MOLECULAR AND CELLULAR BIOLOGY, Sept. 1988, p. 3582–3590
0270-7306/88/093582-09$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 8, No. 9

Multiple *cis*-Acting Sequence Elements Are Required for Efficient Splicing of Simian Virus 40 Small-t Antigen Pre-mRNA

XIN-YUAN FU,† JOHN D. COLGAN, AND JAMES L. MANLEY*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 1 February 1988/Accepted 19 May 1988

We have determined the effects of a number of mutations in the small-t antigen mRNA intron on the alternative splicing pattern of the simian virus 40 early transcript. Expansion of the distance separating the small-t pre-mRNA lariat branch point and the shared large T-small t 3′ splice site from 18 to 29 nucleotides (nt) resulted in a relative enhancement of small-t splicing in vivo. This finding, coupled with the observation that large-T pre-RNA splicing in vitro was not affected by this expansion, suggests that small-t splicing is specifically constrained by a short branch point–3′ splice site distance. Similarly, the distance separating the 5′ splice site and branch point (48 nt) was found to be at or near a minimum for small-t splicing, because deletions in this region as small as 2 nt dramatically reduced the ratio of small-t to large-T mRNA that accumulated in transfected cells. Finally, a specific sequence within the small-t intron, encompassing the upstream branch sites used in large-T splicing, was found to be an important element in the cell-specific pattern of early alternative splicing. Substitutions within this region reduced the ratio of small-t to large-T mRNA produced in HeLa cells but had only minor effects in human 293 cells.

One of the most intriguing questions concerning pre-mRNA splicing in higher eucaryotes is how splice sites are selected. The conserved consensus sequences found at splice sites (6, 26) are of course important factors. However, they are unlikely to be the sole determinants, since "cryptic" splice sites, which are sequences in pre-mRNA that resemble a consensus sequence, are not normally used. More interestingly, some pre-mRNAs undergo alternative splicing, which results in the production of different mature mRNAs from a common precursor by use of different splice sites (for reviews, see references 7, 17, 20, and 29). What factors dictate how splice sites are chosen in such pre-mRNAs, especially in those in which splice site selection is regulated?

The view that sequences besides the conserved splice site sequences are involved in alternative splicing has received some experimental support. For example, the alternative splicing patterns of simian virus 40 (SV40) late pre-mRNAs (36) and adenovirus early-region 3 pre-mRNAs (2, 3) were altered in viral mutants containing deletions within exon sequences. In addition, in mutants containing tandem duplications of 5′ and 3′ splice sites of the β-globin first intron, splice site selection in vitro appeared to be influenced by the length of adjacent exon sequences (30). One way in which distal sequences might influence splice site selection is through higher-order RNA structure. In an effort to understand the possible role of RNA secondary structure in alternative splicing, pre-mRNAs containing artificial inverted repeats were studied. The results of these analyses indicated that the alternative splicing patterns of such pre-mRNAs could be influenced by certain RNA secondary structures, both in vitro (34) and in vivo (14). The relative contribution of RNA secondary structure to splice site selection in vivo may be low, however, because in one study (35) such inverted repeats affected splicing at most only slightly.

To study the mechanism of alternative splicing, we have been using the SV40 early gene as a model, because it displays perhaps the simplest form of alternative splicing. Large-T and small-t mRNAs are alternatively spliced from a common pre-mRNA by utilizing either of two 5′ splice sites and a shared 3′ splice site. We showed previously that selection of the 5′ splice site could be influenced by several factors, including the relatively short distance (48 nucleotides [nt]) separating the small-t 5′ splice site and lariat branch site (14). Based on differences in the pattern of alternative splicing observed in HeLa and 293 cells (a human embryonic kidney cell line transformed with the early region of adenovirus 5 [16]), we suggested that cell-specific *trans*-acting factors might be involved in determining the pattern of alternative splicing of this pre-mRNA (14). A possible molecular mechanism for this stems from the observation that, unlike other pre-mRNAs studied previously, large-T mRNA splicing involves multiple lariat branch sites in vitro and in vivo (28). The predominant branch sites used were found to be cell type specific and to correlate well with cell-specific patterns of alternative splicing (12, 28). In this paper we describe in more detail some of the *cis*-acting factors that modulate alternative splicing of SV40 early pre-mRNA, specifically those that do so by influencing the efficiency of small-t mRNA splicing.

MATERIALS AND METHODS

Mutant construction. The wild-type plasmid pSTER contains the entire SV40 early region (nt 294 to 2533) inserted between the *Eco*RI and *Bam*HI sites of pBR322 (21). Oligonucleotide-directed mutagenesis (45) was carried out with a single-stranded M13 derivative containing SV40 early-region sequences from nt 4002 to 5171.

To create the polypyrimidine stretch expansion mutants, plasmid pSVx, which contains a single base substitution (T to A) 3 nt upstream of the 3′ splice site, creating a new *Xba*I site, (13) was cleaved with *Xba*I, blunt-ended by treatment with S1 nuclease, and then ligated with synthetic double-stranded 15-mers. These 15-mers corresponded to the sequence from −15 to −1 relative to the 3′ splice sites of the SV40 early gene itself (pSV-T), of the first intron of the

* Corresponding author.
† Present address: Rockefeller University, New York, NY 10027.

human β-globin gene (pSV-G), or of the first intron in the adenovirus late mRNA leader (pSV-L). All mutants were recloned into pSTER.

Small-t intron deletion mutants were derived from the plasmid pSTER-ill, which contains an inserted *Eco*RI linker at position −41 in the small-t intron (14). pSTER-ill was cleaved with *Eco*RI and digested with exonuclease Bal 31, and the shortened DNA fragments were ligated with *Eco*RI linkers. Deletions extending only downstream of the original *Eco*RI site were obtained by recloning the downstream fragments (*Eco*RI to *Bam*HI, SV40 nt 2533) into *Eco*RI- and *Bam*HI-digested pSTER-ill to generate plasmids pSV14S, pSV32S, pSV16S, and pSV57S. Similarly, pSV61L was also constructed by recloning but by substituting an upstream deleted fragment (*Eco*RI to *Bst*XI, SV40 nt 4759) into the original *Eco*RI site.

The substitution mutant pSV32K was made by cleaving pSV32S with *Eco*RI, followed by filling in the sticky end with the Klenow fragment of DNA polymerase I and religation. pSV32B was constructed by a procedure similar to that used to construct the S series of mutations, except that a longer *Eco*RI linker was used. The relevant nucleotide sequences of all mutants were verified by DNA sequence analysis (24) and are indicated in Fig. 1.

Cell transfection and RNA analysis. The methods used for transfection of HeLa and 293 cells and for extraction and analysis of RNA by S1 nuclease mapping were those described previously (1, 14, 43). Over the course of many experiments, we observed that transfection efficiencies were variable, with fluctuations as great as threefold in a given experiment. This would make comparisons of absolute amounts of SV40 RNA synthesized from different plasmids difficult. However, in these studies we compared only the ratios of small-t to large-T mRNAs produced in single transfections. These ratios were extremely reproducible and independent of transfection efficiency, eliminating potential problems that could have arisen from experimental variations. All autoradiograms used for quantitation were produced without intensifying screens and were scanned with a Gilford model 250 densitometer.

In vitro splicing and analysis of RNA products. The methods of in vitro splicing analysis were exactly as described previously (28). Briefly, HeLa cell nuclear extracts were prepared by a modification of the procedure of Dignam et al. (10), and in vitro splicing was carried out as described (28). Pre-RNAs were prepared from derivatives of the relevant plasmids in which SV40 fragments extending from SV40 nt 4002 to nt 5171 were cloned into the bacteriophage SP6 promoter-containing plasmid SP64 (25). To study small-t splicing, 5′ sequences between nt 5171 and 4698 were deleted from the wild type (pSPHdA) and one of polypyrimidine stretch insertion mutants (pSP-G) to generate pSPHdAΔT and pSP-GΔT, respectively. Debranching analysis was performed as described by Ruskin and Green (31). Splicing products and intermediates were analyzed on denaturing 6% polyacrylamide–urea gels. Small-t introns were purified from gels and analyzed by primer extension as described (Noble et al., submitted for publication). Primers were a 14-mer (wild type) and 15-mer (pSP-G) complementary to the 3′ ends of wild-type and mutant introns, respectively. Extended products from lariat and debranched introns were analyzed on 15% polyacrylamide–urea gels.

RESULTS

Distance between 3′ splice site and small-t mRNA lariat branch point can influence alternative splicing of SV40 early pre-mRNA in vivo. The distance between the pre-mRNA lariat branch point and 3′ splice site appears to be constrained in introns in higher eucaryotes. Most branch points mapped thus far are located between 24 and 37 nt upstream from the 3′ splice site, although distances as short as 18 nt have been reported (17). Because the distance between the 3′ splice site and unique small-t branch site in SV40 early pre-mRNA is only 18 nt (28), we wished to determine whether this factor might influence the alternative splicing pattern of SV40 pre-mRNA, as does the short distance between the small-t 5′ splice site and branch point. To address this question, the mutant pSV-T, which essentially duplicates the wild-type polypyrimidine stretch, was constructed. This mutation increases the distance between the small-t branch site and 3′ splice site to 29 nt (Fig. 1B and below). The effects of this mutation, as well as of others described below, on alternative splicing of SV40 early pre-mRNA were determined by measuring the ratios of small-t to large-T mRNA produced in HeLa and 293 cells. Cells were transfected with wild-type (pSTER or pSTER-B [14, 21]) or mutant plasmids, and RNA extracted from cytoplasmic or nuclear fractions was analyzed by quantitative S1 nuclease mapping (Fig. 2). Cleavage at the 5′ splice site of large-T (254-nt protected fragment) or small-t (534 nt) mRNA was detected by using a 3′-end-labeled DNA probe, as described previously (14). The results obtained were expressed as ratios of small-t to large-T mRNAs, and the effects of a mutation on alternative splicing were determined by comparing mutant and wild-type ratios. Because absolute amounts of RNA were not compared, potential errors due to variations in transfection efficiencies were avoided (see Materials and Methods).

The results of such an analysis with the mutant pSV-T show that the ratio of small-t to large-T mRNA (t/T) in pSV-T transfected HeLa cells was 0.7 in the cytoplasmic fraction and approximately 1.0 in the nuclear fraction (Fig. 2A). These values are considerably higher than the 0.15 ratio observed with both cytoplasmic and nuclear RNA from pSTER-transfected cells, suggesting that the distance between the branch site and 3′ splice site is a limiting factor in small-t mRNA splicing.

Although pSV-T was constructed to prevent changes in the sequence content in the region separating the 3′ splice site and the small-t branch point, it was conceivable that the increased t/T mRNA ratio brought about by the mutation in pSV-T was sequence specific. Indeed, we have shown that changes which alter the pyrimidine content of the SV40 early pre-mRNA polypyrimidine stretch can have significant effects on the alternative splicing of this pre-mRNA (13). To test this possibility, two additional intron expansion mutants were constructed. pSV-G was created by insertion of an 11-nt pyrimidine-rich fragment corresponding to the 3′ splice site of the first intron of the human β-globin gene (19), and pSV-L was created by insertion of sequences from the 3′ splice site of the adenovirus type 2 late mRNA leader first intron (11). The altered sequences in these mutants are shown in Fig. 1B. S1 nuclease mapping of cytoplasmic and nuclear RNA from HeLa cells transfected with these mutants revealed t/T mRNA ratios very similar to, although perhaps slightly higher than, that observed with pSV-T (Fig. 2A). These findings argue that in these mutants the increased distance between the branch site and 3′ splice junction is the factor responsible for the elevated t/T mRNA ratios.

We showed previously that mutants in which the distance between the small-t mRNA 5′ splice site and branch point was increased gave rise to elevated t/T mRNA ratios in 293

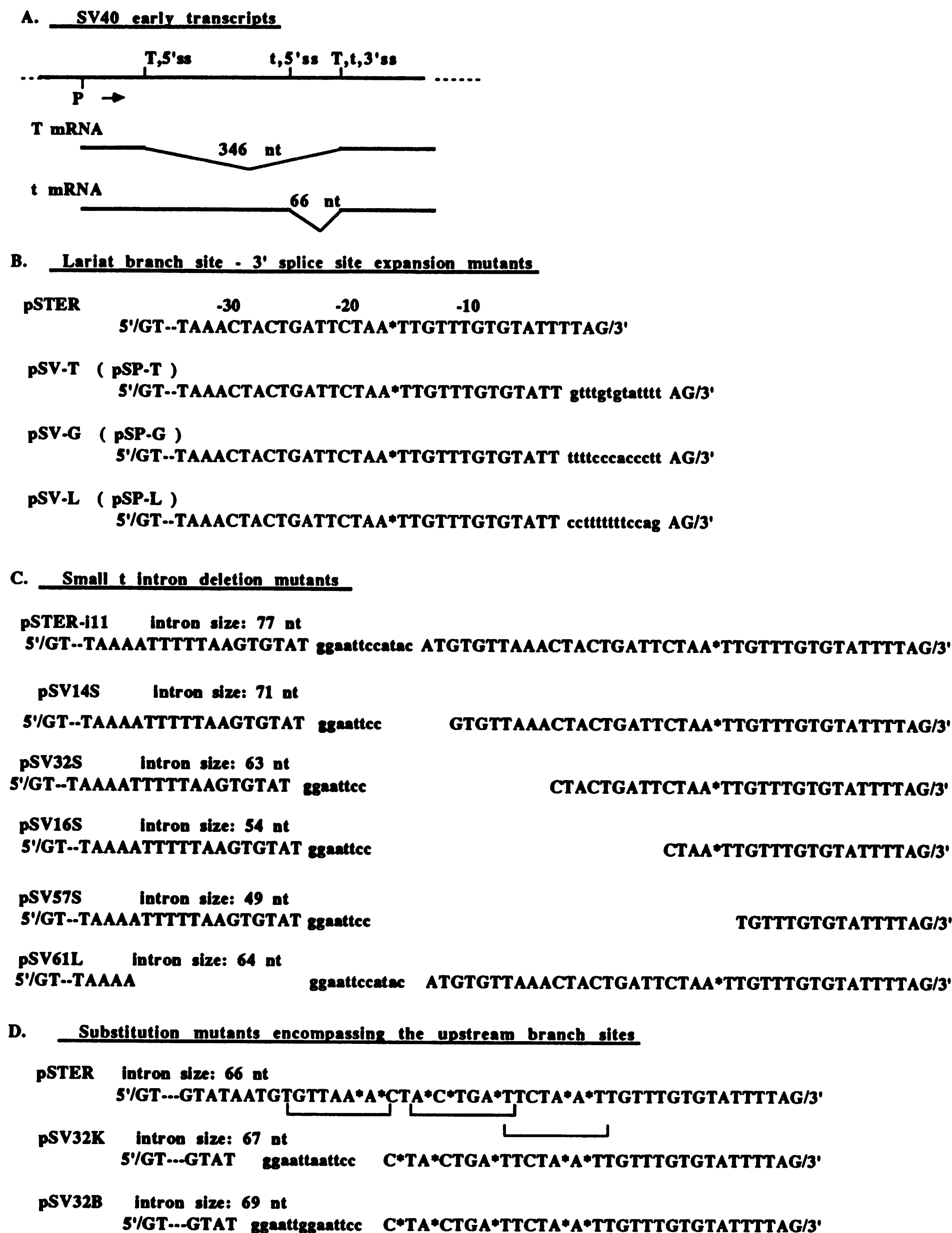

FIG. 1. Nucleotide sequence of small-t intron mutants. (A) Schematic diagram of the SV40 early region and its transcripts. ss, Splice site; carets, introns; P, transcriptional promoter. (B) pSTER is the wild-type plasmid used in these studies and contains the entire SV40 early region (21). The numbers indicate the distance in nucleotides upstream of the 3′ splice site. The small-t mRNA branch site adenosine is indicated by an asterisk. The inserted *Eco*RI linker sequences are represented by lowercase letters. (C) pSTER-ill is an insertion mutant that contains an *Eco*RI linker at position −41 (14). All deletion mutants were made by Bal 31 nuclease digestion from this site. The small-t intron size in each mutant is indicated. Gaps indicate deletions. (D) Multiple branch sites used in large-T mRNA splicing are indicated by asterisks. The three sequences that match the branch site consensus sequence are indicated by brackets. The small-t intron size in each mutant is indicated.

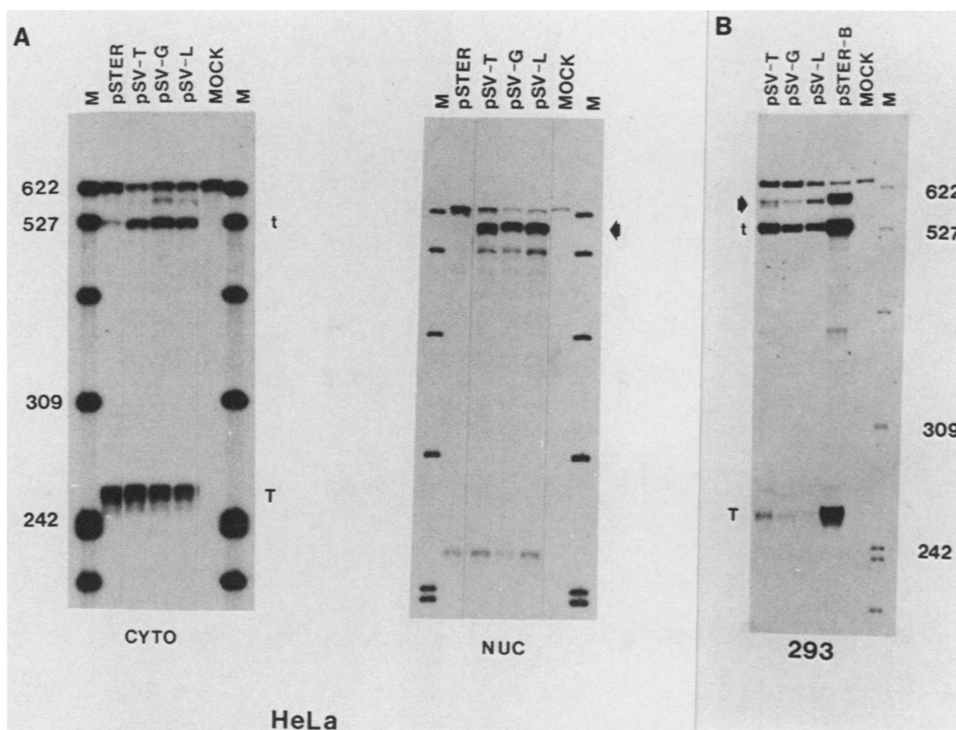

FIG. 2. Effects of lariat branch site-3′ splice site expansion on splice site selection in transfected HeLa cells. Wild-type plasmid pSTER and mutant plasmids were used to transfect HeLa (A) and 293 (B) cells, and RNA from cytoplasmic (CYTO) and nuclear (NUC) (HeLa only) fractions was purified. S1 nuclease analysis was used to detect SV40-specific RNAs, and S1-resistant fragments were resolved by electrophoresis through a 5% polyacrylamide–urea gel. The probe was a 3′-end-labeled DNA fragment (SV40 nt 5171 to 4528), which gave rise to a 254-nt protected fragment for large-T mRNA and a small-t protected fragment of 534 nt. The arrows indicate protected fragments corresponding to mutant or pSTER-B precursors, which resulted from divergence of pre-mRNAs and the wild-type probe. P indicates the position of undigested DNA probe or, in the case of pSTER, unspliced precursor. Lanes M display labeled DNA size markers (in nucleotides). MOCK, RNA obtained from untransfected HeLa cells.

cells as well as in HeLa cells, even though this ratio was already very high in 293 cells transfected with wild-type plasmids (14). To determine whether mutants pSV-T, pSV-G, and pSV-L would show similarly increased t/T mRNA ratios in these cells, 293 cells were transfected with each of the mutants and the t/T ratios were determined by S1 nuclease analysis as above. The results (Fig. 2B) revealed that the t/T mRNA ratios were in fact increased approximately threefold compared with that of a wild-type-plasmid, pSTER-B, to 5/1 with pSV-T and 7/1 with pSV-G and pSV-L. The somewhat increased amount of pSV-T-specific RNA in the experiment shown in Fig. 2B was due to variations in transfection efficiencies, as it was not observed in other experiments. The significantly elevated levels of SV40-specific RNA detected in pSTER-B-transfected cells, however, were observed repeatedly (14) (see below). This plasmid contains an insertion of four *Bam*HI linkers at a site just downstream from the 3′ splice junction, and the mutation was shown previously not to affect the pattern of alternative splicing of SV40 pre-mRNA (14). The plasmid's utility in the experiments described here stems largely from the fact that a wild-type DNA probe will distinguish, in S1 nuclease analysis, between undigested probe and unspliced pSTER-B pre-mRNA, due to the sequence divergence that occurs at the site of the insertion in pSTER-B. The experiment shown in Fig. 2B revealed significant amounts of unspliced RNA in all samples. This presumably resulted from nuclear leakage, which we have found occurs relatively frequently with 293 cells. However, the relative amounts of unspliced precursors were similar, consistent with the notion that the polypyrimidine expansion mutations do not significantly affect the overall efficiency of the splicing reaction.

Effects of the polypyrimidine region expansion mutations on SV40 early pre-RNA splicing in vitro. Several previous studies have shown that splicing of SV40 early pre-mRNA in vitro yields almost exclusively large-T RNA and barely detectable levels of small-t RNA (5, 23, 28, 41, 44). In addition, SV40 early pre-RNA is spliced with a lower efficiency than several other well-studied pre-RNAs, including adenovirus late first-leader pre-RNA (Ge and Manley, unpublished results). To determine whether either of these properties might be due to the short distance separating the 3′ splice site and branch point, we transferred the appropriate sequences from pSV-T, pSV-G, and pSV-L into the bacteriophage SP6 promoter-containing plasmid SP64 (25) to create the mutants pSP-T, pSP-G, and pSP-L. Capped pre-RNAs were transcribed from these plasmids as well as from the wild-type, pSPHdA, and processed in a HeLa cell nuclear extract (10) as described previously (28). Incubations were terminated after 40 min or 3 h, and RNA products were analyzed by polyacrylamide gel electrophoresis.

Two splicing products, spliced large-T RNA and large-T lariat intron, and two intermediates, large-T 5′ exon and lariat large-T intron-3′ exon, were detected with wild-type and mutant pre-RNA substrates (Fig. 3). Identification of these products and intermediates was based on several properties, including their relative mobilities compared with size markers, debranching analysis of the lariat large-T intron and large-T intron 3′-exon (data not shown), and additional characterizations described previously (27, 28).

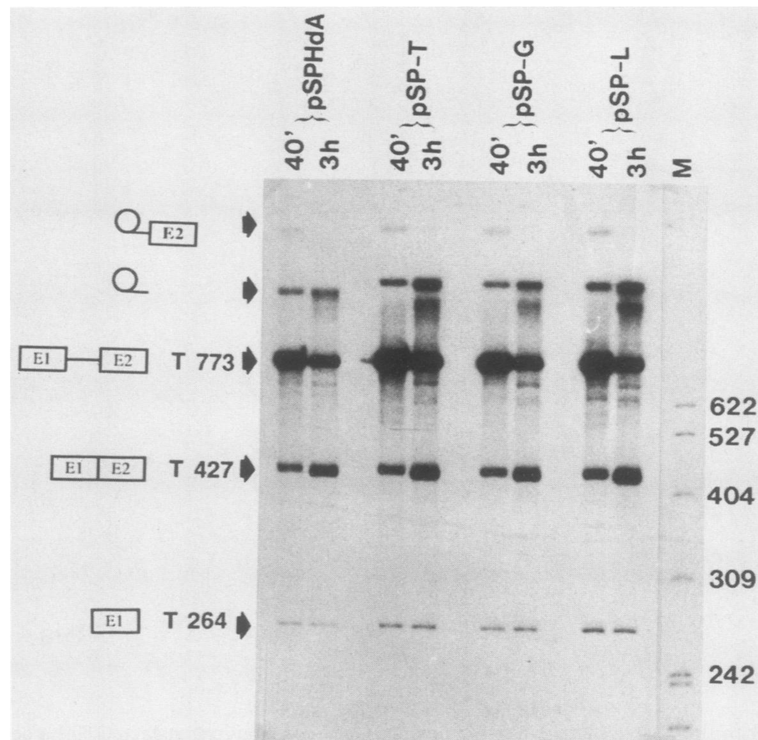

FIG. 3. Effects of lariat branch site-3′ splice site expansion mutations on in vitro splicing. Runoff transcripts from the wild-type plasmid pSpHdA and mutant plasmids were synthesized with SP6 RNA polymerase and extended to an *StyI* site at SV40 nt 4409. Capped pre-RNAs were processed in HeLa nuclear extracts for 40 min or 3 h, as described in Materials and Methods. RNAs were purified, and splicing products and intermediates were resolved by electrophoresis in a denaturing 6% polyacrylamide–urea gel. The corresponding RNA structures and sizes of the RNA species are shown schematically on the left. Boxes indicate exons, and lines show introns. Lane M, DNA size markers (in nucleotides).

The amounts of spliced large-T RNA and remaining unspliced precursor in each sample were determined by densitometric scanning and expressed as a ratio of spliced RNA to precursor. With wild-type pSPHdA, this ratio was 0.58 after a 3-h incubation. A virtually identical ratio was obtained from the products of pSP-T pre-RNA splicing, although the ratio was increased slightly by the other two mutations, to 0.77 with pSP-G pre-RNA and to 0.93 with pSP-L pre-RNA. These results indicate that large-T RNA splicing in vitro is not limited by the distance separating the branch point and the 3′ splice site, consistent with the in vivo results, but that the sequence composition of this region can influence splicing efficiency, consistent with previous results (13).

Although the efficiency of large-T pre-RNA splicing was high in the experiment shown in Fig. 3, no small-t RNA splicing products or intermediates could be detected in any of the processed pre-RNAs. The faint bands located just below the pre-RNA band appear to have resulted from degradation, as they were not detected reproducibly and were generated in the absence of ATP (H. Ge, unpublished results). Therefore, the lack of detectable small-t RNA splicing in vitro is unlikely to be a result of the short distance between the 3′ splice site and branch point. In the mutant samples obtained after 3-h incubations, a species migrating just ahead of the large-T lariat intron was detected. This RNA appears to have been created by partial degradation of the large-T intron by a 3′ exonuclease, because it could be converted to a species migrating slightly faster than the linear large-T intron by enzymatic debranching (data not shown). This RNA was not detected among the products of pSPHdA pre-RNA processing, suggesting that the longer tails of the lariat introns produced by processing of the insertion mutant pre-RNAs were less protected from the 3′ exonuclease by splicing complexes.

We assumed that the expansion mutants analyzed above continued to use the same small-t branch sites used in splicing of the wild-type pre-RNA, resulting in an increased 3′ splice site-to-branch point distance. However, it was possible that a cryptic branch site, closer to the 3′ splice site, may have been used in these mutants. A candidate is the A found in all three mutants at position −18 (Fig. 1B). Although the sequence surrounding this A does not match the lariat branch site consensus (17, 18, 29), its use would have indicated that the insertions had not, in fact, increased the 3′ splice site-branch point distance. Therefore, the branch point(s) used in small-t splicing in vitro of one of the mutants (pSP-G) was determined. For this, SV40 sequences upstream of nt 4698, including the large-T 5′ splice site, were deleted from pSP-G. It has been shown previously that such deletions result in an activation of small-t pre-RNA splicing in vitro (23, 28, 41). Pre-RNAs were synthesized from this mutant DNA, as well as from a similarly deleted wild-type template, and used as substrates for in vitro splicing. Small-t introns were purified, and the branch points were mapped by primer extension of lariat and debranched introns with appropriate DNA primers (see Materials and Methods). The results (Fig. 4) show that, as observed previously (28; Noble et al., submitted), branching in the wild-type occurred primarily to position −18, with a small amount apparently occurring also at the adjacent A at position −19. Branching in the pSP-G small-t intron could be detected only to the same nucleotides, which in this mutant correspond to posi-

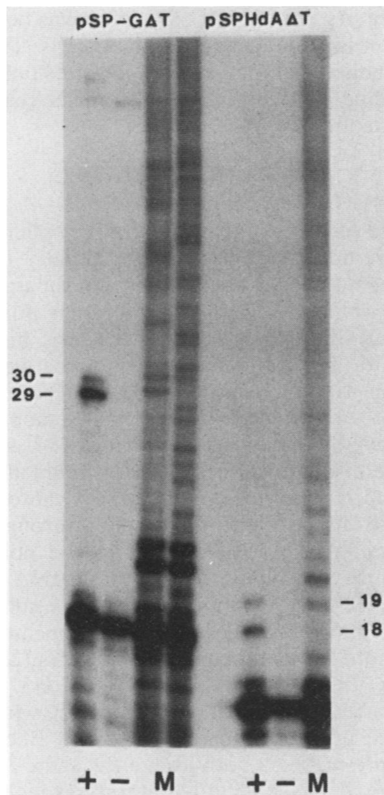

FIG. 4. Small-t intron branch site is shifted 11 nt upstream by an insertion between the 3′ splice site and branch point. Pre-RNAs synthesized from pSPHdAΔT and pSP-GΔT were incubated in nuclear extract, and the small-t introns produced were purified. Lariat (+) or debranched (−) introns were used as templates for reverse transcription with the primers described in Materials and Methods. The numbers on the sides indicate the distance (in nucleotides) upstream of the respective 3′ splice site to which the primer extension products mapped the corresponding branch points. Markers (lanes M) were dideoxy sequencing-type reactions that utilized the appropriate unprocessed-pre RNAs as templates and the same oligonucleotides used in the branch point mapping as the primers.

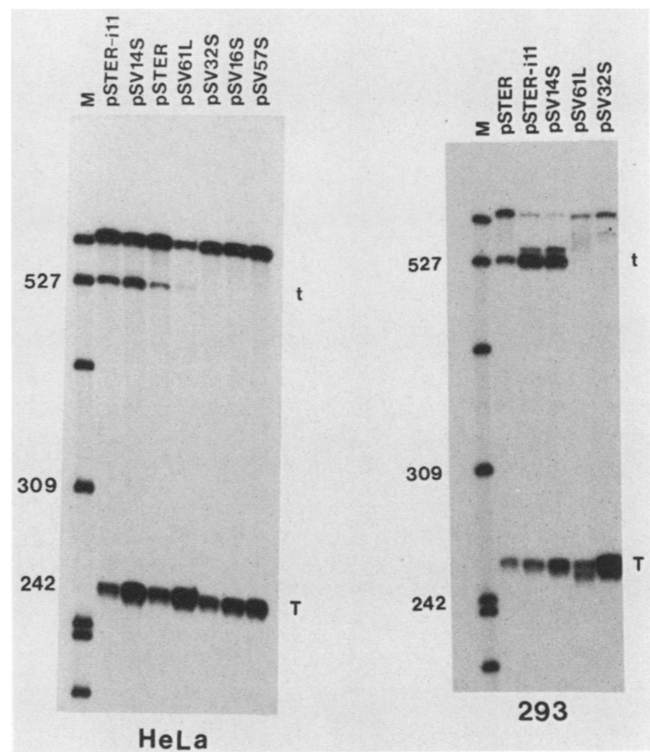

FIG. 5. Deletion mutations in the small-t intron abolish small-t but not large-T mRNA splicing. Wild-type and small-t intron deletion mutant DNAs were used to transfect HeLa and 293 cells. Cytoplasmic RNA was purified, and SV40-specific RNAs were detected by S1 nuclease mapping with either a 3′-end-labeled probe prepared from wild-type DNA or homologous probes prepared from pSV61L and pSV32S. The fluctuations in the amounts of total SV40-specific RNA were due to variations in transfection efficiencies.

tions −29 and −30. Thus, the insertion in pSP-G (and presumably in the other expansion mutants) did increase the 3′ splice site-branch point distance by the predicted 11 nt.

Distance between the small-t mRNA 5′ splice site and lariat branch point cannot be reduced without inhibiting small-t mRNA splicing. Our previous studies indicated that the number of nucleotides separating the small-t 5′ splice site and branch point (48 nt) is a limiting factor in small-t mRNA splicing (14). Mutations that increased this distance by 11 or more nt elevated the t/T mRNA ratio over 10-fold in both HeLa and 293 cells. To analyze this size constraint further, we constructed a series of deletion mutants from plasmid pSTER-ill, which contains an *Eco*RI linker inserted at a position −42 nt from the 3′ splice site, resulting in a 59-nt separation between the 5′ splice site and branch point (14). Deletions were made by Bal 31 nuclease digestion from this *Eco*RI restriction site (see Materials and Methods). Mutants pSV14S, pSV61L, pSV32S, pSV16S, and pSV57S (Fig. 1C), which contained 5′ splice site-to-branch point distances of 53, 46, 45, 36, and 31 nt, respectively, were selected for analysis.

Mutant and wild-type plasmids were used to transfect HeLa and 293 cells, and cytoplasmic RNA was extracted and analyzed by S1 nuclease mapping as described above, except that homologous DNA probes were used for all mutants in which the size of the intron was reduced. As observed previously (14), the mutation in pSTER-ill resulted in a 5- to 10-fold increase in the t/T mRNA ratio in both HeLa and 293 cells (Fig. 5). pSV14S, which contains a smaller, 5-nt insertion, resulted in only a slight (approximately twofold) increase in the ratio. The mutations that reduced the size of the small-t intron all dramatically inhibited small-t mRNA splicing. In HeLa cells, the mutation in pSV61L (2-nt deletion) reduced the t/T mRNA ratio to approximately 1/30, representing at least a sixfold reduction compared with pSTER. Interestingly, approximately the same ratio was observed in 293 cells. However, because the wild-type ratio is significantly greater in these cells, this represents an approximately 50-fold relative reduction. (The doublet seen at the position corresponding to large-T mRNA in the pSV61L-transfected 293 cells is apparently the result of S1 nuclease overdigestion, as it was not observed in other experiments.) The mutant pSV32S (3-nt deletion) produced at most trace amounts of small-t mRNA in both cell types, while the mutants with larger deletions produced no small-t mRNA (results in 293 cells not shown). These findings suggest that the 48 nt separating the small-t 5′ splice site and lariat branch point are at or near a minimum length.

Specific intron sequence is critical for small-t mRNA splicing in HeLa cells but not in 293 cells. SV40 early pre-mRNA splicing is unusual in that splicing to produce large-T mRNA

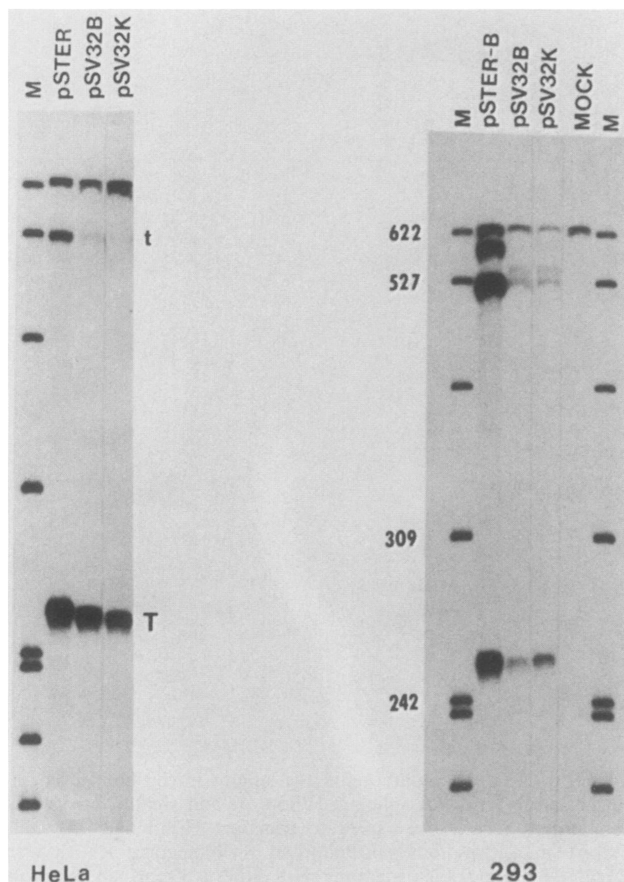

FIG. 6. Specific intron sequence is required for small-t mRNA splicing in HeLa cells but not in 293 cells. Cytoplasmic RNA from HeLa or 293 cells transfected with the indicated plasmids was used for S1 nuclease analysis with a 3'-end-labeled wild-type DNA probe. The bands above the small-t mRNA derived-band in mutant samples resulted from S1 nuclease digestion of mismatches between mutant pre-mRNAs and the wild-type probe.

involves utilization of multiple lariat branch points (28). These sites are situated between 18 and 32 nt upstream of the 3' splice site and have been proposed to play a role in controlling the cell-specific alternative splicing of this pre-mRNA. Small-t splicing, in constrast, appears to use primarily a single branch site, at position −18. To begin to study how these branch site sequences might function in alternative splicing, we constructed two related substitution mutations, pSV32K and pSV32B, in which the upstream-most branch sites are replaced with *Eco*RI linker sequences (see Fig. 1D and Materials and Methods). In addition to disrupting the upstream branch site, these mutations expand the size of the small-t intron slightly, to 67 and 69 nt, respectively. Mutant and wild-type plasmids were introduced into HeLa and 293 cells, and cytoplasmic RNA was subjected to S1 nuclease analysis as above. The results (Fig. 6) were surprising. Small-t mRNA splicing was significantly inhibited in HeLa cells, but only very slightly reduced in 293 cells. Specifically, the mutation in pSV32B reduced the t/T mRNA ratio by a factor of approximately 6 in HeLa cells, but only at most 2 in 293 cells. The effects of the pSV32K mutation were more striking, perhaps because of the smaller size of the small-t intron (see Discussion). Small-t mRNA splicing in HeLa cells was barely detectable, but the t/T mRNA ratio was reduced by only a factor of 3 in 293 cells. We estimate that the inhibitory effect of this mutation was between 5- and 10-fold greater in HeLa cells than in 293 cells. Thus, destruction of a sequence that includes branch sites only for large-T mRNA splicing can inhibit small-t mRNA splicing in a cell-specific manner.

DISCUSSION

We showed previously that the relative splicing efficiency of small-t pre-mRNA could be substantially increased by increasing the number of nucleotides that separate the small-t mRNA 5' splice site and lariat branch point (14). We suggested that this reflects a relief of steric hindrance that interferes with productive interactions between required small nuclear ribonucleoproteins (snRNPs) and the pre-mRNA. The observed relative enhancement of small-t mRNA splicing by mutations increasing the 3' splice site-to-branch point distance described here might also reflect a relief from steric hindrance. Specifically, there is considerable evidence that U5 snRNP, perhaps through a protein-RNA interaction (9, 15, 38), binds to the polypyrimidine stretch-3' splice site region during pre-mRNA splicing (8). Because the 18-nt separation between 3' splice site and branch point in SV40 early pre-mRNA appears to be at or near a minimum for mammalian pre-mRNAs (28), it may be that a steric hindrance exists between U5 snRNP and another splicing factor, such as U2 snRNP, which interacts with the lariat branch point region (9, 32). Because small-t and large-T pre-mRNAs both utilize the same 3' splice site, it might have been expected that both splices would be equally affected by the insertion mutations. We suggest two reasons why only the small-t splice was affected. First, although a fraction of large-T splicing does use the adenosine at position −18 as a branch point, a number of other upstream sites are also utilized (28). Such upstream branching might well not be subject to steric hindrance. Second, because small-t but not large-T splicing is also limited by the 5' splice site-to-branch point distance, it may be that this additional constraint compounds the problems that arise from the short branch site-to-3' splice site distance, perhaps by hindering required snRNP-snRNP or pre-mRNA-snRNP interactions.

The experiments presented here also establish a precise value for the minimum number of nucleotides that must separate the small-t 5' splice site and lariat branch point. Reduction of this number from 48 to 46 nt severely inhibited splicing, so that small-t mRNA was barely detectable. Although the mutations analyzed were not simple deletions, but rather substitutions, we do not believe that the inserted sequence (an *Eco*RI linker) had any deleterious effects on splicing, because its insertion in the absence of deletions actually enhanced small-t mRNA splicing (e.g., mutant pSV14S; see also reference 14). Likewise, the sequences deleted in the mutant pSV6IL (net 2-nt deletion) are not likely to be important, because the effects of this mutation could be completely suppressed by expanding the intron by 4 nt (by filling in the *Eco*RI sticky ends; results not shown). In contrast, the other deletion mutations (Fig. 1C) may, in addition to reducing the intron size, have affected sequences required for small-t mRNA splicing, at least in HeLa cells (see below).

Other studies have revealed minimum 5' splice site-branch point distance requirements of between 47 and 60 nt in the human β-globin first intron (33), between 43 and 55 nt in the rabbit β-globin large intron (42), and approximately 41 nt in the adenovirus E1a intron (40). These studies support our

conclusion that the 48-nt distance in the small-t intron is or is very nearly a minimum. As discussed previously, this probably reflects steric hindrance that prevents functional interactions between *trans*-acting splicing factors (e.g., U1 and U2 snRNPs) and the pre-mRNA. An interesting exception to this rule is the small-t mRNA intron in polyomavirus. The 5′ splice-to-branch point distance is only 29 nt, yet this intron is spliced with moderate efficiency in vivo (39) and at detectable levels in vitro (J. Noble, unpublished data).

The data presented in this paper provide some insight into the *cis*-acting factors that are responsible for the cell-specific differences in SV40 early pre-mRNA alternative splicing that we observed previously in 293 cells and HeLa cells (14). One possibility that we had considered was that a *trans*-acting factor in 293 cells might be less refractory to the size constraints that limit small-t pre-mRNA splicing (22). However, this does not appear to be the case. We showed previously that increasing the small-t 5′ splice site-to-branch point distance led to the same relative increases in the t/T mRNA ratio in both cell types, arguing that the short distance that separates these sites in wild-type SV40 early pre-mRNA is a limiting factor in both cell types. This view is strengthened by the experiments presented here, which show that a 2-nt decrease in intron size greatly reduced the t/T mRNA ratio. It is interesting that the ratio was the same (approximately 1/30) in both cell type, because this represents an approximately 10-fold greater effect relative to that of the wild type in 293 cells. This most likely indicates that interactions that are constrained by the shortened small-t 5′ splice site-branch point distance constitute a rate-limiting step in small-t pre-mRNA splicing and prevent the enhancement normally observed in 293 cells. Increasing the branch point-to-3′ splice site distance resulted in similar increases in the t/T mRNA ratios in both cell types. This finding argues that the enhancement of small-t splicing in 293 cells is not due to the existence of factors present in these cells that are not constrained by this unusually short distance.

We proposed previously that the multiple branch sites used in large-T pre-mRNA splicing and the cell-specific utilization of predominant branch points are important factors in cell-specific alternative splicing of SV40 early pre-mRNA (28). This idea might imply that the *cis*-acting sequences that mediate controlled alternative splicing are located within this region, i.e., between nt −17 and −37 relative to the 3′ splice site. The cell-specific inhibition of small-t pre-mRNA splicing observed with the substitution mutants pSV32B and pSV32K, in which sequences between −30 and −42 were replaced, supports this view. We do not believe that this inhibition was due to possibly trivial effects, such as changes in the structure of the pre-mRNA brought about by the inserted *Eco*RI linker, because a similar substitution replacing sequences between −38 and −42 (pSV14S) had no substantial effects on splicing and because insertions of one or six *Eco*RI linkers or unrelated pBR322 sequences at position −41 all had the same effect on splicing, which was to increase the t/T mRNA ratio (14).

The effect of the mutations in pSV32B and -K, which was to inhibit small-t mRNA splicing dramatically in HeLa cells but much less so in 293 cells, were initially somewhat surprising considered in the context of our previous findings on branchpoint utilization: the sequences removed are used as a branch site only in large-T pre-mRNA splicing and constitute the predominant branch points in 293 nuclear extracts but not in HeLa extracts (28). However, we propose the following model, which explains our current results in light of these previous findings.

The experiments presented here and previously (14) have established that small-t mRNA splicing is severely constrained by the distance separating the 5′ splice site and lariat branch point. To compensate for this, and to allow small-t splicing to compete successfully with large-T splicing, it appears that the small-t intron has a high affinity for splicing factors, the U1 and U2 snRNPs. The small-t 5′ splice site is complementary to the 5′ end of U1 RNA at 9 of 10 positions and has been shown to bind U1 snRNPs with high affinity (37). Likewise, we have presented evidence that the SV40 early branch site region has a high affinity for U2 snRNPs (27; Noble et al., submitted; Z. Pan, unpublished data). Because the branch site consensus sequence (17, 18, 29) is relatively loose, we believe that this high affinity is brought about by the presence of three, rather than the usual one, consensus sequences in this region (Fig. 1D). We propose that this high-affinity branch site region is important in its entirety for recruitment of U2 snRNPs for small-t pre-mRNA splicing, even though small-t branching occurs only at the most-downstream branch site. This provides an explanation for the observed relative inhibition of small-t splicing detected in HeLa cells. In 293 cells, the same considerations apply, but we believe that another factor is also important. Large-T branching in these cells occurs predominantly to the upstream positions (28), which are destroyed in pSV32K and -B. Thus, we suggest that utilization of these branch sites is more important for efficient large-T pre-mRNA splicing in 293 cells than in HeLa cells. The combination of these two effects, a reduced efficiency of small-t splicing due to lowered affinity of U2 snRNPs for the small-t intron and a decreased efficiency of large-T splicing due to the destruction of preferred branch sites, may effectively cancel each other out, leading to the observed small effect of these mutations on the t/T mRNA ratio in 293 cells.

ACKNOWLEDGMENTS

We are grateful to J. Noble and Z. Pan for many useful discussions and for communication of their unpublished results. We thank M. Kapcznski for technical assistance and W. Weast for preparing the manuscript.

This work was supported by Public Health Service grant CA33620 from the National Institutes of Health.

LITERATURE CITED

1. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
2. **Bhat, B. M., H. A. Brady, and W. S. M. Wold.** 1985. Deletion mutants that alter differential RNA processing in the E3 complex transcription unit of adenovirus. J. Mol. Biol. **190**:543–557.
3. **Bhat, B. M., and W. S. M. Wold.** 1987. A small deletion distant from a splice or polyadenylation site dramatically alters pre-mRNA processing in region E3 of adenovirus. J. Virol. **61**:3938–3945.
4. **Bindereif, A., and M. R. Green.** 1986. Ribonucleoprotein complex formation during pre-mRNA splicing in vitro. Mol. Cell. Biol. **6**:2582–2592.
5. **Black, D. L., and J. A. Steitz.** 1986. Pre-mRNA splicing *in vitro* requires intact U4/U6 small nuclear ribonucleoprotein. Cell **46**: 697–704.
6. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
7. **Breitbart, R. E., A. Andreadis, and B. Nadal-Ginard.** 1987. Alternative splicing: a ubiquitous mechanism for generation of multiple protein isoforms from single genes. Annu. Rev. Biochem. **56**:467–495.
8. **Chabot, B., D. L. Black, D. M. LeMaster, and J. A. Steitz.** 1985.

The 3′ splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. Science **230**:1344–1349.
9. **Chabot, B., and J. A. Steitz.** 1987. Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes. Mol. Cell. Biol. **7**:281–293.
10. **Dignam, J. D., R. M. Lebowitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
11. **Flint, S. J., and T. R. Broker.** 1981. Lytic infection by adenoviruses, p. 443–546. *In* J. Tooze (ed.), DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Fradin, A., R. Jove, C. Hemenway, H. D. Keiser, J. L. Manley, and C. Prives.** 1984. Splicing pathways of SV40 mRNAs in *X. laevis* oocytes differ in the requirements for snRNPs. Cell **37**: 927–936.
13. **Fu, X.-Y., H. Ge, and J. L. Manley.** 1988. The role of the polypyrimidine stretch at the SV40 early pre mRNA 3′ splice site in alternative splicing. EMBO J. **7**:809–817.
14. **Fu, X.-Y., and J. L. Manley.** 1987. Factors influencing alternative splice site utilization in vivo. Mol. Cell. Biol. **7**:738–748.
15. **Gerk, V., and J. A. Steitz.** 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3′ splice site of premessenger RNA. Cell **47**:973–984.
16. **Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. **36**:59–72.
17. **Green, M. R.** 1986. Pre-mRNA splicing. Annu. Rev. Genet. **20**: 671–706.
18. **Keller, E. B., and W. A. Noon.** 1984. Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. Proc. Natl. Acad. Sci. USA **81**:7417–7420.
19. **Lawn, R., A. Efstratiadis, L. O'Connell, and T. Maniatis.** 1980. The nucleotide sequence of the human β-globin gene. Cell **21**: 647–651.
20. **Left, S. E., M. G. Rosenfeld, and R. M. Evans.** 1986. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Annu. Rev. Biochem. **55**:1091–1117.
21. **Lewis, E. D., and J. L. Manley.** 1985. Repression of simian virus 40 early transcription by viral DNA replication in human 293 cells. Nature (London) **317**:172–175.
22. **Manley, J. L., J. C. S. Noble, M. Chaudhuri, X.-Y. Fu, T. Michaeli, and C. Prives.** 1986. The pathway of SV40 early mRNA splicing. Cancer Cells **4**:259–265.
23. **Manley, J. L., J. C. S. Noble, X.-Y. Fu, and H. Ge.** 1987. Factors that influence alternative splice site selection *in vitro*, p. 97–112. *In* M. Inouye and V. Dudock (ed.), Molecular biology of RNA: new perspectives. Academic Press, New York.
24. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
25. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 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.
26. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
27. **Noble, J. C. S., C. Prives, and J. L. Manley.** 1986. *In vitro* splicing of simian virus 40 early pre-mRNA. Nucleic Acids Res. **14**:1219–1235.
28. **Noble, J. C. S., Z. Pan, C. Prives, and J. L. Manley.** 1987. Splicing of SV40 early pre-mRNA to large T and small t mRNAs utilizes different patterns of lariat branch sites. Cell **50**:227–236.
29. **Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. Seiler, and P. A. Sharp.** 1986. Splicing of messenger RNA precursors. Annu. Rev. Biochem. **55**:1119–1150.
30. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46**:681–690.
31. **Ruskin, B., and M. R. Green.** 1985. An RNA processing activity that debranches RNA lariats. Science **229**:135–140.
32. **Ruskin, B., and M. R. Green.** 1985. Specific and stable intron-factor interactions are established early during *in vitro* pre-mRNA splicing. Cell **43**:131–142.
33. **Ruskin, B., J. M. Greene, and M. R. Green.** 1985. Cryptic branch point activation allows accurate *in vitro* splicing of human β-globin intron mutants. Cell **41**:833–844.
34. **Solnick, D.** 1985. Alternative splicing caused by RNA secondary structure. Cell **43**:667–676.
35. **Solnick, D., and S. Lee.** 1987. Amount of RNA secondary structure required to induce an alternative splice. Mol. Cell. Biol. **7**:3194–3198.
36. **Somasekhar, M. B., and J. E. Mertz.** 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. Nucleic Acids Res. **13**:5591–5609.
37. **Tatei, T., K. Takemura, H. Tauaka, T. Masaki, and Y. Oshima.** 1987. Recognition of 5′ and 3′ splice site sequences in pre-mRNA studied with a filter binding technique. J. Biol. Chem. **262**:11667–11674.
38. **Tazi, J., C. Alibert, J. Temsamani, I. Reveilland, G. Cattiala, C. Brunnel, and P. Jeanteur.** 1986. A protein that specifically recognizes the 3′ splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. Cell **47**:755–766.
39. **Treisman, R., U. Novak, J. Favaloro, and R. Kamen.** 1981. The structures of the spliced mRNAs encoding polyoma virus early region proteins. J. Mol. Appl. Genet. **1**:183–192.
40. **Ulfendahl, P. J., U. Petterson, and G. Akusjarvi.** 1985. Splicing of the adenovirus-2 EIA 13S mRNA requires a minimal intron length and specific signals. Nucleic Acids Res. **13**:6299–6315.
41. **van Santen, V. L., and R. A. Spritz.** 1986. Alternative splicing of SV40 early pre-mRNA *in vitro*. Nucleic Acids Res. **14**:9911–9926.
42. **Wieringa, B., E. Hofer, and C. Weissman.** 1984. A minimal intron length but no specific internal sequence is required for splicing the large rabbit β-globin intron. Cell **37**:915–925.
43. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA mediated transfer of adenosine phosphoribosyl transferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA **76**:1373–1376.
44. **Zhuang, Y., H. Leung, and A. M. Weiner.** 1987. The natural 5′ splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. Mol. Cell. Biol. **7**:3018–3020.
45. **Zoller, M. J., and M. Smith.** 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Methods Enzymol. **100**:468–500.