Role of the Constitutive Splicing Factors U2AF65 and SAP49 in Suboptimal RNA Splicing of Novel Retroviral Mutants*

John Bouck‡‡, Xiang-Dong Fu‡, Anna Marie Skalka‡§, and Richard A. Katz‡**

From the ‡Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, the §Cell and Molecular Biology Graduate Group, University of Pennsylvania, Pennsylvania 19104, and the ¶Division of Cellular and Molecular Medicine, University of California at San Diego, La Jolla, California 92039-0651

Retroviruses display a unique form of alternative splicing in which both spliced and unspliced RNAs accumulate in the cytoplasm. Simple retroviruses, such as avian sarcoma virus, do not encode regulatory proteins that affect splicing; this process is controlled solely through interactions between the viral RNA and the host cell splicing machinery. Previously, we described the selection and characterization of novel avian sarcoma virus mutants. These viruses were separated into two classes based upon analysis of splicing intermediates produced in infected cells and in a cell-free system. One class, which included mutants with altered polypyrimidine tract or branch point sequences, showed significant accumulation of intermediates, suggesting that splicing was regulated in step 2. The other class, which included mutants with deletions of exonic enhancer sequences, did not accumulate splicing intermediates, suggesting that splicing was regulated before step 1 of the splicing reaction. In this report, we show that a mutant blocked at step 1 fails to form a stable spliceosomal complex, whereas one blocked at step 2 shows a defect in its ability to transit through the last spliceosomal complex. Using UV cross-linking methods, we show that regulation at each step is associated with specific changes in the binding of cellular splicing factors. Regulation at step 1 is correlated with decreased cross-linking of the factor U2AF65, whereas regulation at step 2 is correlated with enhanced cross-linking of the factor SAP49. Because these mutations were isolated by selection for replication-competent viruses, we conclude that retroviral splicing may be regulated in vivo through altered binding of constitutive splicing factors.

Pre-mRNA splicing, the process by which introns are removed, takes place in the nucleus within a large structure termed the spliceosome and can be viewed as a two-step process (for recent reviews, see Refs. 1 and 2) (see Fig. 1B). In step 1, the pre-mRNA is cleaved at the 5'-splice site (5'-ss),1 and a branched lariat-exon 2 intermediate is formed. In step 2, the two exons are ligated together, and the branched intron is released. In vertebrates, several loosely conserved sequence elements within the pre-mRNA are important for recognition by the splicing machinery, including sequences at the 5'-ss, the 3'-ss, the polypyrimidine tract (PPyT), and the branch point (BPS) (3). Additional RNA elements have been described within the intron and the exon that can promote or inhibit splicing (reviewed in Refs. 4 and 5).

The RNA splicing elements are recognized by trans-acting cellular components that include both small nuclear ribonucleoproteins (snRNPs) and proteins known as splicing factors (6–8). The snRNPs are composed of a highly structured RNA to which a number of shared as well as unique proteins are bound. The RNA component of some snRNPs recognizes elements in the splicing signals in pre-mRNAs through base pairing interactions. The U2 snRNP recognizes the BPS, and the U1 snRNP recognizes the 5'-ss. Two other snRNPs required for splicing, U4 and U6, participate in formation of the mature spliceosome (6, 9). It is believed that the RNA components of the snRNPs and pre-mRNA cooperate to catalyze both steps in the splicing reaction (10).

The non-snRNP protein components of the spliceosome appear to assist in splice site selection, organization of snRNPs, and regulation of alternative splicing events. A subset of splicing factors contains motifs of repeating serine-arginine peptides, and these factors are called SR or SR-related proteins (11, 12). SR proteins have distinct as well as overlapping functions; one role of SR proteins appears to be in the activation of splicing by exonic splicing enhancers (ESEs). The ESE elements in the downstream exon promote splicing at the upstream 3'-ss usually through interaction with specific SR proteins (13–15).

The limited base pairing potential between small nuclear RNAs and the pre-mRNA splicing signals, as well as the degenerate nature of these signals, suggests that the binding of snRNPs must be assisted by proteins. Several proteins have been implicated in tethering the U2 snRNP to its target element, the BPS. They include the splicing factor U2AF65 (U2-associated factor of 65 kDa), which binds to the PPyT and has RNA annealing activity (16, 17). Some spliceosome-associated proteins (SAPs) can bind to pre-mRNA as well as to small nuclear RNA (8, 18). For example, the U2 snRNP-associated SAP49 and other SAPs have been shown to bind to pre-mRNA upstream of the BPS (19). It appears that the combined binding of U2AF65 and several SAPs in the vicinity of the BPS may stabilize the binding of the U2 snRNP to the pre-mRNA.

The abbreviations used are: ss, splice site(s); PPyT, polypyrimidine tract; BPS, branch point sequence; snRNP, small nuclear ribonucleoprotein; SR, serine-arginine; ESE, exonic splicing enhancer; SAP, spliceosome-associated protein; ASV, avian sarcoma virus; NEPHGE, non-equilibrium pH gradient electrophoresis; TEMED, N,N,N,N'-tetramethyl-ethylenediamine. This paper is available on line at http://www.jbc.org
The molecular basis for the control of alternative splicing in vertebrate cells has been the focus of several recent in vitro studies (20, 21). In some cases, it appears that a protein factor can inhibit splicing by binding to a nearby RNA sequence (20). In other cases, a protein factor can activate splicing by binding near a particular splice site (21, 22). Retroviruses display a unique form of regulated splicing that allows simultaneous expression of spliced and unspliced RNAs within the same cell. The integrated viral DNA produces a single primary transcript that may either be spliced, to allow for expression of the viral envelope glycoproteins, or may remain unspliced, to allow for expression of viral structural proteins and enzymes. In addition to serving as an mRNA, the unspliced RNA also comprises the genome for progeny virions. In the case of retroviruses with the simplest genomes, such as the avian sarcoma virus (ASV), approximately two-thirds of the primary transcripts remain unspliced, and the limited splicing that takes place uses a common 5′-ss and one of two 3′-ss (see Fig. 1). ASV requires an appropriate balance between spliced and unspliced RNAs for efficient replication. We have exploited this feature as a genetic selection tool to study the control of splicing at the env splice site. We demonstrated that a specific insertion into the BPS, upstream of env, increased the efficiency of splicing and created a replication defect (23). Upon prolonged passage, replication-competent viruses that contained spontaneous second-site suppressor mutations were isolated. The suppressor mutations were of two types: one type comprised base substitutions in the BPS, and the other deletions of the ESE within the env exon. Both types of suppressors down-regulated splicing. Up-regulation of splicing and a concomitant replication defect were also created by lengthening of the PPyT (24). In this case, selection for replication-competent mutants revealed a third type of suppressor comprising U to C transitions within the PPyT, which also down-regulated splicing. When the different types of suppressor mutants were examined in more detail, we found evidence for regulation at two distinct steps in the splicing reaction. The first class (mutants that include ESE deletions) showed a partial block in splicing prior to step 1, and the second class (mutants with alterations in the BPS or PPYT) showed a partial block subsequent to step 1, but prior to step 2 (24, 25). In this report, we describe a more detailed in vitro characterization of RNA splicing control and of the protein factors that form complexes with ASV RNAs that include these two classes of splicing suppressor mutations.

EXPERIMENTAL PROCEDURES

Plasmids—The plasmids used to prepare in vitro RNA splicing substrates have been described elsewhere (24).

In Vitro Transcription and Splicing—In vitro transcription and splicing assays were carried out as described (24). Radioactive [32P]-labeled run-off transcripts were synthesized in vitro using T7 polymerase and were then gel-purified for UV cross-linking experiments. For in vitro splicing reactions, RNA (10⁵ cpm) was incubated with 3–10 μl of nuclear extract (~30 μg, prepared as described previously (24)), ATP (0.5 mM), creatine phosphate (20 mM), MgCl₂ (3.2 mM), RNase inhibitor (5 Prime-3 Prime, Inc.) (0.25 unit/reaction), dithiothreitol (2 mM), KCl (100 mM) and 20 mM HEPES, pH 7.6 in a 25 mM volume. Reactions were incubated at 30 °C for the indicated times and stopped by addition of an equal volume of proteinase buffer. One-hundred μg of proteinase K was then added, and the reaction was incubated for 15 min at 37 °C. The RNA products were then purified by phenol/chloroform extraction and ethanol precipitation. Samples were resuspended in formamide loading buffer (50 mM Tris, pH 7.0 for 5 min prior to loading on 6% polyacryl- amide gel containing 7 M urea. The gels were dried and subjected to autoradiography.

Splicing Complex Formation Assays—Conditions for gel shift assays have been described (26). In vitro splicing reactions were incubated at 30 °C for the indicated times, after which they were brought to a final concentration of 0.5 mg/ml heparin and incubated for 5 min at 30 °C. Samples were then loaded directly onto a 4% polyacrylamide gel (running and gel buffer was 50 mM Tris, pH 8.0, and 50 mM glycine) and subjected to electrophoresis at 200 V for 5 h. The gels were then dried and subjected to autoradiography. Because it is unusual to detect significant amounts of complex C in splicing reactions, we analyzed our most efficiently spliced substrate CG-uu, in a similar manner (data not shown). As we detected amounts of complex C similar to that observed with I(−17), we attribute the increased detection of this complex to some general feature of our system.

UV Cross-linking—For UV cross-linking, the splicing reactions were incubated for the indicated times and then placed on ice in an inverted Eppendorf tube top. UV exposure was for 5 min at 10 cm from the UV lamp (UV-GT6S lamp). After UV exposure, the samples were treated for 15 min with an RNase mixture (Ambion Inc.) that contained RNases A and T1. An equal volume of 2× Laemmli loading buffer was added to the samples, which were heated to 95 °C for 5 min and then loaded onto 10% SDS protein gels. The gels were then fixed, dried, and subjected to autoradiography.

Two-dimensional Gels—Non-equilbrium pH gradient gel electrophoresis (NEPHGE) was carried out as described (27). Cross-linking reactions were treated with RNase, incubated with an equal volume of sample buffer (9.5 M urea, 2% Nonidet P-40, 5% 2-mercaptoethanol, 1.6% ampholyte, pH 5–8, and 0.4% ampholyte, pH 3–10) at 65 °C for 10 min, and then layered on top of a tube gel (9.2 M urea, 4% acrylamide (10:1 acrylamide/bisacrylamide ratio), 20% Nonidet P-40, 1.6% am- pholyte, pH 5–8, 0.4% ampholyte, pH 3–10, 0.01% ammonium persulfate, and 0.1% TEMED). A solution of 100 mM NaOH was placed in the lower chamber, and 10 mM H3PO4 was placed in the upper chamber. Samples were subjected to electrophoresis at 500 V for 1 h and then at 750 V for 2 h. Each tube gel was then transferred to the top of a 10% SDS protein gel without a stacking gel. The protein gels were run until the dye reached the bottom, at which time they were transferred to nitrocellulose for blotting. To determine the location of cross-linked species and the location of specific proteins in the nuclear extract on the gels, the membrane was subjected to chemiluminescence autoradiography (Pierce). The membrane was then washed overnight in a large volume of phosphate-buffered saline to remove chemiluminescence re- agents and exposed to x-ray film to detect the location of [32P]-labeled cross-linked species.

Production of Anti-SAP49 Antibodies and Immunoblotting—An anti-SAP49 rat serum was prepared using a fragment of SAP49 comprising the RNA-binding domains (amino acids 1–180) (the SAP49 cDNA was kindly provided by Robin Reed (28)). The pET28 system (Novagen) was used for Escherichia coli expression and purification through histidine tags at both termini. The bacteria were lysed with a French press, and the recombinant protein was purified by TALON metal chelate chromatography (CLONTECH). Polyclonal antibodies were raised in rats by the Laboratory Animal Facility at the Fox Chase Cancer Center. Immunoblotting and chemiluminescence detection were performed according to the recommendations of the manufacturer (Pierce).

RESULTS

Substrates and Experimental Design—Fig. 1A shows the organization of the ASV genome and the location of RNA elements involved in splicing and the minigene RNA substrates used in this study. The nucleotide sequences of the regions spanning the BPS to the env 3′-ss for RNAs analyzed in this study are shown in Fig. 1C. Two mutations, I(−17) and IS1cu (24), were constructed by site-directed mutagenesis. Both increase splicing efficiency and cause a replication defect. The suppressor mutations listed were obtained by selecting for replication-competent viruses in cultures transfected with viral DNAs that contain the I(−17) or IS1cu mutations (Fig. 1C). These suppressor mutations all down-regulate splicing. Several of the suppressor mutations cause a partial block in step 2 of the splicing reaction detected by the accumulation of lariat-exon 2 intermediates in both cell-free splicing extracts and in cells infected with the mutant viruses (Fig. 1C) (24). An additional substrate used in our studies is not described in Fig. 1, but has been described elsewhere (24). This substrate replaces the weak wild type BPS with the strong β-globin BPS and adds a three-uridine extension onto the PPyT and it is thus denoted BG-uu. The splicing of BG-uu is very efficient compared with
the wild type, and it was therefore used as a positive control when appropriate.

**Suppressor Mutations Affect the Formation of Splicing Complexes**—We were interested in determining if suppressor mutants with partial blocks at either step in splicing could be distinguished from each other, or from an efficiently splicing RNA substrate, in their ability to form splicing-specific complexes. For these analyses, we chose the parental, oversplicing mutant I(−17) and the IΔ51 and I51 suppressor derivatives that cause partial blocks in splicing before step 1 and at step 2, respectively (24, 25). As illustrated in the *in vitro* splicing assay shown in Fig. 2A, the lariat-exon 2 intermediates accumulated with I51, but not with IΔ51, and spliced product was detectable only with the I(−17) parent. The association of a substrate RNA with the splicing machinery is known to occur through a stepwise process; several stages in this process can be visualized on a native gel. In the initial stage, the RNA is found in a heterogeneous, but fast migrating complex, designated H (26, 29), which contains mainly heterogeneous nuclear ribonucleoproteins (30). After the H complex, RNAs can be found in splicing complexes, designated A, B, and C in order of increasing size and time of appearance. Step 2 splicing takes place in complex C, after which the spliced RNA is released, and the spliceosome disassembles (31). This complex is usually short-lived and often difficult to detect.

Using the standard mobility shift assay, we found most of the RNA of the efficiently spliced substrate I(−17) in the fast migrating H form at the earliest time point (15 min), with only a small amount of complex A visible. With increasing time, the complex increased in size, first accumulating at the position of complex B and then complex C. Similar kinetics were observed with step 2-blocked IS1 RNA, despite the fact that most of the RNA remained unspliced (Fig. 2A). Quantitation of complex C formation is shown in Fig. 2C. The efficiently spliced substrate I(−17) formed increasing amounts of complex C for the first 45 min. Thereafter, the amount remained relatively constant, despite the continued accumulation of spliced product. We interpret this plateau to reflect a steady state in which new C complexes are formed at the same rate as others are disassembled, following completion of the splicing reaction. The I51 substrate was initially assembled into complex C at the same rate as I(−17). However, with I51, a plateau was not observed. Instead, the amount of complex C increased linearly throughout the course of the reaction. This would be expected if the I51 substrate is as efficient as I(−17) for step 1 of the splicing reaction, but is inhibited from completing step 2.

As shown in Fig. 2B (lanes 11–15), formation of splicing complexes was very inefficient with IΔ51 RNA, and the limited amount of complex H formed disappeared quickly. This result is not unexpected as we have shown by sedimentation analysis that RNA substrates with suppressor mutations that are ESE deletions (I36 and IΔ51) do not form spliceosomes efficiently, but rather accumulate smaller complexes (25). Thus, the complex formation data are consistent with a model in which splicing is regulated very early in the pathway by the ESE deletion suppressors and later with the step 2 suppressors.

![Diagram](image_url)

**Fig. 1. RNA substrates and mutations.** A: top, shown is a schematic diagram of the genome of ASV. Viral genes are indicated (gag, pol, env, and src) as well as elements previously demonstrated to be important for splicing: 5′-ss, negative regulator of splicing (NRS), 3′-ss, ESE, and src splicing silencer (SSS). Bottom, the minigene construct used for *in vitro* studies is 513 nucleotides long and contains all elements near the 3′-ss, including the ESE. B: shown is a schematic diagram of the two chemical steps that take place in pre-mRNA splicing. C: the nucleotide sequence spanning the region from the BPS to the PPyT to the 3′-ss is shown. The mutations that activate splicing are indicated in boldface, and the genetically isolated suppressor mutations are underlined. I36 and IΔ51 contain suppressor mutations that correspond to deletions of 36 and 51 nucleotides, respectively, within the exonic splicing enhancer (23). Accordingly, the absence of the ESE sequences is indicated. The *black dot* indicates the location of the branch point adenosine used in lariat formation for most of the substrates. The relative efficiency of viral growth in tissue culture is indicated (*Growth*). The step of the splicing reaction that is blocked is indicated (*Step Blocked*); scoring is based upon the ability to detect stable lariat-exon 2 intermediates *in vivo* (24). Viruses that do not replicate in tissue culture could not be tested for the production of lariat-exon 2 intermediates *in vivo* (not tested (*nt)).
to UV light (32). Following digestion with RNase, proteins that contained cross-linked [32P]RNA adducts were analyzed by polyacrylamide gel electrophoresis followed by autoradiography. We first examined the cross-linked products from RNA substrates containing the ESE deletion suppressor mutations (Fig. 3B). Typical patterns from in vitro splicing reactions with these RNA substrates are shown in Fig. 3A. The block in splicing was less severe with the IA51 suppressor compared with the IA51 suppressor. When normalized for total radioactive content, only an -2-fold reduction in splicing was typically observed between IA36 and the I−(17) parent. This is consistent with our previous observation that IA36 only partially suppresses I−(17) oversplicing in vivo (23).

In the gel analysis of cross-linked proteins (Fig. 3B and C), we observed differential cross-linking of a 65-kDa protein. Substrates corresponding to the I−(17) oversplicing parental mutant and the step 2-blocked IS1 RNA substrates exhibited stronger cross-linking to the 65-kDa protein, whereas step 1-blocked RNAs (IA36 and IA51) showed weaker cross-linking to this protein. The reduced cross-linking of the 65-kDa protein to the ESE-deleted substrates was observed in numerous experiments and was reproduced in all nuclear extracts tested. This 65-kDa protein was identified as U2AF65 based on the following criteria. (i) The cross-linked species comigrated with U2AF65. The cross-linking gel shown in Fig. 3C (left panel) was transferred to a nitrocellulose filter and subjected to autoradiography to identify the cross-linked species. Subsequently, the filter was immunoblotted with an antibody that recognizes U2AF65 (kindly provided by Dr. James Patton). The band detected by immunoblotting comigrated with the [32P]-labeled, UV-cross-linked band (indicated by an arrow). (ii) The 65-kDa species cross-linked extremely well to the most efficiently spliced substrate, BG-uuu (Fig. 4, lane 1). The U2AF65 protein has been shown to strongly cross-link other efficiently spliced substrates (33). (iii) Cross-linking to the 65-kDa species was observed in the absence of exogenous ATP (data not shown). This is consistent with the finding that U2AF65 binding to RNA substrates does not require ATP (33). (iv) Migration of the cross-linked species in NEPHGE was similar to that of U2AF65 (data not shown). (v) Cross-linking of the 65-kDa species was affected by mutations in the PyP'T that have been shown previously to affect the cross-linking of U2AF65 (see below).

The observation that cross-linking of U2AF65 to the efficiently spliced substrate I−(17) is stronger than to substrates...
containing the ESE deletions (IΔ51 and IΔ36) is consistent with models in which exon enhancers act indirectly to promote U2AF65 binding to the PPyT (see "Discussion"). The IS1 substrate, which is blocked at step 2, cross-links U2AF65 efficiently. This is consistent with the hypothesis that U2AF65 is required very early in the splicing reaction. Because these suppressors were obtained by selection of replication-competent viruses, these data support a model of exon enhancer function in vivo that involves the splicing factor U2AF.

Substrates That Are Blocked at Step 2 in Splicing Show Enhanced Cross-linking of a 50-kDa Protein—

We next asked if a partial block at step 2 is also correlated with differential binding of cellular proteins. In vitro splicing patterns for the seven RNA substrates used in these analyses are shown in Fig. 4A. As expected, the BG-uuu substrate was spliced most efficiently, and no splicing was detected with the wild type substrate (24). Results with I(−17) and IS1 were the same as those shown in Fig. 3A. The substrate IS1cu, with an extended PPyT, was spliced more efficiently than IS1, from which it was derived (Fig. 4A, compare lanes 5 and 6). Splicing patterns of RNA substrates derived from replication-competent viruses that contain additional mutations that suppress IS1cu (IS1cuU(−9,−15)C and IS1cuU(−15)C) are also shown (Fig. 4A, lanes 7 and 8). As can be seen by comparing the patterns, the two suppressors of IS1cu restored a step 2 block in splicing similar to that seen in IS1 (Fig. 4A, compare lanes 5, 7, and 8). These results are consistent with previous results from in vivo analyses (24).

Fig. 4B shows the cross-linking pattern of the substrates shown in Fig. 4A. U2AF65 cross-linked to the BG-uuu substrate very efficiently (upper arrow). Cross-linking to U2AF65 was also affected in some of the other substrates that have alterations in their PPyT. For example, IS1cuU(−9,−15)C cross-linked U2AF65 somewhat less efficiently than IS1cuU(−15)C (Fig. 4B, compare lanes 6 and 7). We noted (Fig. 1C) that the −9 mutation reduced a five-uridine stretch to four. A five-uridine stretch in the PPyT has been shown to be important for efficient splicing as well as cross-linking of U2AF65 (34, 35). In addition to differences in the efficiency of U2AF65 cross-linking, several substrates showed relatively enhanced cross-linking to a 50-kDa protein (Fig. 4B, lanes 4, 6, and 7).

As enhanced cross-linking of the 50-kDa protein to the IS1cuU(−9,−15)C substrate was also observed in the absence of ATP (data not shown), we reasoned that slowing the rate of the splicing reaction might reveal significant differences in cross-linking with different substrates. Accordingly, the concentrations of ATP and creatine phosphate were reduced by one-half, to 250 μM and 10 mM, respectively, and the incubation time was reduced to 10 min. Under these conditions, splicing occurs, but at a significantly reduced rate (data not shown). An example of the cross-linking pattern obtained under these conditions is shown in Fig. 5. The results indicate that the 50-kDa...
protein is the predominant cross-linked species for a number of substrates, and U2AF 65 cross-linking is reduced greatly. Furthermore, substrates that were blocked at step 2, IS1, IS1cuU(−15)C, and IS1cuU(−9,−15)C (lanes 4, 6, and 7), showed enhanced cross-linking to the 50-kDa protein. In contrast, viral RNA substrates that were spliced efficiently or were blocked at step 1 showed less cross-linking (BG-uuu, wild type, I(−17), and IS1cu (lanes 1–3 and 5)). We next examined the kinetics of cross-linking under the conditions of reduced ATP and creatine phosphate. A comparison of the efficiently spliced substrate I(−17) and its step 2-blocked suppressor, IS1, showed that cross-linking of the 50-kDa protein to the IS1 suppressor was strongest at the earliest time point (Fig. 6A, lane 8) and decreased slowly at later time points. Quantitation by phosphoimaging is shown in Fig. 6A (lower panel). Similar comparisons were made with the IS1cuU(−15)C suppressor and its oversplicing parent, IS1cu (Fig. 6B). These results also demonstrate an increased cross-linking of the 50-kDa protein to the suppressor at the earlier time points. Because reactions proceeded at a slower than normal rate under these conditions and the difference in 50-kDa protein cross-linking was observed at the earliest time point, this suggests that the enhanced binding of the 50-kDa protein to the suppressors occurs very early in the splicing reaction.

In the experiments presented above, RNA substrates were uniformly labeled using [α-32P]UTP. In other experiments (data not shown), RNA substrates were double-labeled with [α-32P]UTP + [α-32P]CTP. As similar results were obtained with the single- or double-labeled substrates, we conclude that the differential detection of cross-linked proteins is not due to a labeling bias. Substrates labeled with [α-32P]GTP produced a significantly simpler cross-linking pattern in which no differences were apparent among any of the substrates tested (data not shown).

Identification of the 50-kDa Protein as SAP49—Based on its predicted molecular mass and the observed migration of the 50-kDa cross-linked species on two-dimensional gels, we suspected that this protein might correspond to the splicing factor SAP49 (36). We employed NEPHGE to determine if the 50-kDa cross-linked species comigrated with SAP49. We raised polyclonal antibodies to the bacterially expressed N-terminal domain of SAP49. Serum from one rat contained highly specific antibodies that recognized a single species in one- and two-dimensional Western analyses (two-dimensional analysis shown in Fig. 7, upper panels). An in vitro cross-linking reaction was performed with a novel PPyT mutant RNA substrate that displayed a step 2 block and relatively strong cross-linking to the 50-kDa protein (data not shown) and with a wild type RNA substrate as a control. Following digestion with RNase, proteins were fractionated on a two-dimensional NEPHGE gel. The proteins were subsequently transferred to nitrocellulose, and the membrane was probed with our anti-SAP49 antibody and detected by chemiluminescence (upper panels). The same membrane was subjected to autoradiography to identify the cross-linked species (lower panels). The upper panels show that the anti-SAP49 antibody recognized a single protein on the NEPHGE gels. Two major cross-linked species were detectable.
on the gels shown in the lower panels, as expected from the one-dimensional gels (Fig. 6). Alignment of the two films revealed that the 50-kDa cross-linked species comigrated with SAP49 (Fig. 7, lower panels, closed arrowheads). We conclude that the 50-kDa protein corresponds to SAP49. In a similar NEPHGE experiment, Staknis and Reed (39) found that cross-linked SAP49 demonstrated a slightly slower and more acidic migration than silver-stained SAP49 from nuclear extracts, an alteration attributed to the RNA adducts. In our experiments, RNA nucleotide adducts did not significantly affect the protein migration on the NEPHGE gels. We attribute this to technical differences in our protocol.

The lower panels of Fig. 7 contain a second spot (open arrowheads) that corresponds to a cross-linked species of ~30 kDa and that was found in all substrates tested (see Fig. 6). The 30-kDa species serves as an internal standard as its intensity was similar in analyses with all the substrates. A comparison of the intensities of the two spots in the analysis of Fig. 7 verified that there was relatively stronger cross-linking of SAP49 to the step 2-blocked RNA substrate than to the wild type substrate. Other observations also support the conclusion that the 50-kDa protein is SAP49. First, the 50-kDa protein cross-linked to RNA at an early stage in the splicing reaction, which is in agreement with the observation that SAP49 is found primarily in early splicing complexes (36). Second, a known characteristic of SAP49 is its ability to cross-link to an RNA substrate under splicing conditions (28), which is consistent with our results. From our experiments, we conclude that SAP49 is likely to be involved in the establishment of the block in splicing at step 2 that we observed both in vitro and in vivo.

**DISCUSSION**

In this report, we have examined virus-host interactions through in vitro studies of RNA splicing substrates derived from retroviral mutants selected for their ability to restore viral growth. We have shown previously that the relative splicing efficiencies observed in vitro can be recapitulated in vivo. Here we have demonstrated that one mechanism of regulation, which involves accumulation of splicing intermediates in vivo, is characterized by the accumulation of spliced complexes in vitro. We further characterized the substrates by examining the proteins that are in contact with the RNA. Differential UV cross-linking of two proteins, U2Af65 and SAP49, is described. We also demonstrated that the suppressor mutations comprising ESE deletions result in diminished cross-linking of U2Af65. This supports the findings of others that exonic enhancer elements act indirectly to promote U2Af65 binding to the PPY7 (13, 37–39). We also observed a decrease in U2Af65 cross-linking in the case of a natural suppressor mutation that converts a five-uridine stretch in the PPY to four (Fig. 4B, lanes 6 and 7). This is consistent with in vitro mutagenesis studies (34, 35). With the substrates tested, enhanced cross-linking of SAP49 correlated with accumulation of spliceosomal complex C, indicative of a block at step 2. Furthermore, this increased cross-linking of SAP49 occurred early in the splicing reaction, suggesting that this protein is involved in the early establishment of the block, but that its effect is manifested later in the splicing reaction. Because the suppressor mutations arose naturally during viral replication, the differences in factor binding observed in vitro are likely to contribute to the regulation that was selected for in vivo.

All of the substrates tested that were blocked in step 2 demonstrated increased cross-linking of SAP49. One potential explanation is that this factor is a component of C complexes, and the increased cross-linking merely reflects the accumulation of this complex. We addressed this possibility by examining cross-linking profiles at early time points in the splicing reaction (Fig. 6). Under the conditions used, we found increased cross-linking of SAP49 at the earliest time tested. This time point was prior to detection of any splicing complexes, which indicates that the enhanced association of SAP49 with the substrate occurred before accumulation of C complexes. This suggests that the increased binding of this protein is likely to be involved in the establishment of the block at step 2, as opposed to being a consequence of the block.

The SAP49 protein is a component of the essential splicing factor SF3b and is associated with the U2 small nuclear RNA (8, 40). SAP49 has been shown to cross-link efficiently to RNA substrates in complexes A and B and also to bind both U2 snRNP and the pre-mRNA (36, 41). It has been proposed that SAP49 is involved in tethering the U2 snRNP to the BPS (28). How can it be that a factor involved in constitutive splicing is also involved in the establishment of a block within the splicing pathway? We speculate that SAP49 may bind to the mutant substrates in an inappropriate manner and that this causes the spliceosome to stall after step 1. This interpretation is supported by the following observations. First, UV cross-linking experiments reported in the literature indicate that SAP49 normally binds to RNA substrates only in the presence of ATP. However, one of our step 2-blocked RNA substrates, (IS1cuU(−9,−15)C) can be cross-linked to SAP49 in the absence of exogenous ATP (data not shown). Second, suppressor mutations that affected SAP49 binding in our studies com-
prised U to C substitutions in the PPyT (IS1cuU(−15)C and IS1cuU(−9, −15)C). Previous studies indicate that SAP49 normally binds to RNA substrates upstream of the BPS (28). Although we have not yet mapped SAP49 cross-linking sites on the suppressor RNA substrates, we had not expected that changes in the PPyT would affect its cross-linking. The fact that they do suggests that SAP49 may bind inappropriately to these substrates and that such binding may contribute to the block in step 2.

Another protein that demonstrated variable cross-linking was identified as U2AF65. This essential splicing factor has been shown to bind to the PPyT of pre-mRNAs early in the splicing reaction (33). U2AF65 also contains RNA annealing activity (17) and has been shown to contact the BPS of the pre-mRNA (16). These results suggest that U2AF65 binds to pre-mRNA early and promotes the association of U2 snRNP with the BPS. Thus, both U2AF65 and SAP49, along with other proteins, act to tether the U2 snRNP to the BPS (19). Generally, our data are consistent with this model for U2AF65 and SAP49 functions. However, we did not find an absolute correlation between increased U2AF65 cross-linking and efficiency at step 1. Splicing of one of the substrates tested, IS1cuU(−9, −15)C (Fig. 4), was fairly efficient at step 1, but did not cross-link U2AF65 well. In our system, the substrates that demonstrated low U2AF65 cross-linking and that were efficient for step 1 of splicing also showed significant cross-linking to SAP49. This suggests that SAP49 and U2AF65 may have some complementary functions in step 1 of splicing. We noted that cross-linking of U2AF65 was less efficient than SAP49 under reduced ATP conditions, indicating that SAP49 may bind prior to U2AF65 in our system. Experiments are underway to address this issue directly.

Retroviruses display a unique form of alternative splicing in which both spliced and unspliced RNAs accumulate in the cytoplasm. Here we have performed in vitro analyses on several sets of mutant and suppressor viruses in which the ratio of spliced to unspliced RNA was altered. One set demonstrated decreased splicing efficiency that correlated with decreased association with U2AF65. A second set demonstrated decreased splicing efficiency that correlated with increased association with SAP49. This result suggests that splicing efficiency can be decreased through increased binding of constitutive splicing factors.

Acknowledgments—We thank Drs. Robin Reed and James Patton for providing useful reagents and Sharon Jamison and the Laboratory Animal Facility of the Fox Chase Cancer Center for technical support. HeLa cell pellets were provided by the Cell Culture Center and National Center for Research Resources. We also thank Drs. John Taylor and Robert Perry for helpful comments on this manuscript.

REFERENCES
1. Moore, M. J., Query, C. C., and Sharp, P. A. (1993) in Splicing of Precursors to Messenger RNAs by the Spliceosome (Getzeland, R. F., and Atkins, J. F., eds) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
2. Adams, M. D., Rudner, D. Z., and Rio, D. C. (1996) Curr. Opin. Cell Biol. 8, 331–339
3. Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990) Methods Enzymol. 183, 252–278
4. Chabot, B. (1996) Trends Genet. 12, 472–478
5. Inoue, K., Ono, M., and Shimura, Y. (1995) Gene Expr. 4, 177–182
6. McKeown, M. (1995) Curr. Opin. Cell Biol. 5, 448–454
7. Nilsen, T. W. (1994) Cell 78, 1–4
8. Kramer, A. (1996) Annu. Rev. Biochem. 65, 367–409
9. Newman, A. (1998) Curr. Biol. 48, 462–464
10. Lamond, A. I. (1993) Bioessays 15, 595–603
11. Fu, X. D. (1995) RNA 1, 663–680
12. Valcarcel, J., and Green, M. R. (1996) Trends Biochem. Sci. 21, 296–301
13. Ares, M. J., and Weiser, B. (1995) Science 268, 13472–13478
14. Reed, R. (1996) Curr. Opin. Genet. Dev. 6, 215–220
15. McKeown, M. (1995) Curr. Opin. Cell Biol. 5, 448–454
16. Valcarcel, J., Gaur, R. K., Singh, R., and Green, M. R. (1996) Science 273, 1706–1709
17. Lee, C. G., Zamore, P. D., Green, M. R., and Hurwitz, J. (1993) J. Biol. Chem. 268, 1023–1036
18. Ashiya, M., and Gribowski, P. J. (1997) RNA 3, 996–1012
19. Min, H., Turck, C. W., Nikolic, J. M., and Black, D. L. (1997) Genes Dev. 11, 1023–1036
20. Min, H., Chan, R. C., and Black, D. L. (1995) Genes Dev. 9, 2659–2671
21. Katz, R. A., and Skalka, A. M. (1990) Mol. Cell. Biol. 10, 696–704
22. Bouck, J., Fu, X. D., Skalka, A. M., and Katz, R. A. (1995) Mol. Cell. Biol. 15, 2663–2671
23. Lee, C. G., Zamore, P. D., Green, M. R., and Hurwitz, J. (1993) J. Biol. Chem. 268, 595–603
24. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 87, 9243–9247
25. Reed, R. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
26. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 92, 4379–4384
27. Reed, R. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8031–8035
28. Reed, R. (1996) Curr. Opin. Genet. Dev. 6, 215–220
29. Reed, R. (1996) Proc. Natl. Acad. Sci. U. S. A. 87, 8031–8035
30. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
31. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 92, 4379–4384
32. Reed, R. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
33. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
34. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
35. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
36. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
37. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
38. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
39. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
40. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035
41. Reed, R. (1999) Proc. Natl. Acad. Sci. U. S. A. 93, 5030–5035