finding suggests that these proteins might be SR proteins. The binding of SR proteins to SE1 and SE2 was confirmed by using an SR protein preparation purified from HeLa cells (Fig. 6A). Mutations which abolish the function of SE1 and SE2 also significantly decreased the binding of these proteins, suggesting that the binding of these proteins is functionally relevant (compare SE1 WT and M7 lanes and SE2 WT and M2 lanes in Fig. 6A). However, the mutant RNAs also bound proteins which did not bind to the WT sequences.

SE2 also bound an additional protein with an apparent molecular mass of approximately 28 kDa in HeLa nuclear extracts (Fig. 6A). This protein appeared to be enriched in HeLa S100 extracts and was absent in SR protein preparations, suggesting that the 28-kDa protein is not a classical SR protein. In a separate experiment, the 28-kDa protein from HeLa S100 extracts bound strongly to WT SE2 and the synthetic ESE $(AAG)_8$ (Fig. 6B). In both experiments, SE1 bound the 28-kDa protein only weakly (Fig. 6). Again, mutations which abolished the splicing enhancer activity of SE2 and $(AAG)_8$ also dramatically reduced binding of the 28-kDa protein, suggesting that the binding of this protein is also functionally relevant (M2 and Py3 in Fig. 6B).

We noted that the apparent molecular weights of the proteins bound to the SE elements in nuclear extracts was slightly higher than that of the purified SR proteins (Fig. 6A). Although the reason for this difference is unknown, we speculate that this could be due to differences in the phosphorylation state of the proteins resulting from phosphatase activity during purification. We also noted that both SE1 and SE2 contain relatively long purine-rich stretches which are resistant to digestion with RNase A (Fig. 7). These RNase A-resistant RNA oligonucleotides could alter the mobility of a protein to which they are cross-linked. This may explain several other observations. Most of the proteins labeled by UV cross-linking migrated as broad fuzzy bands or multiple closely spaced bands. In addition, the apparent molecular weights of the purified SR

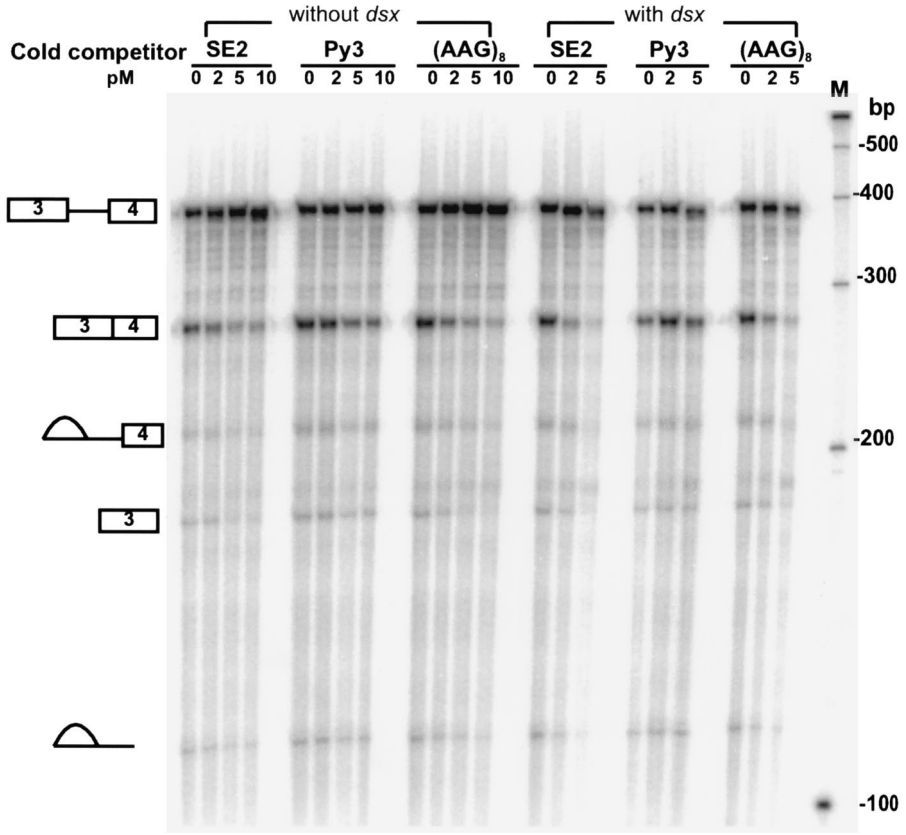

FIG. 5. Splicing factors binding to SE2 can be competed out with cold SE2 and $(AAG)_8$ RNAs or chimeric *dsx*-SE2 or *dsx*-$(AAG)_8$ pre-mRNAs. The in vitro splicing was carried out with ^{32}P-labeled *dsx*-SE2 pre-mRNA as the substrate in the presence of various cold competitor RNAs with or without *dsx*. HeLa cell nuclear extracts (10 μl) were first incubated for 10 min on ice with each competitor RNA indicated above the gel at a concentrations of 2, 5, or 10 pM. The ^{32}P-labeled *dsx*-SE2 pre-mRNA (4 ng) was then added to the reaction mixture, and splicing was carried out at 30°C for a further 2 h. The molar ratio between the cold and hot RNAs was 66-fold (2 pM), 165-fold (5 pM), and 330-fold (10 pM) molar excess of the cold competitors to hot substrates. Electrophoresis was performed with a 5% polyacrylamide gel containing 8 M urea. Positions of DNA size markers corresponding to a 100-bp ladder are shown on the right. The identities of the spliced products and intermediates are shown at the left. Splicing efficiency was calculated as described in Materials and Methods. RNAs without *dsx* sequences were transcribed from the following plasmids: SE2, p3041; Py3, p3040; and $(AAG)_8$, p3039. RNAs with *dsx* sequences were transcribed from the following plasmids: SE2, p3014; Py3, p3019; and $(AAG)_8$, p2701.

proteins were higher than those of the known SR proteins, and the apparent mobilities of the SR proteins labeled by UV cross-linking to SE2 RNA were higher than those of the SR proteins labeled by UV cross-linking to SE1 RNA. To determine if the protein mobilities were influenced by the cross-linked RNA, binding reactions were extensively digested with a combination of RNases (A, T_1, and T_2) which should cut after every nucleotide. For these experiments, SE1 and SE2 RNAs were uniformly labeled with all four α-^{32}P-labeled ribonucleoside triphosphates. This ensured that all UV cross-linked proteins would be labeled with ^{32}P even if only one nucleotide remained covalently attached to the protein. Combined RNase treatment significantly increased the mobility of all the proteins and reduced broad smears to one or more discrete bands (Fig. 7). This result indicates that the RNA does significantly alter the mobility of the proteins. In addition, SE1 and SE2 RNAs gave nearly identical protein profiles (except for a 23-kDa protein) with both HeLa nuclear extracts and HeLa SR protein preparations when all three RNases were used to digest the protein-RNA mixtures after UV cross-linking (Fig. 7). Thus, with the exception of the 23-kDa protein which binds predominantly to SE2, SE1 and SE2 appear to bind the same proteins. The 23-kDa protein seen with combined RNase digestion is most likely the same protein as the 28-kDa protein seen with RNase A digestion alone. The proteins with apparent molecular masses of 35 to 40, 55 to 60, and 70 to 80 kDa seen with RNase A digestion probably correspond to the SR proteins with apparent molecular masses of 35, 55, and 75 kDa which are seen with combined RNase digestion. Although only one protein with molecular mass of 35 kDa was seen in this size range in HeLa SR protein preparations, three proteins with molecular masses of 34, 35, and 36.5 kDa could be cross-linked to SE1 and SE2 in HeLa nuclear extracts. Thus, in addition to the SR proteins, several apparently non-SR proteins also bind to SE1 and SE2.

Identification of SR proteins cross-linked to BPV-1 ESEs by immunoprecipitation. Although the apparent molecular weights of the proteins which bound to BPV-1 ESEs are consistent with those of the classical members of the SR protein family, it is possible that one or more of these proteins are contaminants in the SR protein preparation and are therefore not SR proteins. To confirm that SR proteins do bind to SE1 and SE2, MAbs specific for SR proteins were used to immunoprecipitate the proteins which were labeled by UV cross-linking to radiolabeled SE1 and SE2 RNAs. Three MAbs were chosen for these experiments. MAb 104 recognizes a phosphoepitope in the RS domain of all classical SR proteins (36, 54). MAb 16H3 recognizes an epitope composed of alternating

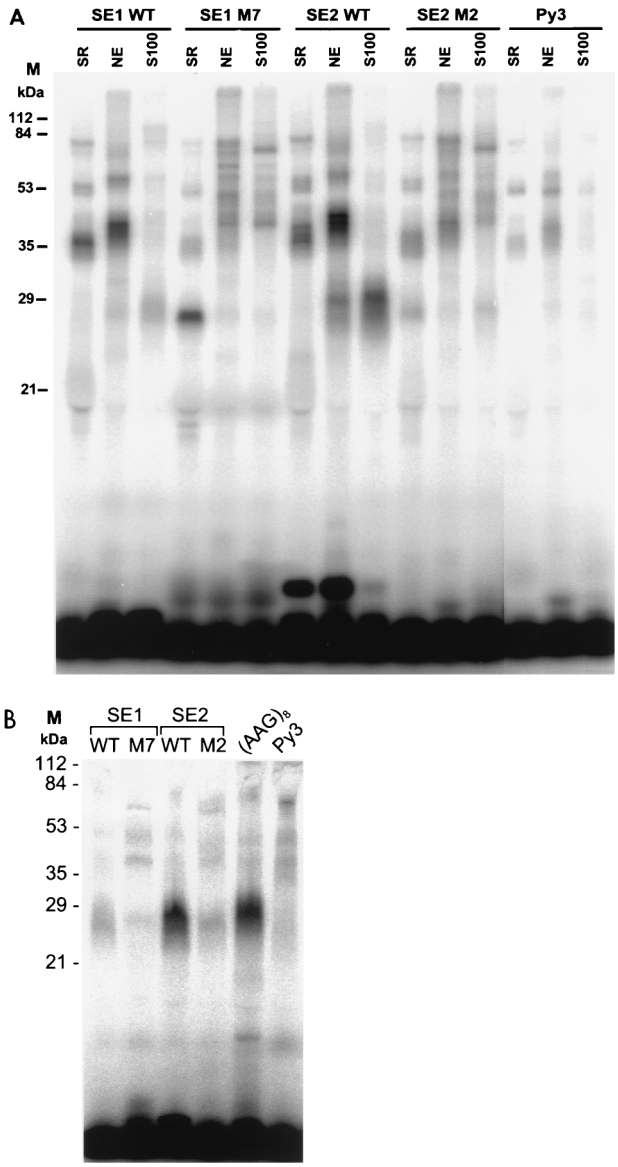

FIG. 6. UV cross-linking analysis of proteins that bind to SE1 and SE2. (A) UV cross-linking of BPV-1 SE1 and SE2 to purified HeLa SR proteins, nuclear extracts, and S100 preparations. Purified HeLa SR proteins, HeLa nuclear extracts, or HeLa S100 extracts were cross-linked by UV irradiation (see Materials and Methods) to pre-mRNAs labeled with [α-^{32}P]ATP and digested with RNase A. The protein-RNA mixtures were then heated at 100°C for 7 min in SDS sample buffer and analyzed by SDS-PAGE on a 12% gel. (B) UV cross-linking of a 28-kDa protein present in HeLa S100 extracts to BPV-1 SE2. RNAs were labeled with [α-^{32}P]ATP. UV cross-linking, RNase A digestion, and electrophoresis were carried out as described in Materials and Methods. Pre-mRNAs were transcribed from the following plasmids: SE1 WT, p3053; SE1 M7, p3055; SE2 WT, p3041; SE2 M2, p3056; $(AAG)_8$, p3039; and Py3, p3040. M, size markers.

arginine and glutamate or aspartate residues which is found in a family of splicing factors that includes SRp20, SRp40, SRp55, SRp75, U2AF^{65}, U2AF^{35}, U1-70K, and HRH1 (human RNA helicase 1), as well as additional proteins (30). However, this MAb does not recognize SRp30a (ASF/SF2) or SRp30b (SC35). MAb αSF2/ASF recognizes only ASF/SF2 (SRp30a) (42). A MAb to the BPV-1 major capsid protein (L1) and fetal bovine serum were used as nonspecific controls. For these experiments, we used SE1 or SE2 RNA labeled with [α-^{32}P] ATP and digested the UV cross-linked RNA-protein mixtures with RNase A. Although RNase A digestion gives broader bands than digestion with a mixture of RNase A, T_1, and T_2, it also leaves larger RNA fragments attached to the protein and therefore gives a stronger signal.

The results of the immunoprecipitation experiments show that the major proteins which cross-linked to both SE1 and SE2 are SR proteins SRp30a (ASF/SF2), SRp55, and SRp75 (Fig. 8). Each of these proteins was UV cross-linked much more efficiently to WT SE sequences than to mutant SE sequences, again suggesting that the binding of these proteins is functionally relevant. Each antibody gave a distinct protein profile. MAb 104 preferentially immunoprecipitated SRp55 better than SRp75 and SRp30a/b (Fig. 8A), while MAb 16H3 precipitated more SRp75 than SRp55 (Fig. 8B). No significant band was seen in the region of 30 to 35 kDa after immunoprecipitation with MAb 16H3 (Fig. 8B). This result was expected since MAb 16H3 does not recognize SRp30a/b (30). The binding of ASF/SF2 to both SE1 and SE2 was confirmed by immunoprecipitation with MAb αSF2/ASF (Fig. 8C). Surprisingly, this antibody also precipitated a 28-kDa protein which specifically cross-linked to WT SE2 but not to mutant SE2 or to SE1. Unlike ASF/SF2, which was more abundant in the purified SR protein preparations, the 28-kDa protein was more abundant in nuclear extracts. These two observations suggest that it is the same 28-kDa protein that was abundant in S100 extracts (Fig. 6). Further immunoprecipitation assays showed that the 28-kDa protein seen in HeLa nuclear extracts could also be immunoprecipitated with MAbs 104 and 16H3 (data not shown), suggesting that this protein might actually be an SR or SR-related protein even though it does not copurify

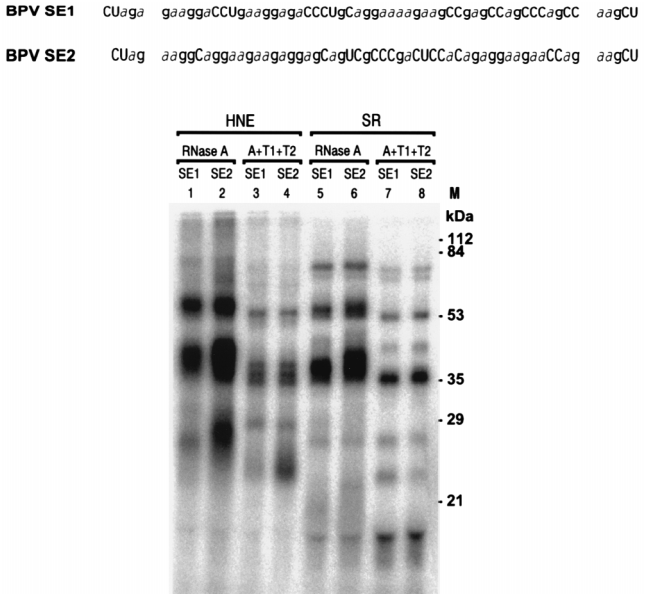

FIG. 7. Choice of RNase affects mobility of proteins labeled by UV cross-linking. The SE1 and SE2 sequences are shown at the top, with the flanking polylinker sequences at each end separated by a space. The RNases used digest after the following nucleotides and leave a 3′ phosphate: capital letters, RNase A (pyrimidine residues); bold lowercase letters, RNase T_1 (G); or italicized lowercase letters, RNase T_2 (A). RNAs used for the experiments were uniformly labeled with [α-^{32}P]ATP, [α-^{32}P]UTP, [α-^{32}P]GTP, and [α-^{32}P]CTP. UV cross-linking of SE1 and SE2 to HeLa nuclear extracts (HNE) or SR proteins is described in Materials and Methods. The RNA-protein mixtures were digested at 37°C for 30 min with RNase A (20 μg) only or with a mixture of RNases A (20 μg), T_1 (200 U), and T_2 (40 U), heated at 100°C for 7 min in SDS sample buffer, and then analyzed by SDS-PAGE on a 12% gel.

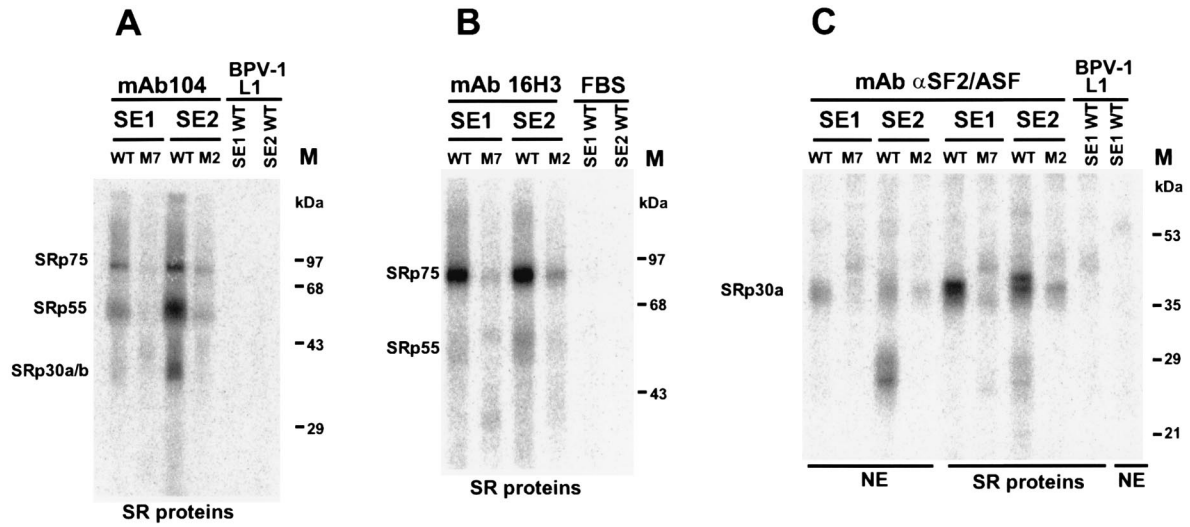

FIG. 8. Identification by immunoprecipitation of SR proteins cross-linked to BPV-1 SE1 and SE2. Purified HeLa SR proteins (A to C) or HeLa nuclear extracts (NE; C) were used for the UV cross-linking experiments. The hot RNAs used for the experiments were labeled with [α-^{32}P]ATP. The cross-linked protein-RNA mixtures were digested by RNase A (20 μg) only. After precleaning with GammaBind Plus Sepharose beads (no antibody bound), the protein-RNA mixtures were immunoprecipitated overnight at 4°C, using MAb 104 (A), MAb 16H3 (B), or MAb αSF2/ASF (C) prebound to GammaBind Plus Sepharose beads (Pharmacia). The immunoprecipitates on the beads were washed extensively, heated at 100°C for 7 min in SDS sample buffer, and fractionated by SDS-PAGE electrophoresis on a 12% gel. MAbs against BPV-1 L1 protein (A and C) or 20% fetal bovine serum (FBS) in culture medium (B) were used as negative controls. M, size markers.

with the classical SR proteins and is abundant in SR deficient HeLa S100 preparations (Fig. 6).

We have tried to use immunoprecipitation to determine if SRp30b (SC35) binds to either SE1 or SE2. Unfortunately, our immunoaffinity-purified MAb αSC35 did not specifically precipitate SRp30b but rather appeared to have a specificity similar to that of MAb 104 (data not shown). Thus, we are unable to say whether SRp30b binds to either of the BPV-1 ESEs.

DISCUSSION

Purine-rich ESEs (previously termed exon recognition sequences) have now been described in many cellular and viral pre-mRNAs. These elements are generally paired with suboptimal splice sites (typically 3′ splice sites) in pre-mRNAs that are subject to alternative splicing. In most cases, these elements stimulate splicing of the upstream intron or inclusion of an internal exon. However, exceptions to this rule were noted recently in adenovirus and Rous sarcoma virus pre-mRNAs, where purine-rich elements in introns have been shown to serve as splicing repressors (21, 27). The characteristics of BPV-1 SE1 and SE2 are typical of other ESEs. Both elements are purine rich and contain several clusters of the purine repeat GAR or GRA (R = G or A) which has been identified in other ESEs (33, 52). Each is capable of enhancing the splicing of a heterologous pre-mRNA containing a suboptimal splice site (56). In BPV-1 pre-mRNAs, SE1 lies 31 nt downstream of the nt 3225 3′ splice site and has previously been shown to be essential for utilization of this suboptimal 3′ splice site in vitro (56). Here we show that SE1 is also essential for utilization of this 3′ splice site in the context of an alternatively spliced BPV-1 late pre-mRNA in vivo (Fig. 4). Although SE2 functions as an ESE in heterologous pre-mRNAs, an enhancer-like role for SE2 in BPV-1 late pre-mRNA splicing could not be demonstrated in vitro because of the dominant action of a suppressor element that lies between SE1 and SE2 (56). Our in vivo data now indicate that SE2 is also required for preferential utilization of the nt 3225 3′ splice site in an alternatively spliced late pre-mRNA. Deletion or mutation of either SE1 or SE2 switches splicing to a downstream alternative late-specific 3′ splice site at nt 3605 (Fig. 4). However, the exact mechanism by which SE2 functions is not known. Although the data are consistent with a role as an ESE that acts on the nt 3225 3′ splice site, SE2 lies 252 nt downstream of this 3′ splice site and may not be able to function as an ESE at that distance (46). Alternatively, SE2 lies only 80 nt upstream of the nt 3605 3′ splice site and by analogy with the adenovirus 3RE (21) may function as an intronic splicing repressor by blocking the binding of U2 snRNP at the branch point of the nt 3605 3′ splice site. Although our in vivo and in vitro data do not distinguish between these two mechanisms, we favor the latter mechanism. A proposed model for the regulation of BPV-1 late pre-mRNA alternative splicing is shown in Fig. 9. In this model, splicing is delicately balanced by the suboptimal nature of both 3′ splice sites and by the presence of additional *cis* elements. SE1 and the ESS act as positive and negative modulators, respectively, of the nt 3225 3′ splice site, while SE2 functions as either a repressor of the nt 3605 3′ splice site or an enhancer for the nt 3225 3′ splice site. The role of *trans*-acting factors in this regulation will be discussed further below.

In this report, we have also carried out an in vitro structure function analysis of both SE1 and SE2 and used this information to correlate protein binding with function. Analysis of the sequences of SE1 and SE2 revealed two potential ASF/SF2 binding sites in each ESE (43). Substitution of U residues for all G residues in and around these potential binding sites inactivated both SE1 and SE2 (Fig. 2 and 3), suggesting that ASF/SF2 binding may be required for the function of these ESEs. These results are consistent with those of Tanaka et al. (44), who showed that U residues in an ESE are deleterious. Surprisingly, however, substitution of U residues for A residues frequently had only a relatively minor effect on BPV-1 SE function (Fig. 2 and 3). Similar results were seen in a study by Staknis and Reed (40): the avian sarcoma-leukosis virus ESE containing six A to U substitutions still efficiently assembled the enhancer complex in vitro, although SR protein composition of this complex differed from the WT composition. The mutational analysis of SE1 and SE2 may indicate that the G

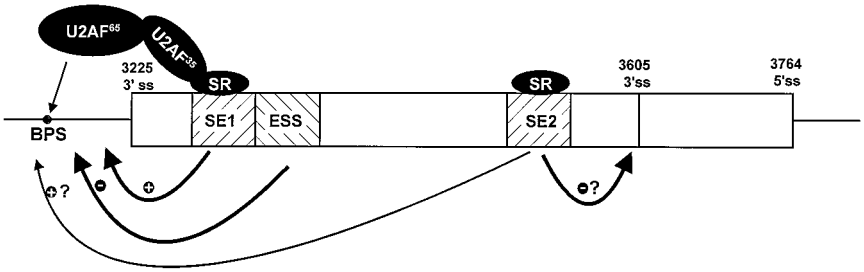

FIG. 9. A proposed model for regulation of BPV-1 alternative 3′ splice site selection. Rectangles, exon; lines, intron; 5′ss, 5′ splice site; 3′ss, 3′ splice site; SR, SR proteins; ?, hypothetical interactions.

residues make more important RNA-protein contacts than do the A residues. An alternative explanation is suggested by the observation that the A-to-U mutations in SE1 and SE2 create multiple GU sequences which may form cryptic 5′ splice sites. A downstream 5′ splice site has been shown to activate the splicing of an upstream intron (reviewed in reference 4) and can function as a splicing enhancer (49).

As mentioned previously, the function of ESEs is mediated by the SR proteins which bind to the ESE and facilitate early steps of spliceosomal assembly by recruiting U2AF to suboptimal 3′ splice sites. SR proteins have been shown to bind several ESEs, including those in the human fibronectin alternative exon ED1 (23), bovine growth hormone exon 5 (40, 42), avian sarcoma-leukosis virus *env* 3′ exon (40), equine infectious anemia virus exon 3 (16), chicken cardiac troponin T exon 5 (33), and *Drosophila dsx* exon 4 (24, 45, 46). Although ASF/SF2 is the best-studied example, other SR proteins (SRp40, SRp55, and SRp75) also bind to these enhancer elements (16, 33, 53). SR protein binding to these elements can be either direct (23, 33, 40, 42) or mediated through other sequence-specific RNA-binding proteins (24, 45, 46, 53).

Consistent with these studies, we have shown by UV cross-linking and immunoprecipitation that at least three SR proteins (SRp30a, SRp55, and SRp75) bind specifically to both BPV-1 SE1 and SE2 (Fig. 6 to 8). Furthermore, this binding was decreased significantly by mutations that block the function of the enhancers, strongly suggesting that SR protein binding is required for the function of these enhancers. Three bands which run with mobilities expected for SRp30 were seen when SE1 and SE2 RNAs were incubated with HeLa nuclear extracts and digested with a combination of RNase A, T_1, and T_2 after UV cross-linking (Fig. 7). One of these proteins was demonstrated to be SRp30a, using an ASF/SF2-specific antibody (Fig. 8). However, we were unable to use immunoprecipitation to determine if SRp30b (SC35) or SRp30c binds to SE1 or SE2, since the αSC35 antibody was nonspecific and very inefficient in our assays and an antibody against SRp30c was unavailable during this study. The other two bands may be non-SR proteins, since only one distinct band was seen in assays using SR proteins purified from HeLa cells (Fig. 7). It is possible that one of these proteins is the 37-kDa protein which has been shown to bind to repetitive GAA enhancer elements and promote the binding of SRp40 (53). The binding of ASF/SF2 (SRp30a) to SE1 and SE2 is consistent with the observation that each element contains two potential ASF/SF2 binding sites (Fig. 2A and 3A). Furthermore, G-to-U mutations in both binding sites in SE1 eliminated both function (Fig. 2) and ASF/SF2 binding (Fig. 8). G-to-U mutations in either one of the two ASF/SF2 binding sites in SE2 also dramatically reduced splicing enhancer activity (SE2 M2 and M3 [Fig. 3]), and mutation of only one of these sites (SE2 M3 [Fig. 3 and 8]) also reduced ASF/SF2 binding considerably more than twofold. One explanation for this is that more than one copy of the binding sites may be required for cooperative SR protein binding and splicing activation as reported by Tacke and Manley (43).

In addition to the SR proteins mentioned above, an approximately 23-kDa protein (28 kDa without combined RNase A, T_1 and T_2 digestion) which binds to SE2 but not to SE1 or mutant SE2 was identified. Binding of this protein was seen in assays using HeLa nuclear extracts and S100 extracts but not purified SR protein preparations. This result indicates that the protein is not a typical SR protein. Surprisingly, the protein-RNA complexes were immunoprecipitable by all three MAbs, 104, 16H3, and αSF2/ASF, suggesting that it may actually be an SR or SR-related protein (Fig. 8 and data not shown). This protein is unlikely to be an alternatively spliced form of ASF/SF2 because ASF-2 has an approximate molecular mass of 40 kDa and ASF-3 lacks the RS domain which contains the phosphoepitope recognized by MAb 104 (15). It is possible that this protein is an SR protein that is insoluble in 65% ammonium sulfate or is not precipitated by 20 mM $MgCl_2$. Alternatively, it may simply be coimmunoprecipitated as part of an enhancer complex. The identity of this protein remains unknown. However, since the specific interaction between the 23-kDa protein and SE2 occurred in the absence of functional splice sites, this protein is not likely to be the same 28-kDa protein which binds to the branch point (14).

There is now extensive evidence that SR proteins mediate the function of ESEs by facilitating the recruitment of splicing factors to a suboptimal splice site at early stages of spliceosomal assembly. Our data suggest a working model for the splicing of BPV-1 late pre-mRNAs (Fig. 9). In this model, initial stages of 3′ splice site recognition first involve the binding of one or more SR proteins to high-affinity binding sites on SE1 or SE2. This may be followed by the recruitment of additional SR proteins to the splicing enhancer with an increase in the stability of the complex. In the case of SE1, this SR protein complex then recruits the heterodimeric protein U2AF to the 3′ splice site through protein-protein interactions between the RS domain of the SR proteins and the RS domain of $U2AF^{35}$ (51, 57). Thus, $U2AF^{35}$ functions as a bridge between $U2AF^{65}$ and the enhancer complex to recruit $U2AF^{65}$ to the adjacent intron (57). $U2AF^{65}$ then interacts with the polypyrimidine tract and branch point of the pre-mRNA (14, 47, 55, 57) and promotes the base pairing of the U2 snRNA with the RNA branch point (14, 47). The location of SE2 in the intron near the nt 3605 3′ splice site suggests that SR protein binding to SE2 may simply block interaction of the U2 snRNP with the branch point of this splice site, possibly by steric hindrance. The mechanism by which the BPV-1 ESS functions remains to be understood. This arrangement of *cis* elements provides a

mechanism for the coordinated regulation of alternative splicing by the same set of proteins. When the activities of SR proteins are high, the nt 3225 3′ splice site will be used preferentially through enhancement of spliceosomal assembly at the nt 3225 3′ splice site and possibly also by repression of assembly at the nt 3605 3′ splice site. As SR protein activity decreases, repression of nt 3225 3′ splice site utilization by the ESS dominates over SE1-mediated enhancement and repression of the nt 3605 3′ splice site by SE2 is lost. The result is a switch in 3′ splice site utilization from nt 3225 to 3605. This switch in 3′ splice site utilization takes place at late stages of keratinocyte differentiation, suggesting that SR protein activity or abundance may be regulated by keratinocyte differentiation. This possibility is currently being investigated.

ACKNOWLEDGMENTS

We thank Yoshiro Shimura for providing the p*dsx*-Sa construct, Adrian Krainer for providing the SF2 MAb, Karla Neugebauer and Mark Roth for providing MAb 16H3, Richard Gontarek and David Derse for providing mouse hybridoma cell line MAb 104, and Mark McNally for providing HeLa SR proteins used in our initial experiments. We also thank Steve Mount and Susan Haynes for critical reviews of the manuscript.

REFERENCES

1. **Baker, C. C., and P. M. Howley.** 1987. Differential promoter utilization by the bovine papillomavirus in transformed cells and productively infected wart tissues. EMBO J. **6:**1027–1035.
1a. **Baker, C. C.** Unpublished results.
2. **Barksdale, S. K., and C. C. Baker.** 1993. Differentiation-specific expression from the bovine papillomavirus type 1 P_{2443} and late promoters. J. Virol. **67:**5605–5616.
3. **Barksdale, S. K., and C. C. Baker.** 1995. Differentiation-specific alternative splicing of bovine papillomavirus late mRNAs. J. Virol. **69:**6553–6556.
4. **Berget, S. M.** 1995. Exon recognition in vertebrate splicing. J. Biol. Chem. **270:**2411–2414.
5. **Black, D. L.** 1995. Finding splice sites within a wilderness of RNA. RNA **1:**763–771.
6. **Bruzik, J. P., and T. Maniatis.** 1995. Enhancer-dependent interaction between 5′ and 3′ splice sites in trans. Proc. Natl. Acad. Sci. USA **92:**7056–7059.
7. **Caceres, J. F., and A. R. Krainer.** 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. EMBO J. **12:**4715–4726.
8. **Colwill, K., T. Pawson, B. Andrews, J. Prasad, J. L. Manley, J. C. Bell, and P. I. Duncan.** 1996. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J. **15:**265–275.
9. **Cote, G. J., D. T. Stolow, S. Peleg, S. M. Berget, and R. F. Gagel.** 1992. Identification of exon sequences and an exon binding protein involved in alternative RNA splicing of calcitonin/CGRP. Nucleic Acids Res. **20:**2361–2366.
10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
11. **Fu, X. D.** 1995. The superfamily of arginine/serine-rich splicing factors. RNA **1:**663–680.
12. **Fu, X. D., and T. Maniatis.** 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature **343:** 437–441.
13. **Fu, X. D., and T. Maniatis.** 1992. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3′ splice site. Proc. Natl. Acad. Sci. USA **89:**1725–1729.
14. **Gaur, R. K., J. Valcarcel, and M. R. Green.** 1995. Sequential recognition of the pre-mRNA branch point by $U2AF^{65}$ and a novel spliceosome-associated 28-kDa protein. RNA **1:**407–417.
15. **Ge, H., P. Zuo, and J. L. Manley.** 1991. Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. Cell **66:** 373–382.
16. **Gontarek, R. R., and D. Derse.** 1996. Interactions among SR proteins, an exonic splicing enhancer, and a lentivirus Rev protein regulate alternative splicing. Mol. Cell. Biol. **16:**2325–2331.
17. **Gui, J.-F., W. S. Lane, and X.-D. Fu.** 1994. A serine kinase regulates intracellular localization of splicing factors in the cell cycle. Nature **369:**678–683.
18. **Gui, J.-F., H. Tronchère, S. D. Chandler, and X.-D. Fu.** 1994. Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. Proc. Natl. Acad. Sci. USA **91:**10824–10828.
19. **Hodges, D., and S. I. Bernstein.** 1994. Genetic and biochemical analysis of alternative RNA splicing. Adv. Genet. **31:**207–281.
20. **Jamison, S. F., A. Crow, and M. A. Garcia-Blanco.** 1992. The spliceosome assembly pathway in mammalian extracts. Mol. Cell. Biol. **12:**4279–4287.
21. **Kanopka, A., O. Muhlemann, and G. Akusjarvi.** 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. Nature **381:**535–538.
22. **Kohtz, J. D., S. F. Jamison, C. L. Will, P. Zuo, R. Luhrmann, M. A. Garcia-Blanco, and J. L. Manley.** 1994. Protein-protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. Nature **368:**119–124.
23. **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.
24. **Lynch, K. W., and T. Maniatis.** 1995. Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes Dev. **9:**284–293.
25. **MacMillan, A. M., C. C. Query, C. R. Allerson, S. Chen, G. L. Verdine, and P. A. Sharp.** 1994. Dynamic association of proteins with the pre-mRNA branch region. Genes Dev. **8:**3008–3020.
26. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. Genes Dev. **10:**1569–1579.
27. **McNally, L. M., and M. T. McNally.** 1996. SR protein splicing factors interact with the Rous sarcoma virus negative regulator of splicing element. J. Virol. **70:**1163–1172.
28. **Michaud, S., and R. Reed.** 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes Dev. **5:**2534–2546.
29. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to messenger RNAs by the spliceosome, p. 303–357. *In* R. F. Gestland and J. F. Atkins (ed.), The RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
30. **Neugebauer, K. M., J. A. Stolk, and M. B. Roth.** 1995. A conserved epitope on a subset of SR proteins defines a larger family of pre-mRNA splicing factors. J. Cell Biol. **129:**899–908.
31. **Nielsen, D. A., and D. J. Shapiro.** 1986. Preparation of capped RNA transcripts using T7 RNA polymerase. Nucleic Acids Res. **14:**5936.
32. **Nilsen, T. W.** 1994. RNA-RNA interactions in the spliceosome: unraveling the ties that bind. Cell **78:**1–4.
33. **Ramchatesingh, J., A. M. Zahler, K. M. Neugebauer, M. B. Roth, and T. A. Cooper.** 1995. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. Mol. Cell. Biol. **15:**4898–4907.
34. **Rossi, F., E. Labourier, T. Forne, G. Divita, J. Derancourt, J. F. Riou, E. Antoine, G. Cathala, C. Brunel, and J. Tazi.** 1996. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. Nature **381:**80–82.
35. **Roth, M. B., C. Murphy, and J. G. Gall.** 1990. A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. J. Cell Biol. **111:**2217–2223.
36. **Roth, M. B., A. M. Zahler, and J. A. Stolk.** 1991. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. J. Cell Biol. **115:**587–596.
37. **Ruskin, B., P. D. Zamore, and M. R. Green.** 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell **52:**207–219.
38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1992. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
39. **Sharp, P. A.** 1994. Split genes and RNA splicing. Cell **77:**805–816.
40. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of pre- mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol. **14:**7670–7682.
41. **Stamm, S., M. Q. Zhang, T. G. Marr, and D. M. Helfman.** 1994. A sequence compilation and comparison of exons that are alternatively spliced in neurons. Nucleic Acids Res. **22:**1515–1526.
42. **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.
43. **Tacke, R., and J. L. Manley.** 1995. The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. EMBO J. **14:**3540–3551.
44. **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.
45. **Tian, M., and T. Maniatis.** 1993. A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. Cell **74:**105–114.
46. **Tian, M., and T. Maniatis.** 1994. A splicing enhancer exhibits both constitutive and regulated activities. Genes Dev. **8:**1703–1712.
47. **Valcarcel, J., R. K. Gaur, R. Singh, and M. R. Green.** 1996. Interaction of $U2AF^{65}$ RS region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. Science **273:**1706–1709.
48. **Valcárcel, J., and M. R. Green.** 1996. The SR protein family: Pleiotropic

functions in pre-mRNA splicing. Trends Biochem. Sci. **21:**296–301.
49. **Watakabe, A., K. Tanaka, and Y. Shimura.** 1993. The role of exon sequences in splice site selection. Genes Dev. **7:**407–418.
50. **Will, C. R., B. Kastner, and R. Luhrmann.** 1994. Analysis of ribonucleoprotein interactions, p. 141–177. *In* S. J. Higgins and B. D. Hames (ed.), RNA processing: a practical approach. Oxford University, New York, N.Y.
51. **Wu, J. Y., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75:**1061–1070.
52. **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.
53. **Yeakley, J. M., J.-P. Morfin, M. G. Rosenfeld, and X.-D. Fu.** 1996. A complex of nuclear proteins mediates SR protein binding to a purine-rich splicing enhancer. Proc. Natl. Acad. Sci. USA **93:**7582–7587.
54. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev. **6:**837–847.
55. **Zamore, P. D., and M. R. Green.** 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. USA **86:**9243–9247.
56. **Zheng, Z. M., P. He, and C. C. Baker.** 1996. Selection of the bovine papillomavirus type 1 nucleotide 3225 3′ splice site is regulated through an exonic splicing enhancer and its juxtaposed exonic splicing suppressor. J. Virol. **70:**4691–4699.
57. **Zuo, P., and T. Maniatis.** 1996. The splicing factor U2AF^{35} mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes Dev. **10:**1356–1368.