Two Strong 5' Splice Sites and Competing, Suboptimal 3' Splice Sites Involved in Alternative Splicing of Human Immunodeficiency Virus Type 1 RNA

MARINA M. O'REILLY,* MARK T. McNALLY,† and KAREN L. BEEMON*,[1]

*Department of Biology, Johns Hopkins University, Baltimore, Maryland 21218; and †Department of Microbiology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226

The human immunodeficiency virus type 1 (HIV-1) genome contains 20 exons that are alternatively spliced from 16 splice sites to generate more than 40 different mRNAs, including incompletely spliced and unspliced mRNAs. In contrast to avian retroviral RNA, which has a *cis*-acting element in *gag* that negatively regulates splicing (NRS), HIV-1 RNA did not have any NRS sequences in the *gag* or *pol* genes detectable by a splicing inhibition assay. However, this assay demonstrated that the HIV-1 first 5' splice site competed with a cellular 5' splice site, suggesting that HIV-1 may have some strong splice sites. To extend this observation, we used a splice site swapping strategy to determine the efficiency of 14 HIV-1 splice sites in human **β** globin chimeras tested in transient transfection experiments. While the 1st HIV-1 5' splice site used in all spliced transcripts and the 4th 5' splice site used in most of the 2-kb transcripts were efficient, the other splice sites, including all the 3' splice sites, were less efficient, ranging in use from 25 to 60%. We propose that this range of splice site efficiencies contributes to the regulation of alternative splicing of HIV-1 mRNAs.   © 1995 Academic Press, Inc.

INTRODUCTION

Retroviral RNA is transcribed from a provirus integrated into the host chromosome and processed by cellular machinery (reviewed in Coffin, 1990). The capped and polyadenylated retroviral transcript is alternatively and incompletely spliced. Mutations that increase splicing decrease viral infectivity, suggesting that a balance of unspliced and spliced transcripts is essential for retroviral replication (Katz *et al.*, 1988; Zhang and Stoltzfus, 1995). The complex genome of human immunodeficiency virus type 1 (HIV-1) (reviewed in Cullen, 1991; Felber and Pavlakis, 1993) includes 20 exons, bounded by 5 5' and 11 3' splice sites, which are alternatively spliced to yield more than 40 transcripts (Felber *et al.*, 1990; Muesing *et al.*, 1985; Purcell and Martin, 1993; Robert-Guroff *et al.*, 1990; Schwartz *et al.*, 1990). As in the simpler retroviruses, the ~9-kb primary transcript serves as genomic RNA for progeny virions and as mRNA to generate the Gag and Pol proteins. The Env, Vif, Vpr, and Vpu proteins are encoded by an ~4-kb class of partially spliced mRNAs (Garrett *et al.*, 1991; Schwartz *et al.*, 1990), and a multiply-spliced, ~2-kb class of mRNAs encodes the regulatory proteins Tat, Rev, and Nef (Feinberg *et al.*, 1986; Felber *et al.*, 1990; Kim *et al.*, 1989).

Maintenance of the proper balance of retroviral mRNAs requires mechanisms that prevent the RNA from splicing to completion while still permitting its capping,

polyadenylation, transport to the cytoplasm, and translation. In the case of the simple avian retroviruses, such as Rous sarcoma virus (RSV), both the *env* and the *src* 3' splice sites are suboptimal, but the single 5' splice site is strong (Berberich and Stoltzfus, 1991; Fu *et al.*, 1991; Katz and Skalka, 1990; McNally and Beemon, 1992; Zhang and Stoltzfus, 1995). Different mechanisms seem to be involved in inhibition of splicing at the 3' sites. The RSV *env* pre-mRNA has a suboptimal branchpoint site located at −16 nt relative to the 3' splice site and a suboptimal pyrimidine tract (Fu *et al.*, 1991; Katz *et al.*, 1988; Katz and Skalka, 1990; Bouck *et al.*, 1995). The *src* 3' splice site also has a nonconsensus pyrimidine tract (Zhang and Stoltzfus, 1995). In addition, a *src* suppressor of splicing has been identified more than 70 nucleotides upstream of the 3' splice site (Berberich and Stoltzfus, 1991; McNally and Beemon, 1992). Splicing is also controlled in RSV by a *cis*-acting RNA sequence in the *gag* gene termed the negative regulator of splicing (NRS), which decreases splicing at both 3' splice sites (Arrigo and Beemon, 1988; Stoltzfus and Fogarty, 1989). Insertion of the NRS into heterologous introns inhibits splicing in an orientation- and position-dependent manner *in vivo* and *in vitro* (Arrigo and Beemon, 1988; McNally *et al.*, 1991; Gontarek *et al.*, 1993). Thus, multiple *cis*-acting sequences and suboptimal 3' splice sites act additively to control splicing of RSV RNA (McNally and Beemon, 1992).

An additional level of complexity exists in HIV where unspliced and incompletely spliced RNA is not expressed in the cytoplasm in the absence of a viral-encoded regulatory protein, Rev (Emerman *et al.*, 1989;

Felber *et al.*, 1989; Hammarskjold *et al.*, 1989; Malim and Cullen, 1993; Malim *et al.*, 1989). Rev binds to a highly structured RNA sequence, the Rev response element (RRE), located in the *env* gene region that is retained in both the unspliced and the partially spliced (4 kb) HIV-1 transcripts (Cochrane *et al.*, 1989; Hadzopoulou-Cladaras *et al.*, 1989; Hauber *et al.*, 1988; Malim *et al.*, 1989; Rosen *et al.*, 1988). Rev regulation requires either inefficient splicing (Chang and Sharp, 1989; Stutz and Rosbash, 1994; Lu *et al.*, 1990; Hammarskjold *et al.*, 1994); elements within *gag, pol,* and *env* genes termed *cis*-acting repressive sequences (CRS) (Rosen *et al.*, 1988); inhibitory regions (IR) (Maldarelli *et al.*, 1991); or inhibitory sequences (Schwartz *et al.*, 1992) or both inefficient splicing and *cis*-acting repressive sequences. While some workers observe untranslated CRS-containing RNA in the cytoplasm in the absence of Rev (Arrigo and Chen, 1991; Cochrane *et al.*, 1991; D'Agostino *et al.*, 1992), others found that these RNAs were restricted to the nucleus (Cochrane *et al.*, 1991; Emerman *et al.*, 1989; Hadzopoulou-Cladaras *et al.*, 1989; Hammarskjold *et al.*, 1989; Maldarelli *et al.*, 1991; Malim and Cullen, 1993).

To investigate the regulation of alternative splicing in HIV-1, we first asked whether *cis*-acting sequences analogous to the RSV NRS element were present in HIV-1 RNA. Sequences in the *gag* and *pol* genes of HIV-1 were tested for splicing inhibition using a heterologous c-*myc* intron splicing assay previously employed to characterize the RSV NRS element (Arrigo and Beemon, 1988; McNally *et al.*, 1991). An HIV-1 *gag* gene fragment containing a CRS element (IR-1) appeared to inhibit splicing as well as the NRS; however, the sequences responsible for splicing inhibition by IR-1 mapped to a small 5′ fragment containing the first HIV-1 5′ splice site. Unlike the RSV NRS which inhibits splicing of the intron it occupies, the HIV-1 first 5′ splice site within IR-1 competed for splicing with the *myc* 5′ splice site.

Having found that the first HIV-1 splice site was capable of competing with a cellular splice site, we proceeded to investigate the efficiency of the other HIV-1 splice sites. Previous studies of the efficiency of HIV-1 splice sites have involved only a fraction of the 16 identified HIV-1 splice sites. Several groups have studied the pair of splice sites involved in removal of the second intron of the 2-kb *tat/rev* mRNAs (Chang and Sharp, 1989; Krainer *et al.*, 1990; Staffa and Cochrane, 1994), as well as the cluster of alternative 3′ splice sites used to generate *tat, rev, nef,* or *vpu/env* mRNAs (Amendt *et al.*, 1994; Riggs *et al.*, 1994). However, no global functional comparisons have been made between all of the HIV-1 splice sites.

In this study, we investigated the control of alternative splicing in HIV-1, using a splice site swapping strategy to compare the efficiency of 14 HIV-1 splice sites in chimeric substrates. Individual 5′ and 3′ HIV-1 splice sites and flanking sequences, as well as two clusters of 3′ splice sites, were tested *in vivo* in combination with effi-

cient splice sites from human β globin. A fragment bearing the 5′ or 3′ splice site was deleted from constructs containing the human β globin first intron and replaced with an HIV-1 fragment bearing a 5′ or 3′ splice site. Nine chimeric plasmids containing 14 HIV-1 splice sites were transfected into 293 cells, and splicing was assayed by RNase protection. We found that the 1st 5′ splice site used in all spliced transcripts and the 4th 5′ splice site used in all but one of the 2-kb transcripts were efficiently used; however, the other splice sites were not. Thus, a range of splice site efficiencies was observed; we propose that this contributes to splicing regulation of HIV-1 mRNAs.

MATERIALS AND METHODS

Plasmid constructs

Sequences from the *gag* and *pol* regions of HIV-1 BH10 (Ratner *et al.*, 1986) were tested for their ability to inhibit splicing using a heterologous assay developed to study splicing inhibition by the RSV NRS (Arrigo and Beemon, 1988; McNally *et al.*, 1991). HIV-1 fragments containing IR-1 (BH10 nt 709–2003, where 1 corresponds to the beginning of the U3 long terminal repeat (LTR) sequence), IR-2 (BH10 nt 2008–3859), and CP490 (BH10 nt 4192–4685) were cloned into a *Bst*XI site within the *myc* intron of pRSVNeo-int (Linial, 1987) to generate pNeomycIR1, pNeomycIR2, and pNeomycCP490. Additional constructs were made using subfragments of IR-1: pNeomyc3′IR1, (BH10 nt 840–2003) and pNeomyc5′IR1(BH10 nt 708–840). The NRS was contained in an *Mro*I–*Sph*I (RSV nt 706–1006) fragment in pNeomyc(B)MS (McNally *et al.*, 1991).

pRSV2Hβ was constructed by insertion of the *Hin*dIII–*Bam*HI fragment of plasmid SP64 108 containing the first and second exons and the intervening intron of human β globin (Reed, 1989; kindly provided by R. Reed) into the *Hin*dIII and *Bam*HI sites of pRSV2 (McNally *et al.*, 1991). Transcription of this plasmid initiates in the RSV LTR and is terminated by the simian virus 40 (SV40) early poly(A) site.

pHIVHβ 5′ splice site plasmids. The *Nco*I–*Xho*I region of pRSV2Hβ was deleted to remove the human β globin 5′ splice site, together with 94 nt of exon 1 and 77 nt of adjacent intron sequence. Fragments of HIV-1 BH10 sequence containing 5′ splice sites were excised from plasmids pSVGagPol-RRE-R, pSVSX1, or pBH10-flop (Smith *et al.*, 1990; Hammarskjold *et al.*, 1989; Rekosh *et al.*, 1988; all kindly provided by M. L. Hammarskjold). After the ends of these fragments were repaired, they were inserted into pRSV2Hβ Δ*Nco*I–*Xho*I to construct the following plasmids: pHIV1Hβ, containing the first HIV-1 5′ splice site at nt 740 upstream of the *gag* gene in the *Bss*HII–*Xmn*I fragment of pSVGagPol-RRE-R (BH10 nt 709–839); pHIV2Hβ, containing the second 5′ splice site at nt 4995 in the *Ava*II–*Nde*I fragment of pSVGagPol-RRE-R (BH10 nt 4971–5158); pHIV3Hβ, containing the

third 5′ splice site at nt 5486 in the StuI–EcoRI fragment of pSVGagPol-RRE-R (BH10 nt 5440–5780); and pHIV5Hβ, containing the 4th 5′ splice site at nt 6078 in the BsmAI–NdeI fragment of pSVSX1 (BH10 nt 6022–6437).

pHβHIV 3′ splice site plasmids. The human β globin 3′ splice site, together with 53 nt of intronic sequence and 119 nt of exon 2 sequence, was removed from pRSV2Hβ by deleting the XhoI–DraIII fragment. HIV-1 BH10 fragments bearing 3′ splice sites were blunted and inserted into pRSV2Hβ ΔXhoI–DraIII to construct the following plasmids: pHβHIV2, containing the 3′ splice site of exon 2 at nt 4945 in fragment EcoRI–AvaII of pSVGagPol-RRE-R (BH10 nt 4682–4974); pHβHIV3, containing the 3′ splice site of exon 3 at nt 5422 in the NdeI–StuI fragment of pSVGagPol-RRE-R (BH10 nt 5156–5440); pHβHIV4, containing the 3′ splice site of exon 4 at nt 5810 in the StuI–EcoNI (ΔAvrII–AvrII) fragment of pSVGagPol-RRE-R (BH10 nt 5438–5825, Δ5464–5694); pHβHIV4c-5, containing the 3′ splice sites of exons 4c, 4a, 4b, and 5 at nt 5969, 5987, 5993, and 6010, respectively, in the SalI–HinfI fragment of pBH10-flop (BH10 nt 5819–6052); and pHβHIV7, containing the 3′ splice sites of exons 7, 7a, and 7b at nt 8380, 8384, and 8408 in the HindIII–BamHI fragment of pSVSX1 (BH10 nt 8172–8509).

Cell culture and transfection

The 293 cells, human embryonic kidney cells transformed by human adenovirus type 5 (kindly provided by G. Ketner), were maintained in minimum essential medium with 10% fetal calf serum and penicillin–streptomycin. Cells on a 6-cm dish were transfected with 5 µg plasmid DNA by the calcium phosphate method (Promega). To assay for transfection efficiency, the cells were cotransfected with a second plasmid, pMyc23 (McNally et al., 1991). Expression of the test plasmids and the myc plasmids was assayed simultaneously by RNase protection with different riboprobes.

RNA isolation and mapping

Total cellular RNA was isolated 48 hr after transfection with RNAzol (Chomczynski and Sacchi, 1987; Cinna/Biotecx). Hybridization and RNase protection were performed as described previously (Arrigo and Beemon, 1988; McNally et al., 1991; Melton et al., 1984) using probes described below. A PhosphorImager (Molecular Dynamics) or InstantImager (Packard) was used to quantify RNA levels, in most cases on 6% polyacrylamide–8 M urea sequencing gels (Maxam and Gilbert, 1980). Adjustments were made for the size and base composition of the respective fragments, when the percentage of spliced transcripts was calculated.

Riboprobes for analysis of the myc–CRS constructs included the pRSVNeo-int myc 5′ splice site probe, which was described previously (McNally et al., 1991). The myc 3′ splice site probe template was made by inserting a 332-bp AflII–PstI fragment from pRSVNeo-int, which spans the myc 3′ splice site, into the SmaI site of pGEM-4Z (Promega). Transcription of the HindIII-cut plasmid with T7 RNA polymerase yielded a 396-nt probe that protected bands of 332 and 186 nt, corresponding to unspliced and spliced RNA, respectively. The HIV-1 first 5′ splice site probe template was constructed by insertion of a 130-nt BssHII–XmnI blunted fragment into pGEM-3Z after the ends were repaired. This plasmid was cut with HindIII and transcribed with T7 RNA polymerase to generate a 191-nt antisense riboprobe. After hybridization, a 131-nt unspliced and a 32-nt spliced RNA were visualized on a 12% polyacrylamide–8 M urea sequencing gel.

Analysis of globin–HIV chimeric transcripts was carried out by hybridization to an anti-sense transcript of pHβ2, constructed by insertion of the HindIII–BamHI fragment of human β globin into the EcoRI–BamHI sites of pGEM-4Z. pHβ2 was linearized by digestion with HindIII and transcribed with SP6 to generate a riboprobe which spanned the human β globin exon 1–intron 1–exon 2 region, allowing for the simultaneous detection of unspliced and spliced RNAs at both the 5′ and the 3′ human β globin splice sites. A second riboprobe, pMyc2-3ΔSal (McNally et al., 1991), was synthesized to detect expression of the transfection control plasmid, pMyc23.

RESULTS

CRS elements do not affect RNA splicing

Unlike viral unspliced RNAs, the unspliced RNAs from heterologous genes containing the RSV NRS in the intron are restricted to the nucleus where they accumulate to high levels (Arrigo and Beemon, 1988). This is similar to the nuclear retention observed with the HIV-1 cis-acting repressive elements and suggested that the RSV and HIV-1 elements may have similar or overlapping activities. To test this, CRS elements from the HIV-1 gag and pol genes were tested for splicing inhibition using a heterologous intron splicing assay previously employed to characterize the RSV NRS element (Arrigo and Beemon, 1988; McNally et al., 1991).

Three different previously characterized HIV-1 CRS elements (IR-1, IR-2, and CP490) (Maldarelli et al., 1991; Cochrane et al., 1991) were inserted into the myc intron of pRSVNeo-int (Linial, 1987) at a position 163 nt downstream from the 5′ splice site (Fig. 1A). After transfection of these constructs into 293 cells, the relative amounts of spliced and unspliced RNA transcripts were determined by RNase protection assays, using a probe spanning the myc 5′ splice site (Fig. 1A). We observed that control constructs containing no intronic insert produced less than 5% unspliced RNA, whereas constructs containing the RSV NRS at this position produced approximately 95% unspliced RNA (Fig. 1B). Similar results were previously obtained when these constructs were
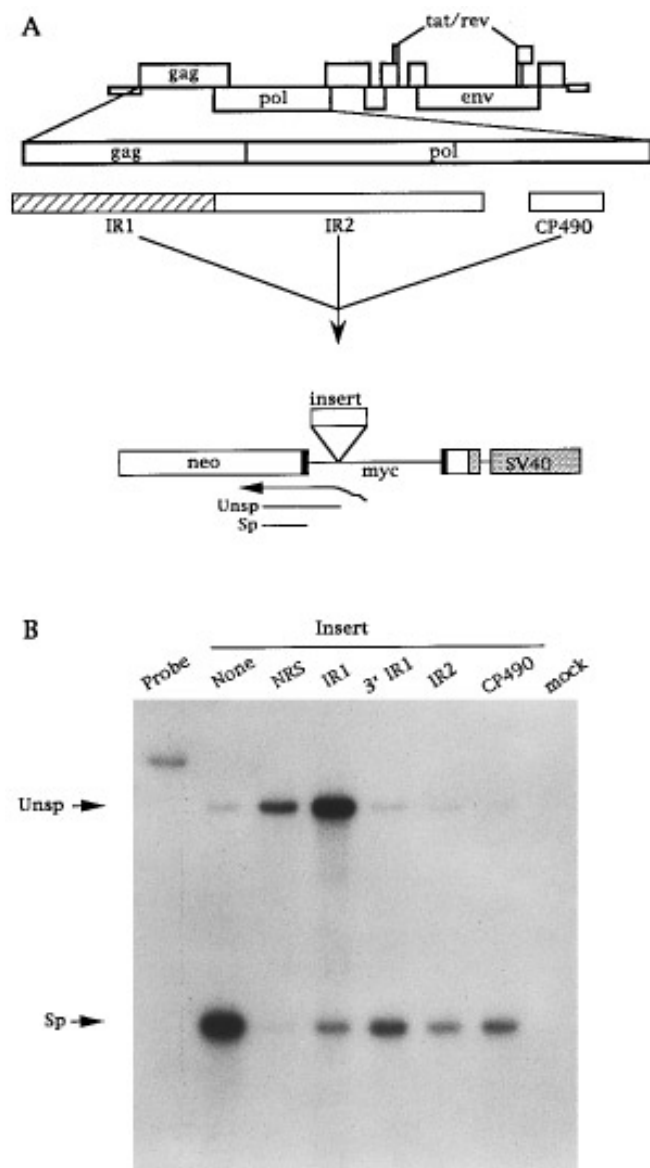
FIG. 1. Test of splicing inhibition of a heterologous intron by HIV-1 CRS elements. (A) Schematic diagram of the HIV-1 genome, showing a blowup of the *gag–pol* region indicating the location of the three CRS elements used in this study. IR-1 and IR-2 were defined by Maldarelli *et al.* (1991) and CP490 was defined by Cochrane *et al.* (1991). Each of the CRS elements was inserted into the *myc* intron of the pRSVNeo-int construct (Linial, 1987) at a site 167 nt from the 5' splice site. The relative positions of the RNase protection probe and the two digestion products are shown. Open boxes represent neo sequences, black boxes the *myc* exon sequences flanking the *myc* second intron (thin line), and stippled boxes depict SV40 sequences surrounding the small t intron. (B) RNase protection analysis of RNA from 293 cells transfected with pRSVNeo-int and its derivatives. RNA from the parental pRSVNeo-int construct is in the insert lane marked None. The other lanes indicate the fragment inserted into the *myc* intron in the sense orientation. The 3' IR-1 insert is a truncated derivative of IR-1, lacking 130 nt at the 5' end. The positions of protected fragments representing unspliced (602 nt) and spliced (441 nt) RNA are indicated at the left.

transfected into chicken embryo fibroblasts (McNally *et al.,* 1991). No inhibition of splicing was detected with constructs containing the *pol* gene CRS elements IR-2

or CP490 inserted into the *myc* intron (Fig. 1B). However, levels of unspliced RNA comparable to that generated with the NRS were seen when IR-1 was present (Fig. 1B). Control constructs containing these same inserts in the anti-sense orientation were also tested; none of these showed an increase in unspliced RNA levels (data not shown).

To localize the sequences within IR-1 that mediated splicing inhibition, we tested subfragments of IR-1 in the *myc* splicing assay. Using the *myc* 5' splice site probe, we found that a 132-nt fragment from the 5' end of IR-1 was as active at inhibiting *myc* splicing as was either the entire 1295-nt IR-1 element or the RSV NRS (Fig. 2A). No inhibition of splicing was seen with a 1163-nt fragment derived from the 3' portion of IR-1 (Fig. 2A). This 5' IR-1 fragment contains the HIV-1 first 5' splice site that is used in all spliced HIV-1 transcripts, raising the possibility that the decreased splicing observed from the *myc* 5' splice site may have resulted from competition between the *myc* and the HIV-1 5' splice sites rather than from the presence of a *cis*-acting inhibitory element like the NRS. Since the initial RNase protection assays (Figs. 1B, and 2A) were carried out with a *myc* 5' splice site probe, it was not possible to distinguish between these possibilities. Any splicing of transcripts from the HIV 5' splice site to the *myc* 3' splice site would appear as unspliced RNA using the *myc* 5' splice site probe, which is upstream of the HIV-1 insert.

Competition between the myc 5' splice site and the HIV 5' splice site within IR-1

To determine if the unspliced RNA observed in the presence of IR-1 resulted from splicing inhibition or competition between the *myc* 5' splice site and the site in IR-1, additional RNase protection assays were performed using probes for both *myc* 5' and *myc* 3' splice sites and for the HIV-1 5' splice site. If splicing were inhibited, a similar high amount of protected probe corresponding to unspliced RNA would be present with both *myc* splice site probes. In contrast, competition for splicing by the HIV-1 5' splice site would result in higher levels of fully protected probe (apparent unspliced RNA) using the *myc* 5' splice site probe than with the *myc* 3' splice site probe.

When RNA from 293 cells transfected with RSV NRS-containing constructs was analyzed, the 5' and 3' *myc* splice site probes both protected a substantial amount of unspliced RNA, as expected from a general splicing defect, and probes specific for the NRS indicated no splicing within the NRS itself (data not shown). With the IR1 and 5'IR1 constructs, the *myc* 5' splice site probe generated substantial signal representing unspliced RNA (Fig. 2A), but the 3' splice site probe suggested splicing was nearly complete (Fig. 2B). This is consistent with efficient splicing occurring between the *myc* 3' splice site and another 5' splice site, most likely that in the 5' end of the HIV IR-1 construct. This interpretation
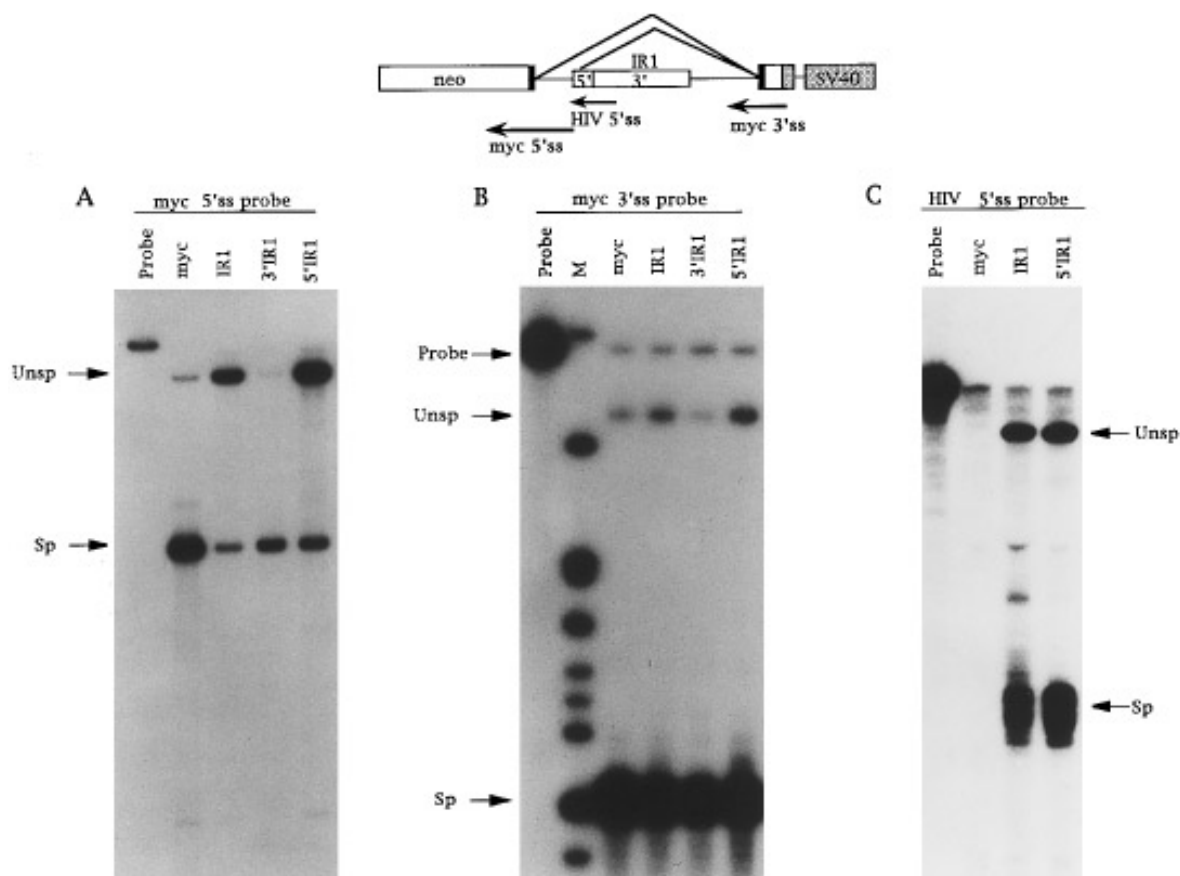
FIG. 2. The 5′ splice site within HIV-1 IR-1 competes with the *myc* 5′ splice site in pRSVNeo-int. The RNA and probes used in RNase protection assays to assess use of the first HIV 5′ splice site in a heterologous construct are illustrated at the top. Probes flanking the three splice sites are shown beneath the RNA diagram. Potential splicing from either the *myc* or the HIV 5′ splice site is depicted by lines connecting each to the *myc* 3′ splice site. (A) RNase protection assay of RNA from transfected 293 cells using a *myc* 5′ splice site probe, as described in Fig. 1. Protected RNAs were resolved on a 6% polyacrylamide–8 *M* urea gel. The lane labeled myc contains the parental pRSVNeo-int construct. IR1 contains the entire HIV-1 IR-1 fragment inserted into the intron of pRSVNeo-int as described in Fig. 1. 3′ IR1 contains the 3′-terminal portion of IR-1 (nt 840–2003); 5′IR1 contains the 5′-terminal portion of IR-1 (nt 708–840), including the first HIV-1 5′ splice site. (B) RNAs from (A) were also analyzed with a *myc* 3′ splice site probe. The *myc* 3′ splice site probe is 396 nt and protects 332- and 167-nt fragments representing unspliced and spliced RNA, respectively. M lane contains markers. (C) RNAs from (A) were also analyzed with a HIV-1 5′ splice site probe. The HIV-1 5′ splice site probe is 191 nt, and protects 131- and 32-nt spliced bands. These bands were resolved on a 12% polyacrylamide–8 M urea gel.

was tested by RNase protection analysis with a probe specific for the HIV-1 first 5′ splice site; this probe showed high levels of spliced RNA with both the IR1 and the 5′IR1 constructs (Fig. 2C). In conclusion, none of the three CRS elements tested inhibited splicing, and the first HIV 5′ splice site competed efficiently with the *myc* 5′ splice site in these constructs.

Fourteen HIV-1 splice sites assayed in chimeric constructs

The HIV-1 genome contains 20 exons, which are alternatively spliced using 16 splice sites, shown in Fig. 3, to generate more than 40 different transcripts (Felber *et al.,* 1990; Muesing *et al.,* 1985; Purcell and Martin, 1993; Robert-Guroff *et al.,* 1990; Schwartz *et al.,* 1990). To extend the observation that the first HIV-1 5′ splice site is very strong (Fig. 2C), we investigated the relative strengths of the other HIV-1 splice sites. To compare

these splice sites to one another, a set of chimeric constructs containing 14 HIV-1 splice sites paired with human β globin splice sites was generated, and the splicing efficiencies of these constructs were compared in transfected cells.

The nucleotide positions of the HIV-1 BH10 splice sites were determined by comparison of the BH10 sequence (Ratner *et al.,* 1986) to that of HIV-1 HXB2 splice sites (Schwartz *et al.,* 1990, 1991) and to the splice sites of pNL4-3 in the case of 3′ splice site 4c (Purcell and Martin, 1993). In all cases, the BH10 splice junctions were identical in sequence to those of HXB2. Table 1 displays a comparison of the HIV-1 BH10 splice site sequences to consensus splice sites (Senapathy *et al.,* 1990). While all of the 5′ splice sites match the consensus sequence by 6 to 8 of 9 nucleotides, none of the 3′ splice sites are good matches to the consensus because their pyrimidine tracts are very short and interrupted by purines. Since none of the HIV-1 branchpoints have been identified ex-
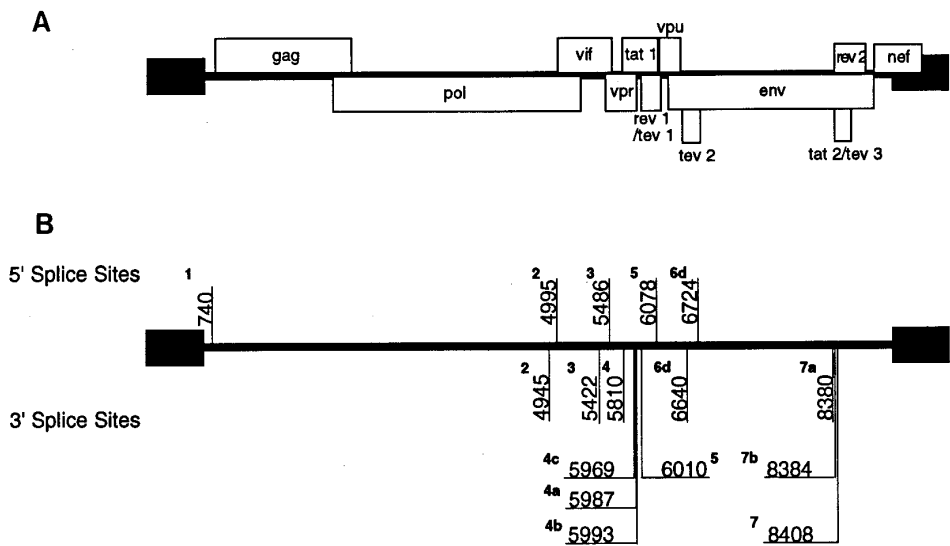
A

B

FIG. 3. HIV-1 BH10 genome organization. (A) Open reading frames. (B) Splice sites. BH10 splice sites were determined by sequence comparison to splice sites identified in HXB2 and pNL4-3 (Purcell and Martin, 1993; Schwartz *et al.,* 1990, 1991). The locations of the 5′ splice sites are indicated above and of the 3′ splice sites below the diagram. 5′ splice site 5 is used by multiple exons, containing 3′ splice sites 3–5.

perimentally and these can be as far as 100 nt upstream of the 3′ splice sites (reviewed in Smith *et al.,* 1989), we have not tried to match putative branchpoints to the branch consensus sequence.

To examine the efficiency of individual HIV-1 splice sites *in vivo,* a splice site swapping strategy was used in which a fragment containing a 5′ or 3′ HIV-1 splice site was paired with an efficient splice site from human β globin. The parental β globin construct, pRSV2Hβ, shown in Fig. 4, was constructed by the insertion of human β globin sequence spanning the entire first exon

TABLE 1

Comparison of HIV-1 BH10 Splice Site Sequences with Consensus Sequence Splice Sites[a]

5′ splice site	Sequence
Consensus	(C/A)AG/GT(A/G)AGT
HIV1Hβ	CTG/GTGAGT
HIV2Hβ	AAG/GTGAAG
HIV3Hβ	AAG/GTAGGA
HIV5Hβ	GCA/GTAAGT

3′ splice site	Sequence
Consensus	(Py)$_n$N(C/T)AG/G
HβHIV2	GGGTTTATTACAG/G
HβHIV3	GACTGTTTTTCAG/A
HβHIV4	TATCCATTTTCAG/A
HβHIV4c	TTTCATTGCCAAG/T
HβHIV4a	TTCATAACAAAAG/C
HβHIV4b	AACAAAGCTTAG/G
HβHIV5	TCTCCTATGGCAG/G
HβHIV7a	TGAATAGAGTTAG/G
HβHIV7b	TAGAGTTAGGCAG/G
HβHIV7	ATTATCGTTTCAG/A

[a] Nucleotides which match the consensus are underlined.

and intron and most of the second exon (Reed, 1989) into the pRSV2 expression vector (McNally *et al.,* 1991), so that transcription of this plasmid initiates in the RSV long terminal repeat and is terminated by the SV40 early polyadenylation signal (lacking any SV40 intron). The second β globin exon was truncated at its 3′ end by 14 nt in this construct to avoid the inclusion of the 5′ splice site of exon 2 upstream of the polyadenylation site; this has been shown to interfere with the definition of the ultimate exon and polyadenylation in other constructs assayed *in vivo* (Niwa *et al.,* 1992).

To generate chimeric splicing substrates, a fragment containing the 5′ or 3′ splice site of human β globin was deleted from pRSV2Hβ and an HIV-1 fragment bearing a 5′ or 3′ splice site was inserted in its place (Fig. 4). Four HIV-1 5′ splice site/β globin chimeric plasmids were constructed after the globin 5′ splice site, together with 94 nt of exon 1 and 77 nt of intron 1, was deleted. So that the splice sites are in their normal, viral context, each 5′ HIV-1 splice site chimeric construct contained from 24 to 56 nt of adjacent HIV-1 exonic sequence and >90 nt of intronic sequence. For example, a 130-nt sequence including the first HIV-1 5′ splice site, containing 31 nt of exon 1 and 99 nt of adjacent intron, was substituted for the globin 5′ splice site and flanking sequences to generate pHIV1Hβ. Similarly, fragments bearing the 5′ splice site regions of splice sites 2 and 3 (Fig. 3B) were used to generate pHIV2Hβ and pHIV3Hβ, respectively. The 5′ splice site region 5 cloned into pHIV5Hβ was used in a number of different exons with alternative 3′ splice sites, including 3, 4, 4c, 4a, 4b, and 5 (Fig. 3B).

Five β globin/HIV-1 3′ splice site chimeric plasmids were also constructed. As with the 5′ splice site HIVHβ plasmids, the numbers in the names correspond to the numbers of the splice sites shown in Fig. 3B. pHβHIV2
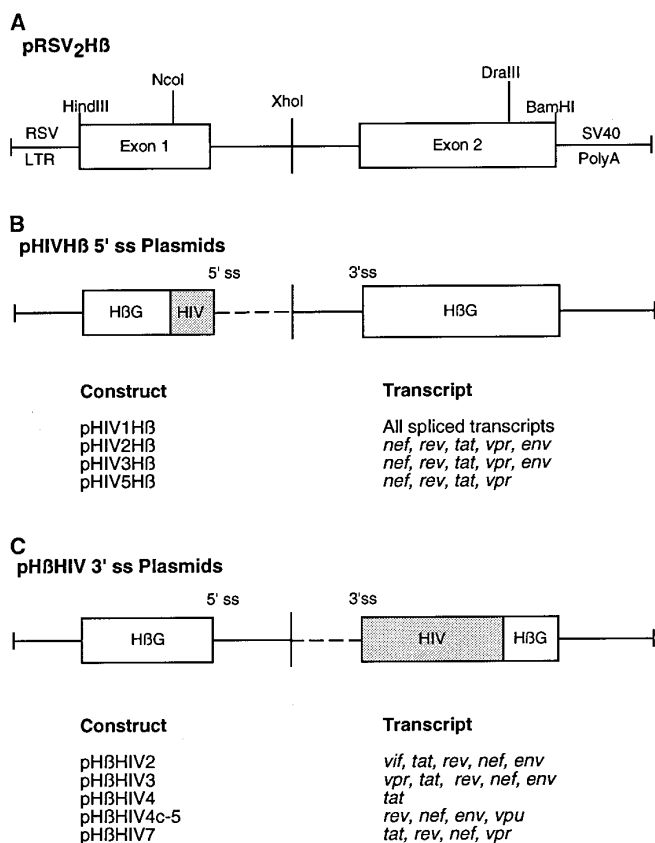
FIG. 4. Plasmid constructs. (A) pRSV2Hβ. The human β globin exon 1–intron 1–exon 2 region was inserted into the expression vector pRSV2 such that transcription initiates in the RSV LTR and is terminated by the SV40 poly(A) site. (B) pHIVHβ 5′ splice sites. The human β globin 5′ splice site region was deleted from pRSV2 and replaced with a 5′ splice site region from HIV. The numbers in the plasmid names refer to the splice sites, numbered as in Fig. 3. Also indicated are the transcripts generated by use of that splice site. (C) pHβHIV 3′ splice site plasmids. The β globin 3′ splice site region was deleted from pRSV2Hβ and replaced with a 3′ splice site from HIV.

and pHβHIV3 include sequences bearing the 3′ splice sites of exons 2 and 3, respectively. The fragment cloned into the pHβHIV4 plasmid contains the 3′ splice site of exon 4. The plasmid pHβHIV4c-5 includes the 3′ splice sites of exons 4c, 4a, 4b, and 5 that are clustered within a 41-nt region. Because 3′ splice sites include significant upstream intronic sequence (branchpoint and pyrimidine tracts) which would overlap with these splice sites, it was not possible to separate them. For the same reason, three 3′ splice sites, 7, 7a, and 7b, that are clustered within a 28-nt region, were cloned into pHβHIV7. Because exon 6D is used infrequently (Benko *et al.*, 1990; Purcell and Martin, 1993), we chose not to investigate its splice sites. Since none of the HIV-1 branchpoints have been mapped, enough intronic sequence (>100 nt) was included in the HIV-1 3′ splice site constructs to contain any distant branchpoints.

To analyze the splicing efficiency of the chimeric HIV-1/human β globin transcripts *in vivo*, the constructs were transfected into 293 cells with calcium phosphate, and total RNA was isolated 48 hr after transfection. RNase protection assays were performed using an anti-sense riboprobe that was complementary to human β globin exons 1 and 2, as well as the intervening intron 1, allowing unspliced and spliced transcripts using either splice site to be distinguished by gel electrophoresis (Figs. 5 and 6). Thus, this same probe could be used to assay the splicing efficiency of all nine chimeras, allowing direct comparisons between them without variations due to the use of different probes. Multiple transfection experiments were performed, and the splicing level of each construct was quantified. Table 2 displays the average percentage of spliced RNA obtained with each construct.

The HIV-1 5′ splice sites of exons 1 and 5 are used efficiently

The transcripts generated by the control plasmid, pRSV2Hβ, were spliced quite efficiently (89 ± 12% spliced) as expected (Fig. 6, Table 2). However, the HIV1Hβ transcript bearing the first HIV-1 5′ splice site, upstream of the *gag* coding region, was spliced even more efficiently (95 ± 4% spliced) than the human β globin first 5′ splice site. We have observed similar results with a second HIV1Hβ construct containing 264 nt of additional HIV-1 exonic sequence upstream of the 5′ splice site, extending to include the authentic HIV-1 RNA 5′ terminus (C. T. O'Sullivan and K. L. Beemon, unpublished results). This suggests that there are no inhibitory elements upstream of this 5′ splice site that are active in 293 cells. These results extend and confirm our earlier finding of a strong first 5′ splice site, based on splice site competition experiments in a construct with two 5′ splice sites (Fig. 2C). In that experiment, different amounts of HIV-1 downstream intronic sequence (100 to 1263 nt) in constructs 5′IR1 and IR1 did not affect usage of this 5′ splice site, suggesting that neither upstream nor downstream sequences regulate its activity. This splice site, which is used in all the spliced mRNAs, exhibits the closest match to a consensus 5′ splice site (8/9 nt, including all of the intronic nt) of all the HIV-1 splice sites (Table 1). The presence of an efficient first 5′ splice site in HIV-1 reiterates the finding of a strong 5′ splice site in avian retroviruses (Katz and Skalka, 1990; McNally and Beemon, 1992).

The 5′ splice site in pHIV5Hβ, which is used in multiple mRNAs to remove the intron nearest the 3′ end of the RNA to generate the 2-kb class of HIV-1 mRNAs (*tat, rev, nef*), was 83 ± 10% spliced (Fig. 6, Table 2). It may be significant that this splice site is not as strong as the first 5′ splice site, since many HIV-1 mRNAs are singly spliced, and all of these have spliced out only the first intron. Thus, the 4-kb class of HIV-1 mRNAs all use the first 5′ splice site and do not use the second major 5′ splice site at nt 6078. The HIV-1 5′ splice sites 2 (38 ± 11% spliced) and 3 (45 ± 6% spliced) of exons 2 and 3,
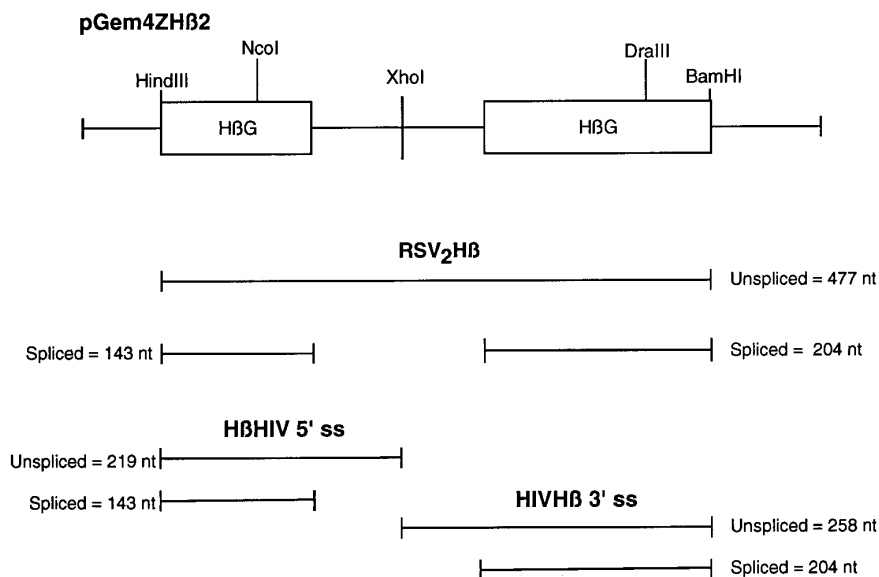
FIG. 5. Probes used for RNase protection assay. Linearized pGEM4ZHβ2 was used as a template for the synthesis of a riboprobe which spanned both β globin intron 1 splice sites. The sizes of unspliced and spliced pRSV2Hβ, HIVHβ, or HβHIV fragments protected by this probe are indicated.

which are present in the 2-kb transcripts as optional, upstream, noncoding exons, were less efficiently used in the hybrid transcripts than were the 5' splice sites 1 or 5.

All the HIV-1 3' splice sites are suboptimal

The levels of splicing observed with the HIV-1 3' splice site chimeric constructs were also variable (Fig. 6, Table 2). The 3' splice site 2 was the least efficiently used of all the splice sites (26 ± 16% spliced). Splicing to this site is necessary to generate the 4-kb *vif* mRNA. Exon 2, bordered by 3' splice site 2 and 5' splice site 2, is also present as a noncoding exon in many 2-kb mRNAs. A higher level of spliced RNA was obtained with 3' splice site 3 (60 ± 10% spliced), which is used to generate the 4-kb *vpr* mRNA and is also used, in conjunction with 5' splice site 3, as an optional, noncoding exon of the 2-kb mRNAs. HβHIV4 RNA, containing the first *tat* 3' slice site, was spliced at an intermediate level (55 ± 11% spliced). Despite bearing multiple 3' splice sites from exons 4c, 4a, 4b, and 5 (for *rev, nef, env,* and *vpu*), pHβHIV4c-5 transcripts (48 ± 11% spliced) were spliced with slightly lower efficiency. The level of splicing obtained with pHβHIV7 transcripts, which contain the 3' splice sites of *tat, rev,* and *nef* 3'-terminal exons, was even lower (41 ± 11% unspliced). Since HβHIV7 contains three alternative 3' splice sites (7, 7a, and 7b) but only use of splice site 7 results in functional mRNAs (Schwartz *et al.,* 1990), the level of splicing to 7 is lowered due to competition between these three splice sites, which are present in a 28-nt region. In contrast to the HIV-1 5' splice sites, none of the 3' splice sites yielded more than 60% spliced RNA. Therefore, in HIV-1 there are no possible combinations of strong 5' and 3' splice sites that can be used together.

DISCUSSION

The HIV-1 primary RNA transcript is used in its unspliced form as mRNA and genomic RNA; it is also a pre-mRNA that is alternatively spliced to generate more than 40 different mRNAs. As an initial step in dissecting the mechanisms involved in this very complex splicing pathway, we have compared the relative efficiencies of the HIV-1 splice sites in the absence of viral proteins and have searched for *cis*-acting splicing regulatory sequences in the HIV-1 *gag* and *pol* genes. In contrast to previous conclusions that HIV-1 splice sites are suboptimal, based on analysis of a subset of these splice sites, we have observed that the first 5' splice site is very strong in several different contexts. Further, the 5' splice site used to generate the 2-kb mRNAs is relatively strong, but weaker than the first 5' splice site. However, all of the 3' splice sites are suboptimal (<60% splicing); thus, there are no strong pairs of splice sites in HIV-1 pre-mRNA. Comparison of the relative strengths of splice sites of HIV-1 and RSV RNAs showed striking similarities, in that both have strong 5' splice sites and weaker 3' splice sites (Fig. 7).

Comparison of the relative efficiencies of 14 HIV-1 splice sites in globin chimeric constructs showed considerable variation. Transcripts containing HIV-1 5' splice sites ranged from 38 to 95% spliced, while the HIV-1 3' splice site transcripts varied from 26 to 60% spliced. We propose that this inherent variability in the efficiency of the HIV-1 splice sites contributes to the control of alternative splicing in HIV-1. In most cases, the efficiency of use of a splice site in the chimeric transcript correlated with that splice site's usage in HIV-infected cells, as determined by the relative abundance of transcripts containing that splice site (Guatelli *et al.,* 1990; Purcell and Martin, 1993; Robert-Guroff *et al.,* 1990). The most efficient splice
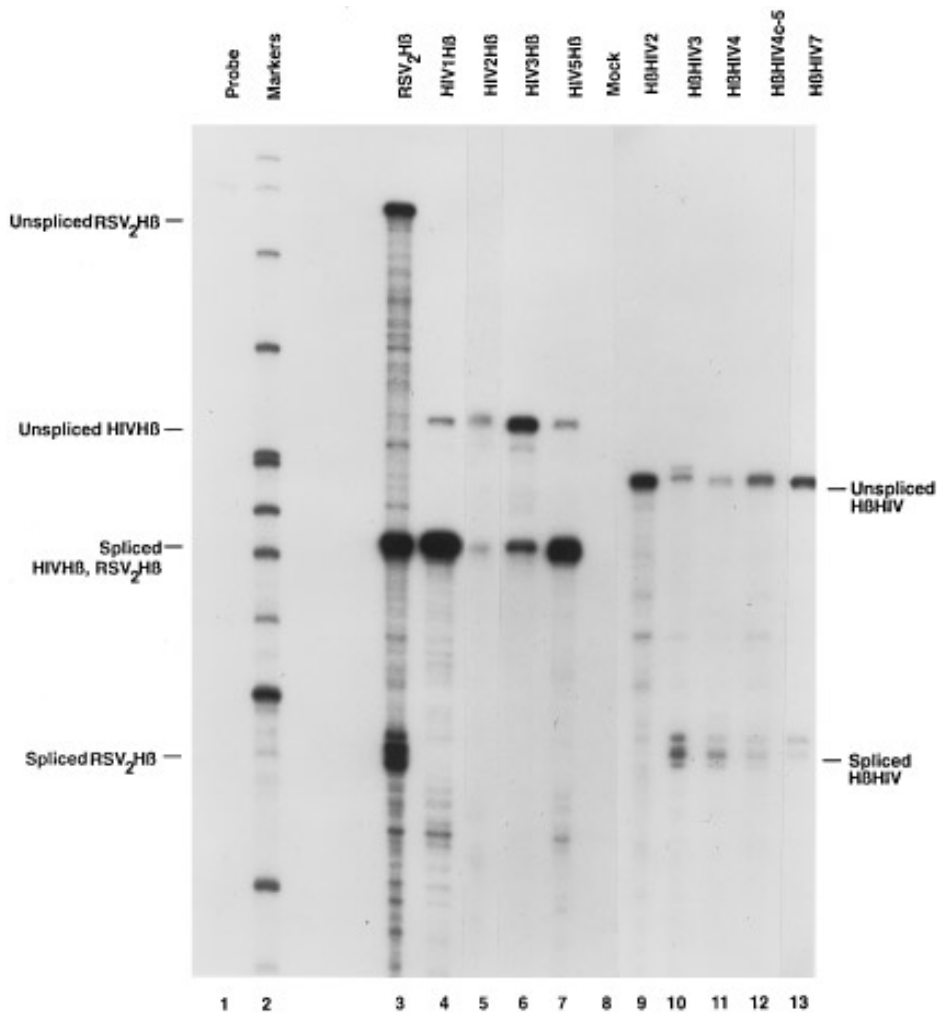
FIG. 6. Splicing efficiency of chimeras assessed by RNase protection assays. RNase protection analysis was performed on total RNA from 293 cells transfected with the parental human β globin construct, pRSV2Hβ (lane 3); with the 5′ HIV splice site constructs, pHIVHβ (lanes 4–7); or with 3′ HIV splice site constructs, pHβHIV (lanes 9–13), using the Hβ2 riboprobe to distinguish unspliced from spliced RNA. Lane 1 contains the β globin riboprobe. The β globin riboprobe protected two spliced bands representing both exons of pRSV2Hβ (lane 3) and is therefore a control for both HIVHβ and HβHIV splicing. Multiple-spliced bands, all of which were quantified together, were observed with all of the HβHIV transcripts (lanes 9–13) and also with the globin control (lane 3). The amount of total RNA recovered after RNase protection with construct HIV2Hβ (lane 5) was reproducibly less than with the other constructs, possibly due to instability of this chimeric RNA. The band showing the control for transfection efficiency was cut off of the bottom of this gel. From the top of the gel, the sizes of the HpaII DNA marker bands are 622, 527, 404, 309, 242, 238, 217, 180, 160, and 123 nt (lane 2).

site, the first 5′ splice site, is used in all spliced transcripts. The second most efficient was the 5′ splice site used in all of the 2-kb transcripts encoding Tat, Rev, and Nef. The 3′ splice sites of exon 7, also used in these 2-kb transcripts, were less efficiently used than the 5′ splice site, consistent with previous studies (Chang and Sharp, 1989; Staffa and Cochrane, 1994). The infrequently used exons 2 and 3 had relatively weak splice sites; these exons had the poorest 5′ splice sites, and exon 2 also had the poorest 3′ splice site (26% spliced), in contrast exon 3 had a relatively strong 3′ splice site (60% spliced).

Some variation in total RNA levels was observed between the different chimeric transcripts shown in Fig. 6. In particular, the level of HIV2Hβ transcripts was repro-

ducibly lower than the others. This may reflect differences in the stability of the transcripts. If degradation occurs prior to splicing or affects spliced and unspliced transcripts similarly, this would have no effect on the levels of splicing observed. However, a difference in the stability of the spliced and unspliced transcripts would alter the interpretation of splicing levels. Additional work is needed to determine the relative stabilities of HIV-1 mRNAs. It is interesting that Rev has been observed to increase the stability of incompletely spliced RRE-bearing HIV-1 transcripts (Felber et al., 1989; Malim and Cullen, 1993).

As in the simple avian retroviruses previously studied (Katz and Skalka, 1990; McNally and Beemon, 1992), the 5′ splice site nearest the 5′ end of HIV-1 RNA was highly

TABLE 2

Average Percentage Spliced RNAs of the
HIV-1/Human β Globin Construct

Construct	Average	SD	N
RSV2Hβ	89	12	7
HIV1Hβ	95	4	7
HIV2Hβ	38	11	8
HIV3Hβ	45	6	8
HIV5Hβ	83	10	6
HβHIV2	26	16	7
HβHIV3	60	10	7
HβHIV4	55	11	8
HβHIV4c-5	48	11	7
HβHIV7	41	11	7

Note. N indicates number of transfections.

efficient (Fig. 7). The importance of a strong first 5′ splice site for viral replication is supported by the observation of Purcell and Martin (1993) that mutation of this splice site leads to activation of a cryptic site 4 nt downstream in a sequence highly conserved among HIV-1 isolates. Since all spliced HIV-1 mRNAs (both 4- and 2-kb classes) use the first 5′ splice site, it is not surprising that this is a strong splice site; otherwise aberrantly spliced mRNAs retaining the first intron might be generated. Further, the presence of alternative weak 3′ splice sites that pair with a single strong 5′ splice site allows multiple mRNAs to be generated by competition between these 3′ splice sites. Work of Hammarskjold and coworkers (Hammarskjold *et al.,* 1994; Lu *et al.,* 1990) also suggests that the presence of an active 5′ splice site may be necessary

for stability and Rev-dependent expression of HIV-1 *env* mRNAs.

The 3′ splice sites of the avian retroviruses are both suboptimal (Berberich and Stoltzfus, 1991; Fu *et al.,* 1991; Katz and Skalka, 1990; McNally and Beemon, 1992; Zhang and Stoltzfus, 1995). The 3′ splice sites of HIV-1 were also suboptimally spliced to a comparable level as shown in Fig. 7. The competing 3′ splice sites (4, 4c, 4a, 4b, and 5) used in *tat, rev, nef,* and *vpu/env* transcripts are closely spaced within a 200-nt region. Chimeric pHβHIV4 transcripts bearing the 3′ splice site nearest the 5′ end of this cluster that is used in *tat* transcripts were 55% spliced in our assays. However, in the viral context this splice site is down-regulated by a splicing regulatory element (SRE) (Amendt *et al.,* 1994), which is not present in our pHβHIV4 construct. The 3′ splice sites of exons 4a, 4c, 4b, and 5 could not be separated since they are clustered within a 41-nt region; together they spliced at 48%. Splicing at the individual sites within this cluster must represent a fraction of the composite we have measured. It would be of interest to map the branchpoints involved in HIV-1 RNA splicing, particularly those involving this cluster of four 3′ splice sites. It is possible that some of these splice sites may share branchpoints or pyrimidine tracts. Temporal analysis of alternative splicing of the competing 3′ splice sites in the transcripts encoding the regulatory proteins Tat and Rev revealed that the relative levels of these transcripts remained constant throughout viral replication, suggesting that alternative splicing was not being regulated by *trans*-acting viral or cellular factors but by *cis*-acting splice site prefer-
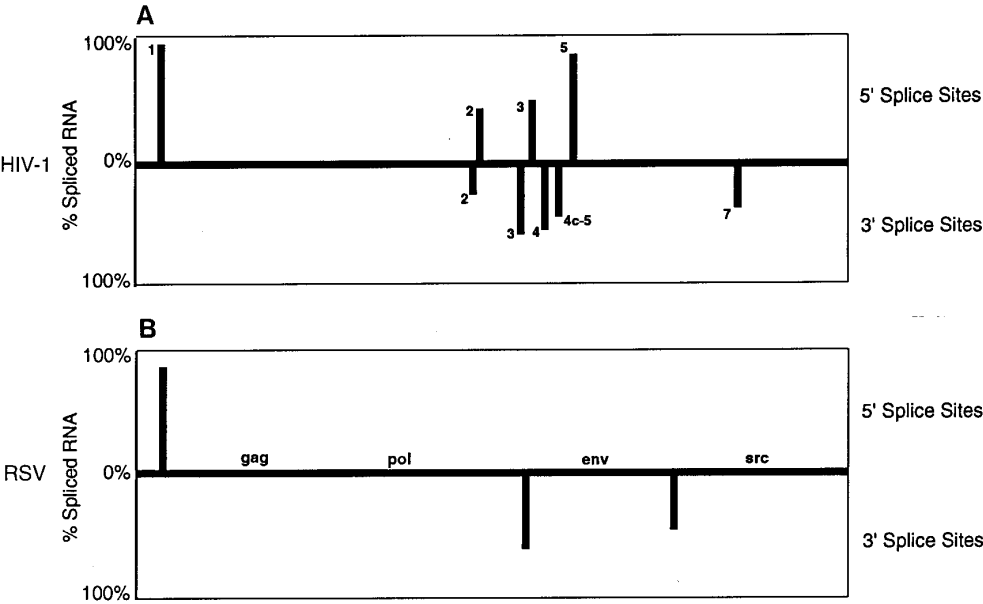

FIG. 7. Comparison of splice site efficiencies for HIV-1 and RSV RNAs. (A) The average percentages of spliced RNAs from Table 2 are plotted for the HIV-1 splice sites positioned on the HIV-1 RNA genome. (B) The average spliced RNA levels are plotted for RSV splice sites. Similar splice site swapping experiments were performed in which the RSV splice sites were paired with the splice sites associated with the chicken c-*myc* second intron (McNally and Beemon, 1992). The RSV major 5′ splice site is 88% spliced, the *env* 3′ splice site is 62% spliced, and the *src* 3′ splice site is 47% spliced; thus in both HIV-1 and RSV, the major 5′ splice site was strong while the 3′ splice sites were weaker.

ences (Guatelli *et al.,* 1990). The existence of variable splice site efficiencies would be consistent with this.

In the avian retroviruses, splicing is regulated both by suboptimal splice sites and by *cis*-acting elements (Katz and Skalka, 1990; McNally and Beemon, 1992; Zhang and Stoltzfus, 1995). In RSV, an intronic sequence in the *gag* gene, the NRS, inhibits splicing (Arrigo and Beemon, 1988; Gontarek *et al.,* 1993; McNally *et al.,* 1991; Stoltzfus and Fogarty, 1989). In this work, analysis of the *gag–pol* region of HIV-1 did not reveal sequences functionally analogous to the NRS. The HIV-1 *gag* CRS element, IR-1, appeared initially to be an effective inhibitor of splicing from the *myc* 5′ splice site, but the inhibition observed was not from activity similar to that of the NRS but rather by competition between the HIV and the *myc* 5′ splice sites. The sequences within the RSV NRS resemble both 5′ and 3′ splice sites, and it has been shown that U1 and U2 snRNPs, as well as U11/U12 snRNPs, interact with the NRS (Gontarek *et al.,* 1993). However, splicing has never been observed into or out of the NRS in RSV constructs, even when the authentic viral 3′ splice sites were deleted to relieve any potential competition (unpublished results). While it is possible that the methods used here would not identify all negative elements, it appears that elements capable of acting over a long distance, like the NRS, are not present in the *gag–pol* region of HIV.

While variation in efficiency of individual splice sites would contribute to the regulation of alternative splicing in HIV-1, it does not preclude the existence of other levels of regulation. Although we did not find an intronic, global, *cis*-acting element in the *gag* or *pol* sequence of HIV-1, Amendt *et al.* (1994) identified a SRE within the first *tat* coding exon of HIV-1 that acted in a position- and orientation-dependent manner to inhibit splicing at the upstream 3′ splice site. Deletion of the small element increases splicing at the 3′ splice site of exon 4 (Amendt *et al.,* 1994). The short exonic portion of our pHβHIV4 construct did not include the SRE, and transcripts of this construct were spliced to a level of approximately 55% in our transfection experiments in 293 cells. Presumably, inclusion of the SRE would further diminish this level of splicing. Thus, splicing at 3′ splice site 4 is controlled both by an inefficient splice site and by a negative downstream element. Although the SRE was present in pHβHIV4c-5, the 3′ splice sites of exons 4c, 4a, 4b, and 5 are not affected by the SRE (Amendt *et al.,* 1994). Since the efficiency of splicing can be modulated by sequences flanking the splice sites that interact with cellular factors (Amendt *et al.,* 1994, and references therein), further work will be needed to determine whether including more or less flanking sequence will alter the efficiency of splicing at some of the HIV-1 splice sites. For the first HIV-1 splice site, no difference in splicing levels was observed between a 132-nt fragment and fragments including considerably more upstream or downstream sequence, suggesting regulatory elements may not be present in the vicinity of the strong, first 5′ splice site of HIV-1 RNA.

Unlike the simple retroviruses, HIV-1 encodes a regulatory protein, Rev, necessary for productive expression of unspliced and partially spliced viral transcripts in the cytoplasm (Emerman *et al.,* 1989; Felber *et al.,* 1989; Hammarskjold *et al.,* 1989; Malim *et al.,* 1989). Multiple functions have been demonstrated for Rev including inhibition of splicing (Felber *et al.,* 1990; Kjems *et al.,* 1992; Kjems and Sharp, 1993), increasing RNA stability (Feinberg *et al.,* 1986; Felber *et al.,* 1989), nuclear export (Emerman *et al.,* 1989; Felber *et al.,* 1989; Hammarskjold *et al.,* 1989; Malim *et al.,* 1989; Fischer *et al.,* 1994), and translation (Arrigo and Chen, 1991; Cochrane *et al.,* 1991; D'Agostino *et al.,* 1992) of RRE-containing transcripts. Rev regulation of β globin constructs requires a single suboptimal splice site that delays intron removal, suggesting that Rev may promote the disassociation of splicing complexes to allow transport of unspliced or incompletely spliced RNA to the cytoplasm (Chang and Sharp, 1989). We have demonstrated that all HIV-1 introns are flanked by at least one suboptimal splice site. Despite the presence of two efficient 5′ splice sites, there were no pairs of efficient splice sites in HIV-1 because all the 3′ splice sites were suboptimal. This may be necessary to maintain the high level of HIV-1 transcripts that remain unspliced and to allow the use of all of the alternative 3′ splice sites. *In vitro* studies have shown that a Rev peptide can inhibit splicing of RRE-containing transcripts (Kjems *et al.,* 1992; Kjems and Sharp, 1993); however, this has not been observed with Rev-defective constructs studied *in vivo* (Malim and Cullen, 1993). While Rev-mediated transport of RRE-containing RNA out of the nucleus will prevent its further splicing, our studies show that splicing is inefficient even in the absence of Rev.

ACKNOWLEDGMENTS

REFERENCES

Amendt, B. A., Hessllein, D., Chang, L-J., and Stoltzfus, C. M. (1994). Presence of negative and positive *cis*-acting RNA splicing elements within and flanking the first *tat* coding exon of human immunodeficiency virus type 1. *Mol. Cell. Biol.* **14,** 3960–3970.

Arrigo, S., and Beemon, K. (1988). Regulation of Rous sarcoma virus RNA splicing and stability. *Mol. Cell. Biol.* **8,** 4858–4867.

Arrigo, S. J., and Chen, I. S. Y. (1991). Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 *vif, vpr,* and *env/vpu* 2 RNAs. *Gene Dev.* **5,** 808–819.

Benko, D. M., Schwartz, S., Pavlakis, G. N., and Felber, B. K. (1990). A novel human immunodeficiency virus type 1 protein, *tev,* shares sequence with *tat, env,* and *rev* proteins. *J. Virol.* **64,** 2505–2518.

Berberich, S. L., and Stoltzfus, C. M. (1991). Mutations in the regions of the Rous sarcoma virus 3' splice sites: Implications for regulation of alternative splicing. *J. Virol.* **65**, 2640–2646.

Bouck, J., Fu, X-D., Skalka, A. M., and Katz, R. A. (1995). Genetic selection for balanced retroviral splicing: Novel regulation involving the second step can be mediated by transitions in the polypyrimidine tract. *Mol. Cell. Biol.* **15**, 2663–2671.

Chang, D. D., and Sharp, P. A. (1989). Regulation by HIV Rev depends upon recognition of splice sites. *Cell* **59**, 789–795.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162**, 156–159.

Cochrane, A., Kramer, R., Ruben, S., Levine, J., and Rosen, C. A. (1989). The human immunodeficiency virus Rev protein is a nuclear phosphoprotein. *Virology* **171**, 264–266.

Cochrane, A. W., Jones, K. S., Beidas, S., Dillon, P. J., Skalka, A. M., and Rosen, C. A. (1991). Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J. Virol.* **65**, 5305–5313.

Coffin, J. M. (1990). Retroviradae and their replication. *In* "Fields Virology" (B. N. Fields and D. M. Knipe, Eds.), 2nd ed., pp. 1437–1500. Raven Press, New York.

Cullen, B. M. (1991). Human immunodeficiency virus as a prototypic complex retrovirus. *J. Virol.* **65**, 1053–1056.

D'Agostino, D. M., Felber, B. K., Harrison, J. E., and Pavlakis, G. N. (1992). The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of *gag/pol* and *vpu/env* mRNAs. *Mol. Cell. Biol.* **12**, 1375–1386.

Emerman, M., Vazeux, R., and Peden, K. (1989). The *rev* gene product of human immunodeficiency virus affects envelope specific RNA localization. *Cell* **57**, 1155–1165.

Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C., and Wong-Staal, F. (1986). HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* **46**, 807–817.

Felber, B. K., and Pavlakis, G. N. (1993). Molecular biology of HIV-1: Positive and negative regulatory elements important for virus expression. *AIDS* **7**(Suppl.), S51–S62.

Felber, B. K., Drysdale, C. M., and Pavlakis, G. N. (1990). Feedback regulation of human immunodeficiency virus type 1 expression by the Rev protein. *J. Virol.* **64**, 3734–3741.

Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T., and Pavlakis, G. N. (1989). Rev protein of human immunodeficiency virus type 1 affects the stability and transport of viral mRNA. *Proc. Natl. Acad. Sci. USA* **86**, 1496–1499.

Fischer, U., Meyer, S., Teufel, M., Heckel, C., Luhrmann, R., and Rautmann, G. (1994). Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *EMBO J.* **13**, 4105–4112.

Fu, X-D., Katz, R. A., Skalka, A. M., and Maniatis, T. (1991). The role of branchpoint and 3' exon sequence in the control of balanced splicing of avian retrovirus RNA. *Genes Dev.* **5**, 211–220.

Garrett, E. D., Tiley, L. S., and Cullen, B. R. (1991). Rev activates expression of the human immunodeficiency virus type 1 *vif* and *vpr* gene products. *J. Virol.* **65**, 1653–1657.

Gontarek, R. R., McNally, M. T., and Beemon, K. (1993). Mutation of an RSV intronic element abolishes both U11/U12 snRNP binding and negative regulation of splicing. *Genes Dev.* **7**, 1926–1936.

Guatelli, J. C., Gingeras, T. R., and Richman, D. D. (1990). Alternative splice acceptor utilization during human immunodeficiency virus type 1 infection of cultured cells. *J. Virol.* **64**, 4093–4098.

Hadzopoulou-Cladaras, M., Felber, B. K., Cladaras, C., Athanassopoulos, A., Tse, A., and Pavlakis, G. (1989). The rev (*trs/art*) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a *cis*-acting sequence in the *env* region. *J. Virol.* **63**, 1265–1274.

Hammarskjold, M. L., Heimer, J., Hammarskjold, B., Sangwan, I., Albert, L., and Rekosh, D. (1989). Regulation of human immunodeficiency virus *env* expression by the *rev* gene product. *J. Virol.* **63**, 1959–1966.

Hammarskjold, M-L., Li, H., Rekosh, D., and Prasad, S. (1994). Human immunodeficiency virus *env* expression becomes Rev-independent if the *env* region is not defined as an intron. *J. Virol.* **68**, 951–958.

Hauber, J., Bouvier, M., Malim, M. H., and Cullen, B. R. (1988). Phosphorylation of the *rev* gene product of human immunodeficiency virus type 1. *J. Virol.* **62**, 4801–4804.

Katz, R. A., and Skalka, A. M. (1990). Control of retroviral RNA splicing through maintenance of suboptimal processing signals. *Mol. Cell. Biol.* **10**, 1805–1810.

Katz, R. A., Kotler, M., and Skalka, A-M. (1988). Cis-acting intron mutations that affect the efficiency of avian retroviral RNA splicing: Implications for mechanisms of control. *J. Virol.* **62**, 2686–2695.

Kim, S., Bym, R., Groopman, J., and Baltimore, D. (1989). Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: Evidence for differential gene expression. *J. Virol.* **63**, 3708–3713.

Kjems, J., and Sharp, P. A. (1993). The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. *J. Virol.* **67**, 769–776.

Kjems, J., Frankel, A. D., and Sharp, P. A. (1992). Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. *Cell* **67**, 169–178.

Krainer, A. R., Conway, G. C., and Kozak, D. (1990). Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.* **4**, 1158–1171.

Linial, M. (1987). Creation of a processed pseudogene by retroviral infection. *Cell* **49**, 93–102.

Lu, X., Heimer, J., Rekosh, D., and Hammarskjold, M-L. (1990). U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus *env* mRNA that remains unspliced. *Proc. Natl. Acad. Sci. USA* **87**, 7598–7602.

Maldarelli, F., Martin, M. A., and Strebel, K. (1991). Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: Novel level of gene regulation. *J. Virol.* **65**, 5732–5743.

Malim, M. H., and Cullen, B. R. (1993). Rev and the fate of pre-mRNA in the nucleus: Implications for the regulation of RNA processing in eukaryotes. *Mol. Cell. Biol.* **13**, 6180–6189.

Malim, M. H., Hauber, J., Le, S. V., Maizel, J. V., and Cullen, B. R. (1989). The HIV-1 *rev trans*-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature (London)* **338**, 254–275.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.

McNally, M., and Beemon, K. (1992). Intronic sequence and 3' splice sites control Rous sarcoma virus RNA splicing. *J. Virol.* **66**, 6–11.

McNally, M., Gontarek, R., and Beemon, K. (1991). Characterization of Rous sarcoma virus intronic sequences that negatively regulate splicing. *Virology* **185**, 99–108.

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**, 7035–7056.

Muesing, M. A., Smith, D. H., Cabradilla, C. D., Benton, C. V., Lasky, L. A., and Capon, D. J. (1985). Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. *Nature (London)* **313**, 450–458.

Niwa, M., McDonald, C. C., and Berget, S. M. (1992). Are vertebrate exons scanned during splice-site selection? *Nature* **360**, 277–280.

Purcell, D. F. J., and Martin, M. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates virus protein expression, replication, and infectivity. *J. Virol.* **67**, 6365–6378.

Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K.,

Ivanoff, L., Petteway, S. R., Jr., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. Complete nucleotide sequence of the AIDS virus, HTLV-III. (1986). *Nature* 313, 277–284.

Reed, R. (1989). The organization of 3' splice-site sequences in mammalian introns. *Genes Dev.* 3, 2113–2123.

Rekosh, D., Nygren, A., Flodby, P., Hammarskjold, M. L., and Wigzell, H. (1988). Coexpression of human immunodeficiency virus envelope proteins and tat from a single simian virus 40 late replacement vector. *Proc. Natl. Acad. Sci. USA* 85, 334–338.

Riggs, N. L., Little, S. J., Richman, D. D., and Guatelli, J. C. (1994). Biological importance and cooperativity of HIV-1 regulatory gene splice acceptors. *Virology* 202, 264–271.

Robert-Guroff, M., Popovic, M., Gartner, S., Markham, P., Gallo, R. C., and Reitz, M. S. (1990). Structure and expression of *tat-, rev-,* and *nef*-specific transcripts of human immunodeficiency virus type 1 in infected lymphocytes and macrophages. *J. Virol.* 64, 3391–3398.

Rosen, C. A., Terwilliger, E., Dayton, A., Sodroski, J. G., and Haseltine, W. A. (1988). Intragenic *cis*-acting art gene-responsive sequences of the human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* 85, 2071–2075.

Schwartz, S., Felber, B. K., Benko, D. M., Fenyo, E-M., and Pavlakis, G. N. (1990). Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. *J. Virol.* 64, 2519–2529.

Schwartz, S., Felber, B. K., and Pavlakis, G. N. (1991). Expression of human immunodeficiency virus type 1 *vif* and *vpr* mRNAs is Rev-dependent and regulated by splicing. *Virology* 183, 677–686.

Schwartz, S., Felber, B. K., and Pavlakis, G. N. (1992). Distinct RNA sequences in the *gag* region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *J. Virol.* 66, 150–159.

Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990). Splice junctions, branch point sites, and exons: Sequence statistics, identification, and applications to genome project. *Methods Enzymol.* 183, 252–278.

Smith, A. J., Cho, M. L., Hammarskjold, M. L., and Rekosh, D. (1990). Human immunodeficiency virus type 1 $Pr55^{gag}$ and $Pr160^{gag-pol}$ expressed from a simian virus 40 late replacement vector are efficiently processed and assembled into virus-like particles. *J. Virol.* 64, 2743–2750.

Smith, C. W. J., Patton, J. G., and Nadal-Ginard, B. (1989). Alternative splicing in the control of gene expression. *Annu. Rev. Genet.* 23, 527–577.

Staffa, A., and Cochrane, A. (1994). The *tat/rev* intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3' splice sites. *J. Virol.* 68, 3071–3079.

Stoltzfus, C. M., and Fogarty, S. J. (1989). Multiple regions in the Rous sarcoma virus *src* gene intron act in *cis* to affect accumulation of unspliced RNA. *J. Virol.* 63, 1669–1676.

Stutz, F., and Rosbash, M. (1994). A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J.* 13, 4096–4104.

Zhang, L., and Stoltzfus, C. M. (1995). A suboptimal *src* 3' splice site is necessary for efficient replication of Rous sarcoma virus. *Virology* 206, 1099–1107.