MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 4395–4405
0270-7306/88/104395-11$02.00/0

Vol. 8, No. 10

Comparison of Intron-Dependent and Intron-Independent Gene Expression

ANDREW R. BUCHMAN† AND PAUL BERG*

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

Received 18 April 1988/Accepted 11 July 1988

Recombinant simian virus 40 viruses carrying rabbit β-globin cDNA failed to express the β-globin sequence unless an intron was included in the transcription unit. The addition of either β-globin IVS1 or IVS2 caused a 400-fold increase in RNA production. Stable β-globin RNA production required sequences in IVS2 that were very close to the splice sites and that coincided with those needed for mRNA splicing. In addition to the recombinant viruses, intron-dependent expression was observed with both replicating and nonreplicating plasmid vectors in short-term transfections of cultured animal cells. Unlike transcriptional enhancer elements, IVS2 failed to increase stable RNA production when it was placed downstream of the polyadenylation site. Using a plasmid vector system to survey different inserted sequences for their dependence on introns for expression, we found that the presence of IVS2 stimulated the expression of these sequences 2- to 500-fold. Sequences from the transcribed region of the herpes simplex virus thymidine kinase gene, a gene that lacks an intervening sequence, permitted substantial intron-independent expression (greater than 100-fold increase) in the plasmid vector system.

The occurrence of intervening sequences (or introns) within eucaryotic genes poses many interesting questions regarding their role in gene expression. The ability to splice RNA transcripts differentially permits great flexibility in the design and expression of eucaryotic genes. Thus, adenovirus expresses a battery of proteins that is involved in capsid formation from a single viral promoter by splicing a common mRNA leader to each of the coding sequences (63). Slightly different patterns of splicing are used to express some cellular genes in different cell types (11, 29, 48). In this way, closely related proteins are produced that differ only in a specific region. In some instances, intervening sequences are the sites where other regulatory elements reside. For example, sequence elements that direct specific DNA rearrangements (52) or that regulate transcription (2, 15, 34, 46) occur in introns.

Some of the earliest studies of eucaryotic vectors uncovered another aspect of introns which still remains unresolved: an apparent necessity for RNA splicing in the production of stable RNA. The first simian virus 40 (SV40) recombinant viruses, which were created before the discovery of splicing, gave somewhat disappointing results in that the foreign sequences were not expressed (17, 18, 25). This failure was later correlated with the absence of introns in these constructions, and it was shown that the deficiency could be corrected by including an intervening sequence from either a viral or cellular gene (23, 24, 41). At that time, it was proposed that splicing is generally required for gene expression (22, 24). Since then, numerous eucaryotic genes have been found that lack intervening sequences (1, 28, 42). Furthermore, while attempting to engineer the expression of different sequences in animal cells, many investigators achieved adequate levels of expression without including introns in their constructions (10, 21, 50, 54). This has left the impression that the requirement for RNA splicing is either a small effect or is restricted to isolated situations.

* Corresponding author.

† Present address: Department of Cell Biology, Stanford University School of Medicine, Stanford, CA 94305.

In the course of our studies of the transcription of the SV40 early region, we noticed that recombinants containing rabbit β-globin DNA failed to express the β-globin sequence unless an intron was included in the transcription unit (7, 8). The size of this effect was unexpectedly large; a difference of several hundred-fold was observed between recombinants containing or lacking an intervening sequence. The behavior of the β-globin recombinants provided an opportunity to explore the basis for intron-dependent gene expression. Two introns, IVS1 and IVS2, interrupt the coding region of the rabbit β-globin gene (26). The nucleotides that are required for accurate processing of these introns have been defined in detail from extensive studies of mutant genes, which were analyzed both in vivo and in vitro (44, 53, 59, 60, 62). Results of these experiments revealed that the sequences required for mRNA splicing are very close to the actual splice sites and are primarily within the intervening sequence. For example, only 6 bases downstream of the 5′ splice site and 24 bases upstream of the 3′ splice site are needed for splicing of IVS2 (59). In our studies, we used SV40-rabbit β-globin recombinants to investigate the relationship between the sequences that are involved in mRNA splicing and those that are required for stable RNA synthesis. A set of specially designed plasmids was constructed to systematically survey different inserted sequences for their dependence on introns for expression. These vectors were used to compare directly the properties of sequences from genes that have introns with sequences from genes that do not.

MATERIALS AND METHODS

Enzymes and reagents. Restriction endonucleases were purchased from New England BioLabs, Inc. (Beverly, Mass.), Bethesda Research Laboratories, Inc. (Gaithersburg, Md.), or Boehringer Mannheim Biochemicals (Indianapolis, Ind.). λ exonuclease was a gift from R. T. White, and T4 DNA ligase was provided by S. Scherer. T4 polynucleotide kinase and DNA ligase were purchased from P-L Biochemicals, Inc. (Milwaukee, Wis.). [^{3}H]leucine, [^{3}H]valine, $^{32}P_i$, [α-^{32}P]dATP, and [γ-^{32}P]ATP were from Dupont, NEN Research Products (Boston, Mass.).

Animal cells and viruses. The origin and maintenance of the CV1 and CV1P cell lines and the procedures for propagating SV40 and recombinants have been described previously (35). CV1, CV1P, and COS (16) cells were maintained in Dulbecco modified Eagle medium containing penicillin, streptomycin, and 5% newborn calf serum. The construction and characterization of the virus recombinants SVGT5-βG (41), SVIV-Raβ2, and SVIV-IVS2 (7) have been presented elsewhere. The recombinant viruses were all propagated with a *tsA*58 helper virus. Derivatives of SVIV-IVS2 that carried deletions of IVS2 were generated by digestion with the indicated restriction endonucleases, electrophoretic purification of the appropriate fragment, and transfection of CV1P cells with the linear DNA along with *tsA*58 helper virus DNA. Three deletions (*dl*17, *dl*18, *dl*13) were created by digestion of SVIV-IVS2 DNA with *Hpa*II endonuclease, followed by brief digestion with λ exonuclease (58). CV1P cells were transfected with the linearized DNA along with *tsA*58 helper DNA. Repair and ligation of the DNA ends occurred in vivo, generating a set of deletions emanating from the *Hpa*II site. The exact structure of each deletion was determined by DNA sequencing (32).

Extraction and analysis of labeled β-globin protein from infected cells. Virus-infected cells were labeled 48 h postinfection for 45 min with [^{3}H]leucine and [^{3}H]valine, proteins were extracted and analyzed as described previously (41). Goat anti-rabbit β-globin immunoglobulin (a gift from S. H. Boyer, Johns Hopkins University, Baltimore, Md.) was used to precipitate proteins by binding it to formaldehyde-fixed *Staphylococcus aureus* cells (IgGsorb; Enzyme Center, Boston, Mass.). The cells were washed extensively in a Triton X-100 buffer, and the bound proteins were eluted with sodium dodecyl sulfate-urea and electrophoresed in a 15 to 20% gradient sodium dodecyl sulfate-polyacrylamide gel. The gel was treated with En^{3}Hance (Dupont, NEN Research Products), dried, and fluorographed with XR5 film (Eastman Kodak Co., Rochester, N.Y.) and Cronex Lightning-Plus (E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.) intensifying screens at −70°C.

Extraction and analysis of RNAs. The methods for the extraction of nuclear and cytoplasmic RNAs, treatment with DNase I (Worthington Diagnostics, Freehold, N.J.), and chromatography on oligo(dT)-cellulose (Collaborative Research, Inc., Waltham, Mass.) have been described in detail elsewhere (55, 58). The spliced structures of RNAs were determined either by the method of Berk and Sharp (4) with virus DNA probes that were uniformly labeled with ^{32}P or by our modification (8) of the procedure of Weaver and Weissmann (57) with DNA probes that were end labeled with ^{32}P. Uniformly labeled DNAs were produced by incorporation of $^{32}P_i$ into viral DNA during growth in vivo (55). Restriction fragments were ^{32}P-labeled at the 5′ end by using calf alkaline phosphatase, T4 polynucleotide kinase, and [γ-^{32}P]ATP; or the fragments were labeled at the 3′ end by replacement synthesis by using T4 DNA polymerase and [α-^{32}P]dATP (30). RNAs were annealed to the DNA probes and digested with S1 nuclease, and the products were separated by electrophoresis either in 1.5% agarose gels or in polyacrylamide-urea gels. The gels were dried and analyzed by autoradiography with XR5 film (Kodak) for 1 to 7 days at −70°C with a Cronex Lightning-Plus intensifying screen. Signals on the autoradiograms were quantitated with a densitometer (Quick Scan Jr.; Helena Laboratories).

Plasmid constructions. Standard recombinant DNA methods were used in the plasmid constructions (30). Plasmid pSβ-IVS2 (8) was generated by ligation of a *Hin*dIII-*Bcl*I fragment (1.3 kilobase pairs [kbp]) of SVIV-IVS2 with a *Hin*dIII-*Bam*HI fragment (3.4 kbp) of the plasmid pSV2-*gpt* (40). pSβ-NI was constructed by replacing the *Bam*HI-*Eco*RI fragment of pSβ-IVS2 (640 bp) with a *Bam*HI-*Eco*RI fragment (67 bp) of SVIV-Raβ2 that contained β-globin cDNA sequences.

p*tk*2.0-NI was a derivative of pSβ-NI, in which the *Hin*dIII-*Bam*HI segment of β-globin DNA (339 bp) was replaced by a 2.0-kbp *Pvu*II restriction fragment of herpes simplex virus type 1 DNA that contained the thymidine kinase (*tk*) gene (33, 56). The ends of the 2.0-kbp *Pvu*II fragment were previously modified by the addition of synthetic *Hin*dIII and *Bam*HI decanucleotide linkers, such that a *Hin*dIII site was placed upstream of the *tk* promoter and a *Bam*HI site was placed downstream of the *tk* polyadenylation site (pSVOTKE; M. Fromm, Ph.D. thesis, Stanford University, Stanford, Calif., 1982). p*tk*2.0-NI DNA was used as the source of *tk* restriction fragments for the construction of the p*tk*0.7, p*tk*1.1, and p*tk*βins plasmids. The structures of other plasmids are explained throughout the text.

RESULTS

Intron-dependent mRNA production in SV40-rabbit β-globin recombinant viruses. Originally, for the purpose of studying SV40 early region transcription, two recombinant viruses were constructed in which rabbit β-globin-coding sequences were under the transcriptional control of the SV40 early region promoter (Fig. 1A). One recombinant, SVIV-Raβ2, contained β-globin cDNA sequences (at 486-bp *Hin*dIII-*Bgl*II fragment of rabbit β-globin cDNA) inserted between an SV40 early region promoter and a late region polyadenylation site. Since no SV40 or β-globin intron was included in this construct, it was expected that SVIV-Raβ2 would produce mature β-globin mRNA without the need for splicing. Another recombinant, SVIV-IVS2 (Fig. 1A), was constructed so that a segment of the β-globin cDNA in SVIV-Raβ2 (the 67-bp *Bam*HI-*Eco*RI segment) was replaced by the corresponding fragment of β-globin genomic DNA containing IVS2 (640 bp), which is the second intron of the gene. Hence, it was expected that SVIV-IVS2 would produce the same mRNA as SVIV-Raβ2, but in the case of SVIV-IVS2 splicing of the nascent transcript should occur. Both recombinant viruses were assayed for β-globin protein production and compared with another recombinant virus, SVGT5-βG, which contained β-globin cDNA sequences that were expressed from the SV40 late region promoter, splicing, and polyadenylation signals (41). β-Globin synthesis in infected monkey cells (CV1) was measured by specific immunoprecipitation of ^{3}H-labeled protein with goat anti-rabbit β-globin immunoglobulin, followed by gel electrophoresis and autoradiography (data not shown). SVGT5-βG produced β-globin protein in the amount expected for a viral late region gene product (41). SVIV-IVS2 produced severalfold more β-globin protein than SVGT5-βG. The observed overproduction probably resulted from the presence of two viral replication origins in the SVIV genome (7). The significant result, however, was that β-globin synthesis was not detected in cells that were infected with SVIV-Raβ2, the recombinant that lacked the intervening sequence. To examine the basis for this defect, β-globin RNA levels were measured in cells that were infected with SVIV-Raβ2 and SVIV-IVS2 by using an S1 nuclease protection assay (57). Polyadenylated nuclear and cytoplasmic RNAs were annealed to an SVIV-IVS2 DNA probe (a 3.6-kbp *Bam*HI-

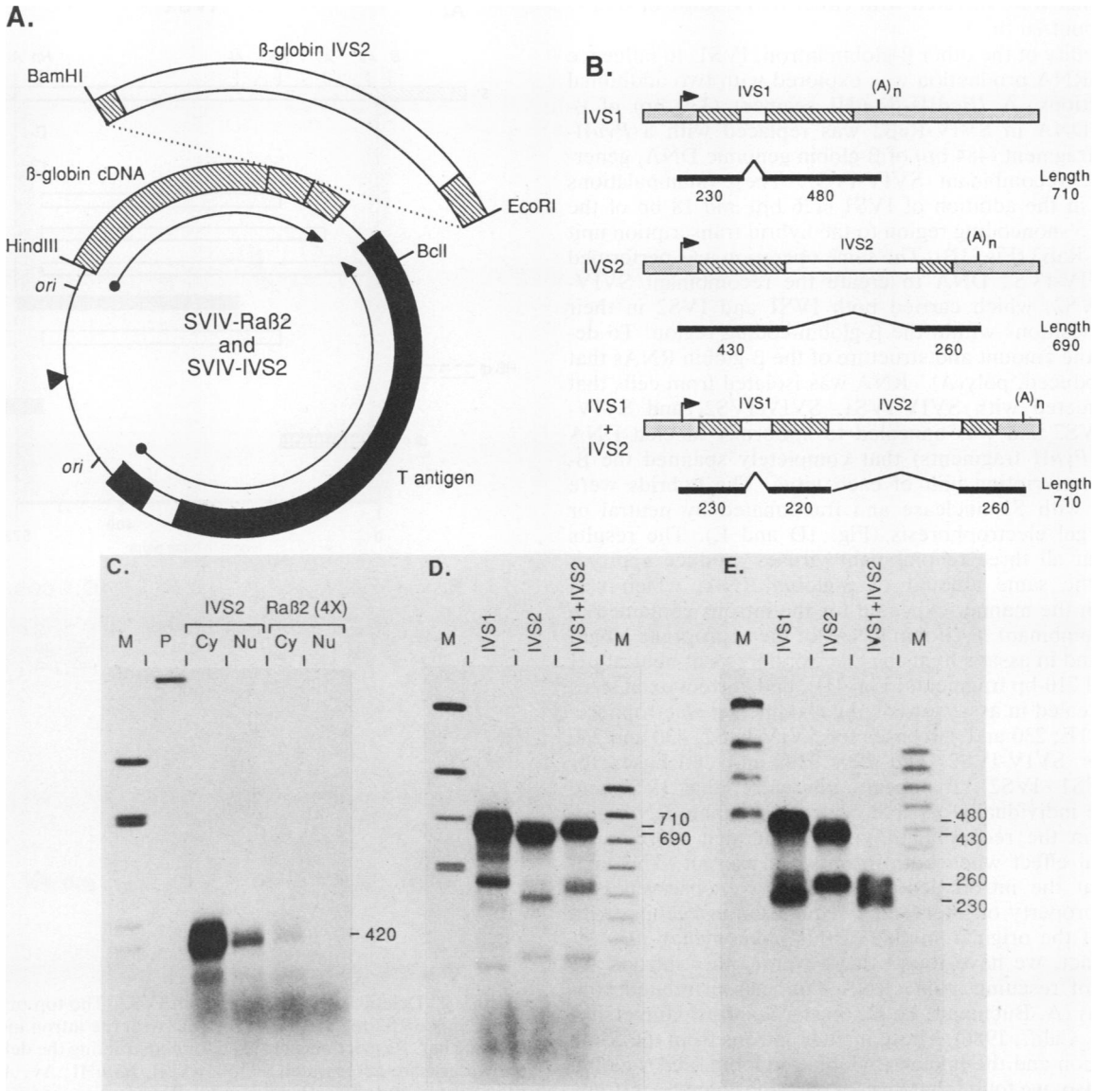


FIG. 1. Intron-dependent gene expression in SV40-rabbit β-globin recombinant viruses. (A) Structures SVIV-Raβ2 and SVIV-IVS2. The SVIV viruses have an inverted duplication of the SV40 origin of DNA replication (*ori*) and the early region promoter (7). Regions transcribed by the duplicate promoters are indicated with arrows. The rabbit β-globin coding region is diagonally hatched, and the viral T-antigen-coding sequences are solid. SVIV-IVS2 was constructed from SVIV-Raβ2 by replacing the *Bam*HI-*Eco*RI region of β-globin cDNA (12) with the *Bam*HI-*Eco*RI segment of β-globin genomic DNA (26). β-Globin IVS2 and T-antigen intron sequences are open. (B) Structures of the β-globin transcription units of SVIV-IVS1, SVIV-IVS2, and SVIV-IVS1+IVS2. SV40 sequences are stippled, β-globin coding regions are hatched, and intervening sequences are open. The structures of the expected RNAs are indicated below the maps of each transcription unit. (C) Poly(A)$^+$ cytoplasmic (Cy) and nuclear (Nu) RNAs were extracted from CV1 cells 48 h after infection with SVIV-Raβ2 and SVIV-IVS2 and annealed to a DNA probe that was 5′-^{32}P-labeled at the *Bam*HI site within the β-globin region of SVIV-IVS2 (lane P). The hybrids were digested with S1 nuclease and fractionated by electrophoresis in a 1.5% agarose gel (7). RNA from 2×10^6 infected cells was used in the analyses of SVIV-IVS2, and RNA from 8×10^6 infected cells was used in the analyses of SVIV-Raβ2. Markers (lane M) produced by digestion of labeled SV40 DNA with *Hin*dIII are shown in the leftmost lane (1,768, 1,169, 1,118, 526, 447, and 215 bp). (D and E) Cytoplasmic polyadenylated RNA was isolated from CV1 cells 48 h after infection with SVIV-IVS1, SVIV-IVS2, and SVIV-IVS1+IVS2; annealed to homologous uniformly labeled DNA probes; and digested with S1 nuclease. The products were fractionated by electrophoresis in 1.5% neutral (D) or alkaline (E) agarose gels. SV40 *Hin*fI markers are shown in the leftmost lane of each panel (1,847, 1,085, 766, 543, 525, 237, and 109 bp). SV40 *Bst*NI markers are shown in the rightmost lane of each panel (993, 823, 673, 552, 444, 369, 311, 249, 200, and 127 bp).

*Eco*RI fragment) that was 5′-^{32}P labeled at the *Bam*HI site. The hybrids were digested with S1 nuclease and fractionated by agarose gel electrophoresis (Fig. 1C). Analysis of SVIV-IVS2 RNA yielded a protected fragment of 420 bp with both the nuclear and cytoplasmic fractions; this is the expected size for RNAs initiated at the SV40 early region promoter (7, 8). We found that SVIV-IVS2 produced 400-fold more cytoplasmic β-globin RNA than SVIV-Raβ2 (note that the analysis of SVIV-Raβ2 was performed with 4 times the amount of total RNA as the analysis of SVIV-IVS2). Longer exposures of the autoradiogram in Fig. 1C showed that the ratio of cytoplasmic to nuclear β-globin RNA was the same

in cells that were infected with either SVIV-Raβ2 or SVIV-IVS2 (about 20 to 1).

The ability of the other β-globin intron, IVS1, to influence stable mRNA production was explored with two additional constructions. A *Hin*dIII-*Bam*HI segment (339 bp) of β-globin cDNA in SVIV-Raβ2 was replaced with a *Pvu*II-*Bam*HI fragment (484 bp) of β-globin genomic DNA, generating the recombinant SVIV-IVS1. These manipulations resulted in the addition of IVS1 (126 bp) and 18 bp of the β-globin 5′-noncoding region to the hybrid transcription unit of SVIV-Raβ2 (Fig. 1B). The same operation was performed with SVIV-IVS2 DNA to create the recombinant SVIV-IVS1+IVS2, which carried both IVS1 and IVS2 in their normal positions within the β-globin-coding region. To determine the amount and structure of the β-globin RNAs that were produced, poly(A)$^+$ RNA was isolated from cells that were infected with SVIV-IVS1, SVIV-IVS2, and SVIV-IVS1+IVS2 and was annealed to uniformly labeled DNA probes (*Pvu*II fragments) that completely spanned the β-globin transcription unit of each virus. The hybrids were digested with S1 nuclease and fractionated by neutral or alkaline gel electrophoresis (Fig. 1D and E). The results show that all three recombinant viruses produce approximately the same amount of β-globin RNA, which was spliced in the manner expected for the introns contained in each recombinant. β-Globin RNAs of the appropriate length were found in assays by using electrophoresis at neutral pH (690- and 710-bp fragments; Fig. 1D), and correct exon sizes were revealed in assays involving alkaline gel electrophoresis (Fig. 1E; 230 and 480 bases for SVIV-IVS1; 430 and 260 bases for SVIV-IVS2; and 220, 230, and 260 bases for SVIV-IVS1+IVS2). It appears, therefore, that IVS1 and IVS2 are individually capable of rescuing stable RNA production in the recombinant viruses and that there is no additional effect when both introns are present. This suggests that the intron dependence that we observed is a general property of intervening sequences, in keeping with results of the original studies of this phenomenon (23, 24, 41). In fact, we have found that several other introns are capable of rescuing stable RNA synthesis in related constructions (A. Buchman, Ph.D. thesis, Stanford University, Stanford, Calif., 1986). These include introns from the SV40 early region and the human β-globin, δ-globin, and γ-globin genes. Also, we found that viruses carrying chimeric introns, which include segments from the SV40 large T-antigen intron and β-globin IVS2 or β-globin IVS1 and IVS2, produce normal amounts of appropriately spliced mRNA.

The intron sequences required for stable RNA production are at the splice sites. Although intervening sequences, in general, were capable of stimulating mRNA production from the recombinant viruses, it was unclear whether the sequences responsible for this effect coincided with those required for mRNA splicing. To answer this question, a series of deletions was constructed in which different sequences within and flanking IVS2 were removed from SVIV-IVS2 (Fig. 2A). Because of the complexity of the manipulations involved, two deletions (*dl*1 and *dl*2) were constructed in a homologous plasmid DNA vector described below. RNAs produced by the different deletions were detected with S1 nuclease protection assays, as described in the legend to Fig. 1D and E (summarized in Fig. 2A legend; data not shown). The deletions that removed only sequences within IVS2 (*dl*17, *dl*18, *dl*6, *dl*10, and *dl*15) all produced substantial amounts of appropriately spliced β-globin RNA. The largest of the internal deletions, *dl*15, produced several-fold less β-globin mRNA than the parent virus, SVIV-IVS2,

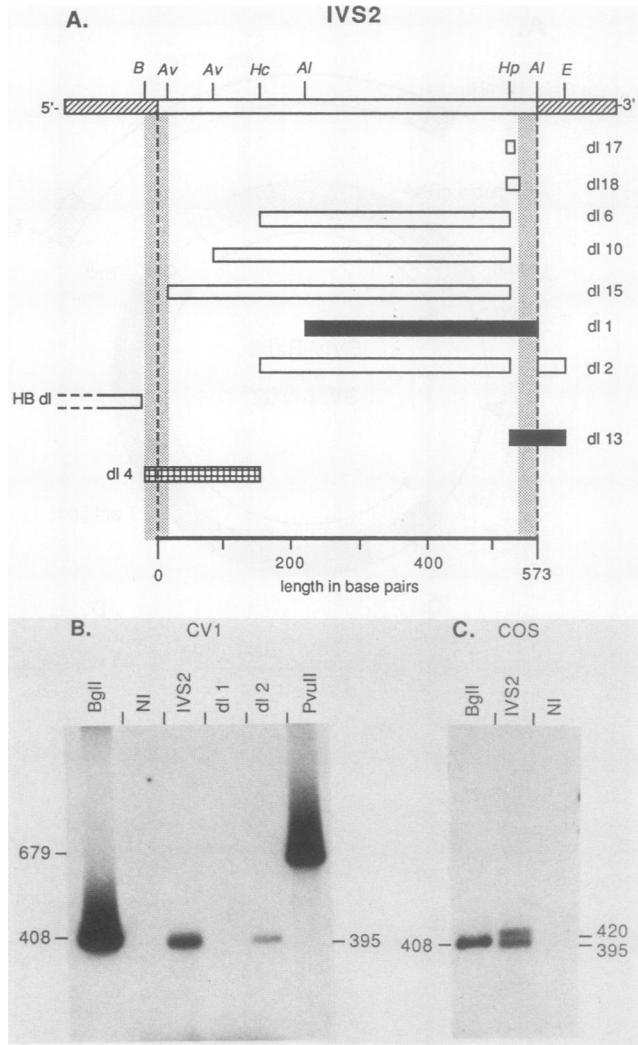


FIG. 2. Deletions of rabbit β-globin IVS2. The top of the figure depicts the IVS2 region of SVIV-IVS2, with the intron indicated as an open bar. Restriction sites used for constructing the deletions are shown and are abbreviated as follows: B, *Bam*HI; Av, *Ava*II; Hc, *Hin*cII; Al, *Alu*I; Hp, *Hpa*II; E, *Eco*RI. The sequences that were removed in each deletion are shown by the bars. Solid bars indicate deletions that disrupted both splicing and stable RNA production. Open bars indicate deletions that produced normal amounts of appropriately spliced RNA. *dl*4, depicted with cross-hatching, produced slightly reduced levels of abnormally spliced RNA. (B and C) Poly(A)$^+$ cytoplasmic RNA was isolated from CV1 (B) and COS (C) cells 48 h after transfection with the plasmids pSβ-NI, pSβ-IVS2, *dl*1, and *dl*2 and were annealed to a pSβ-IVS2 DNA probe that was 5′-^{32}P-labeled at the *Bam*HI site in the β-globin region. The hybrids were digested with S1 nuclease and electrophoresed in a 1.5% agarose gel. The size of marker fragments produced by digestion of the probe with restriction endonucleases, *Bgl*I (408 bp) and *Pvu*II (670 bp), are indicated on the left side of each panel.

did. In this case the intron was reduced from 573 to 66 bp, deleting all but 14 bp next to the 5′ splice site. Other investigators (60) have shown that IVS2 is spliced inefficiently when it is reduced to less than 80 bp.

Deletion of either the IVS2 5′ splice site (*dl*4) or 3′ splice site (*dl*13) had a more noticeable effect on RNA production. *dl*4 produced fourfold less β-globin RNA than SVIV-IVS2 did, and several different RNAs were found. Further analysis of *dl*4 RNA (Buchman, Ph.D. thesis) showed that these

species resulted from the utilization of cryptic 5′ splice sites within both IVS2- and the β-globin-coding sequences. Other studies of mutant rabbit and human β-globin genes with alterations at the IVS2 5′ splice junction revealed the existence of cryptic splice sites (53, 60). A deletion of the 3′ splice site of IVS2 had the most extreme effect on mRNA production. *dl*13 produced only 1/20th the amount of β-globin RNA as SVIV-IVS2 did, and the RNA that was detected appeared to be unspliced.

The importance of the nucleotides flanking the 3′ splice site was examined in more detail with the deletions *dl*1 and *dl*2 (Fig. 2A). In *dl*1, a segment of IVS2 was removed that extended from the *Alu*I restriction site within the intron to the *Alu*I cleavage site, 5′-AG/CT-3′, that overlaps the conserved 5′-AG-3′ dinucleotide that is generally found at 3′ splice sites (38). *dl*2 was derived from *dl*6 by removing sequences between the *Alu*I site at the 3′ splice junction and an *Eco*RI site 50 bp downstream, resulting in the deletion of exon sequences flanking the splice junction. Both of these deletions were constructed in the plasmid vector pSβ-IVS2, which contains the same hybrid SV40-rabbit β-globin transcription unit that is contained in SVIV-IVS2. CV1 cells transfected with *dl*1 or *dl*2 plasmid DNA (8, 45) were examined for β-globin RNA production by S1 nuclease assays by using a pSβ-IVS2 DNA probe that was 5′-^{32}P-labeled at the *Bam*HI site in the β-globin region (Fig. 1A). The RNAs were annealed to the DNA probe, digested with S1 nuclease, and fractionated by gel electrophoresis. *dl*2 produced slightly less β-globin RNA than the control plasmid pSβ-IVS2 did (Fig. 2B). Correct splicing of the *dl*2 RNA was revealed in other assays by using a different fragment as probe (data not shown). By contrast, no β-globin RNA was detected in the cells that were transfected with the *dl*1 plasmid (Fig. 2B), demonstrating that intron sequences adjacent to the 3′ splice site are required for β-globin RNA production.

Intron-dependent mRNA synthesis is observed in short-term transfections with both replicating and nonreplicating plasmid DNAs. For greater flexibility in studying the effects of introns on RNA production, we switched from using SV40 recombinant viruses to using a plasmid vector system. Since the original investigations of intron-dependent gene expression were all performed with recombinant viruses (22–25), it seemed relevant to determine whether any basic property of the vector or host changed the effect of introns on expression. Accordingly, two *Escherichia coli*-animal cell shuttle vectors were constructed. Plasmid pSβ-NI (NI indicates no intron) contained the β-globin transcription unit of SVIV-Raβ2, and pSβ-IVS2 contained the β-globin transcription unit of SVIV-IVS2 (Fig. 3A and B). Two different monkey cell lines were used as hosts for these plasmids. CV1 is the normal host for SV40 virus infections, and COS (16) is a specially constructed derivative of CV1 that contains an integrated copy of the SV40 large T-antigen gene. Since the plasmids pSβ-NI and pSβ-IVS2 contain an SV40 replication origin (*ori*), they are able to replicate in COS cells, but not in CV1 cells, because these cells lack a functional T-antigen gene. CV1 and COS cells were transfected with pSβ-NI and pSβ-IVS2 DNA by the calcium phosphate coprecipitation technique (8, 45). RNA was extracted after 48 h and analyzed with S1 nuclease assays by using DNA probes that were 5′-^{32}P-labeled at the *Bam*HI site of each plasmid (Fig. 2B and C). No β-globin RNA was detected in either COS or CV1 cells that were transfected with pSβ-NI DNA. By contrast, significant amounts of β-globin RNA were produced in cells that were transfected with pSβ-IVS2. A 395-bp fragment was produced with RNA from CV1 cells, and two protected fragments, 395 and 420 bp, were seen with RNA from COS cells. The observed fragment sizes were those expected for initiations at the different components of the SV40 early region promoter (7, 8). The 395-bp fragment was produced by RNAs that were initiated at the downstream early region promoter, which depends on both the TATA and the 72-bp enhancer sequences (8). The 420-bp fragment was produced by RNAs that were initiated at the upstream early region promoter. Upstream initiations are not dependent on the TATA or the enhancer sequences, but rather, they rely on amplification of the template for their production (8). Thus, both enhancer-dependent and amplification-dependent transcripts were greatly affected by the presence of IVS2. The same difference in RNA production between pSβ-NI and pSβ-IVS2 was found when a human cell line (HeLa) was used as the host for transfections (data not shown); therefore, the observed intron requirement is not a peculiarity of CV1 derivatives.

IVS2 fails to rescue stable RNA production when placed outside the transcription unit. From the results with SVIV-IVS1 and SVIV-IVS2, it seemed that the ability of introns to stimulate mRNA production in this system is not critically dependent on their position. In some ways, this behavior is reminiscent of that of transcriptional enhancer elements, which also exhibit a position-independent effect on gene expression. However, since splice site sequences were important for the effect of IVS2, this indicates that the mechanism of IVS2 stimulation is connected to the splicing process and is different than that of transcriptional enhancers. Transcriptional enhancers exert their effects even when they are located downstream and outside of the transcribed gene (2, 3, 14, 37). Accordingly, we examined whether IVS2 would produce its effect when it was located just downstream of the polyadenylation site. This was done by using a set of four plasmids that carried tandem duplications of a *Bam*HI-*Sau*3AI fragment of pSβ-NI or pSβ-IVS2, which spans the IVS2 splice junctions and the SV40 late region polyadenylation site (Fig. 3C through F). Each construct contained two polyadenylation sites and could, in principle, produce two RNAs by processing at either site. The plasmids pSβ-IND and pSβ-IID contained IVS2 inserted upstream of both polyadenylation sites; pSβ-IID had a second copy of IVS2 between the two polyadenylation signals. pSβ-NID had IVS2 inserted just downstream (95 bp) of the first polyadenylation site. No intervening sequence was included in the plasmid pSβ-NND. Because of the nature of these constructions, each plasmid could produce the same two β-globin mRNAs by different combinations of processing events.

COS cells were transfected with pSβ-IVS2 and each of the duplicated plasmid derivatives. RNA was extracted after 48 h and analyzed by S1 nuclease assays by using a pSβ-NND DNA probe that was 3′-^{32}P-labeled at the *Hin*dIII site. The products of nuclease digestion were fractionated by electrophoresis in a neutral agarose gel (data not shown). pSβ-IND and pSβ-IID produced β-globin RNA in an amount comparable to that produced by pSβ-IVS2. The predominant species was polyadenylated at the upstream processing site, yielding a protected fragment of 630 bp. A small amount of RNA (1 to 2%) was found that was processed at the downstream polyadenylation site, producing a 910-bp protected fragment. No β-globin RNA of either size was detected in cells that were transfected with pSβ-NND or pSβ-NID. Hence, IVS2 failed to exert an effect when it was placed downstream of the polyadenylation signal. The lack

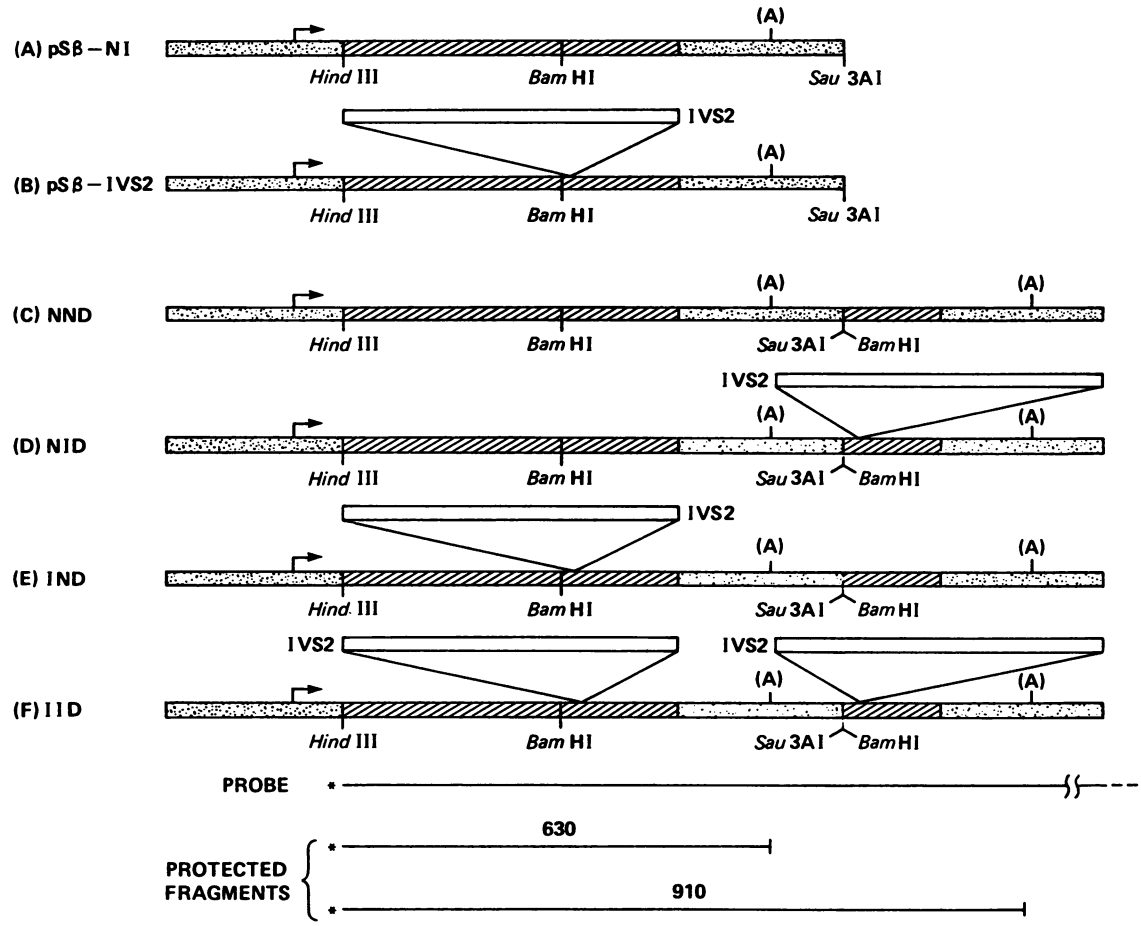
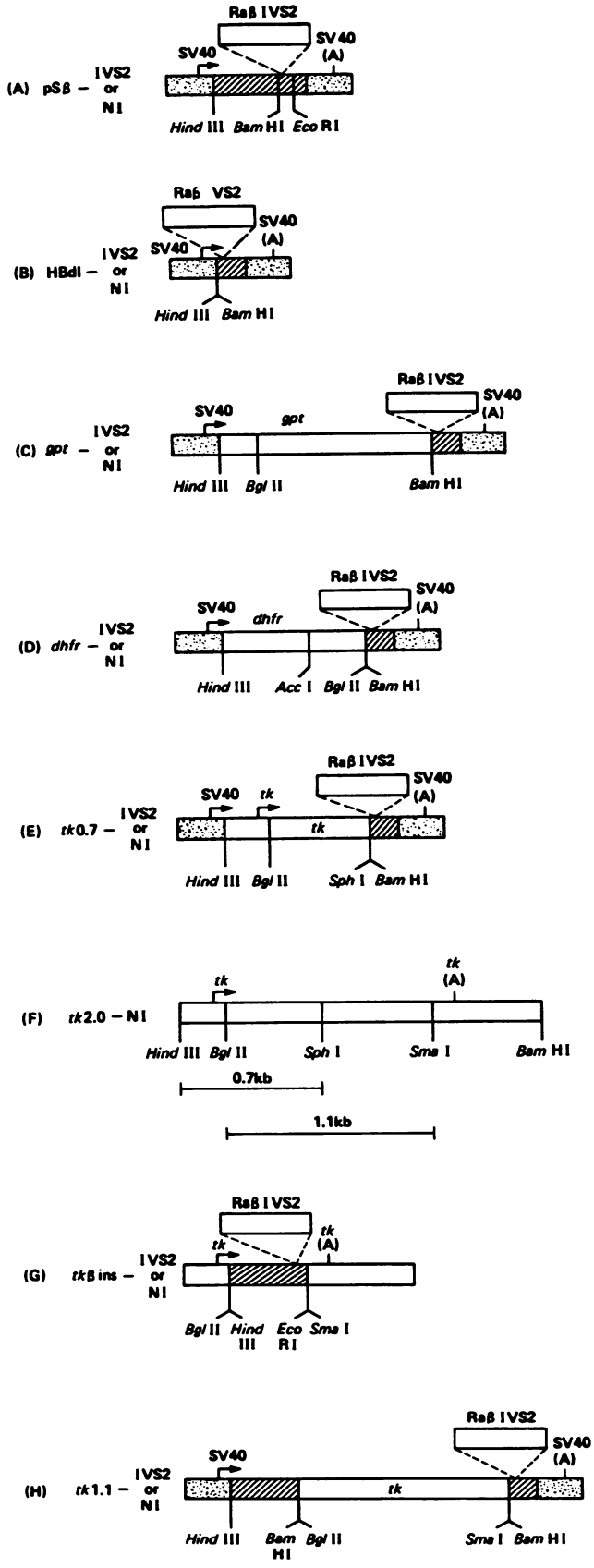


FIG. 3. Testing IVS2 function downstream of a polyadenylation site. (A through F) The β-globin transcription units of pSβ-NI, pSβ-IVS2, and the derivatives with duplicate polyadenylation sites are depicted. SV40 sequences are stippled, β-globin-coding sequences are hatched, and IVS2 sequences are shown above each map as open bars. The rightward arrow shows the location of the SV40 early region promoter, and the A in parentheses indicates the late region polyadenylation sites. The two different protected fragments expected in S1 nuclease assays from RNA processed at the upstream (630 bp) and downstream (910 bp) polyadenylation sites are shown at the bottom.

of stimulation by IVS2 from the downstream position does not appear to be the result of its increased distance from the promoter. A significant effect of IVS2 was found in other constructions, discussed below, where IVS2 was placed at a similar distance (p*dhfr* plasmids; Fig. 4D) or at even greater distance (p*gpt* plasmids; Fig. 4C) from the promoter.

Different transcribed sequences vary greatly in their dependence on introns for stable mRNA production. Although the β-globin constructs that we have described were highly dependent on introns for mRNA production, this requirement was clearly not a general one, since numerous eucaryotic genes lack intervening sequences. We explored the difference between intron-dependent and intron-independent expression by constructing derivatives of the pSβ vectors that carried different sequences inserted in the transcribed region. For each sequence tested, a pair of plasmids was constructed that differed in only one aspect; one contained IVS2 (-IVS2 derivative) and one lacked IVS2 (-NI derivative), so that the influence of the intron could be assessed properly. This was accomplished by performing identical constructions with pSβ-NI and pSβ-IVS2 DNA. COS cells were transfected with each of the plasmids, and RNA production was measured with S1 nuclease assays by using end-labeled DNA probes. For each analysis, at least two different probe DNAs were used to check for the expected mRNA structure. No unexpected transcripts were detected, so only the measurements with one of the probes are shown.

The first two plasmids, pHB*dl*-NI and pHB*dl*-IVS2 (Fig. 4B), were created as a control for the following constructs and were used to investigate the influence of the cryptic 5′ splice sites present in the β-globin coding sequence. It seemed possible that the cryptic 5′ splice sites might promote aberrant or incomplete processing in the absence of a 3′ splice site, resulting in degradation of the transcript. The *Hin*dIII-*Bam*HI region (339 bp) of β-globin cDNA that contained the cryptic splice sites was removed from pSβ-NI and pSβ-IVS2, generating the pHB*dl* derivatives (Fig. 4B). pHB*dl*-IVS2, however, still produced 500-fold more β-globin RNA than pHB*dl*-NI (133-bp protected fragment, Fig. 5A), demonstrating that removal of the cryptic splice sites did not affect intron dependence. In a second pair of constructions (Fig. 4C), the *Hin*dIII-*Bam*HI region of β-globin cDNA was replaced with a *Hin*dIII-*Bam*HI fragment (1.0 kbp) of the *E. coli gpt* gene (39). A substantial effect of the intron was still observed with this procaryotic insert, as p*gpt*-IVS2 produced 10- to 20-fold (the variation between different transfection experiments) more RNA than did p*gpt*-NI (180- and 205-bp protected fragments, Fig. 5B). A different eucaryotic cDNA was examined with the p*dhfr* constructions (Fig. 4D), in which the β-globin *Hin*dIII-*Bam*HI region was replaced



by a *Hin*dIII-*Bgl*II fragment (0.7 kbp) of mouse dihydrofolate reductase (*dhfr*) cDNA containing the entire coding region (9, 51). Like the rabbit β-globin gene, the mouse *dhfr* gene contains intervening sequences (43). In this case, as well, p*dhfr*-IVS2 produced 10- to 20-fold more RNA than did p*dhfr*-NI (465- and 485-bp protected fragments, Fig. 5C). Thus, the stimulation of mRNA accumulation by the presence of an intron was observed with a bacterial coding sequence (*gpt*) and another eucaryotic cDNA (*dhfr*), although the magnitude of the effect was not as great as that seen with the original β-globin plasmids.

The fourth set of plasmids tested the properties of sequences from a eucaryotic gene that lacks introns, the herpes simplex virus *tk* gene (33, 56). In the p*tk*0.7 plasmids (Fig. 4E), the *Hin*dIII-*Bam*HI segment of β-globin cDNA was replaced with a *Hin*dIII-*Sph*I (700 bp) fragment of *tk* DNA. This fragment contains the promoter of the *tk* gene and 500 bp of the transcribed region. Most of the RNA produced in transfections with the p*tk*0.7 plasmids was initiated at the herpes simplex virus *tk* promoter (500-bp protected fragment, Fig. 5D). By contrast with the previous plasmids, substantial amounts of RNA were produced with both p*tk*0.7 derivatives, with p*tk*0.7-IVS2 producing about twice as much *tk* RNA as p*tk*0.7-NI. A somewhat larger effect of the intron was observed for those transcripts that were initiated at the SV40 early region promoter (765-bp protected fragment, Fig. 5D), for which a three- to fourfold difference was found.

Two additional pairs of plasmids were constructed to assess whether the elements directing intron-independent expression of the *tk* gene were within the transcribed region or in the promoter and polyadenylation regions. In the p*tk*βins plasmids (Fig. 4G), most of the transcribed region of the *tk* gene (1.1-kbp *Bgl*II-*Sma*I segment) was replaced with the β-globin *Hin*dIII-*Eco*RI region of either pSβ-NI or pSβ-IVS2. Conversely, the p*tk*1.1 plasmids (Fig. 4F) were generated by inserting the 1.1-kbp *Bgl*II-*Sma*I region of *tk* at the *Bam*HI site of pSβ-NI or pSβ-IVS2. The presence of IVS2 caused a 20-fold increase in RNA production from the p*tk*βins plasmids (340-bp protected fragment, Fig. 5E) but only a 2-fold effect with the p*tk*1.1 constructs (395- and 420-bp protected fragments, Fig. 5F). Thus, it appears that the elements that permitted intron-independent RNA production are located primarily within the transcribed region of the *tk* gene.

DISCUSSION

Certain mammalian transcription units fail to produce stable RNA if they lack an intron. The magnitude of this

FIG. 4. Structures of the hybrid transcription units of pSβ plasmids (A) and their derivatives (B through H) are shown with different regions depicted as follows: SV40 sequences, stippled bars; β-globin-coding sequences, hatched bars; different inserted sequences (*gpt*, *dhfr*, *tk*), open bars. β-Globin IVS2 sequences present in the IVS2 derivatives are indicated above each map as open bars, with the dashed lines showing its position. Rightward-pointing arrows show the location of SV40 or *tk* promoters, and the A's in parentheses show the positions of SV40 or *tk* polyadenylation sites. Restriction sites used for either construction of the plasmids or end-labeling DNA probes for S1 nuclease assays are shown below the maps. The structure of the herpes *tk* gene is depicted in panel F as the 2.0-kbp fragment inserted in the plasmid p*tk*2.0-NI. The 0.7-kbp *Hin*dIII-*Sph*I fragment and the 1.1-kbp *Bgl*II-*Sma*I fragment of *tk* that was used to construct the p*tk*0.7 and p*tk*1.1 plasmids, respectively, are indicated below the map of p*tk*2.0-NI.

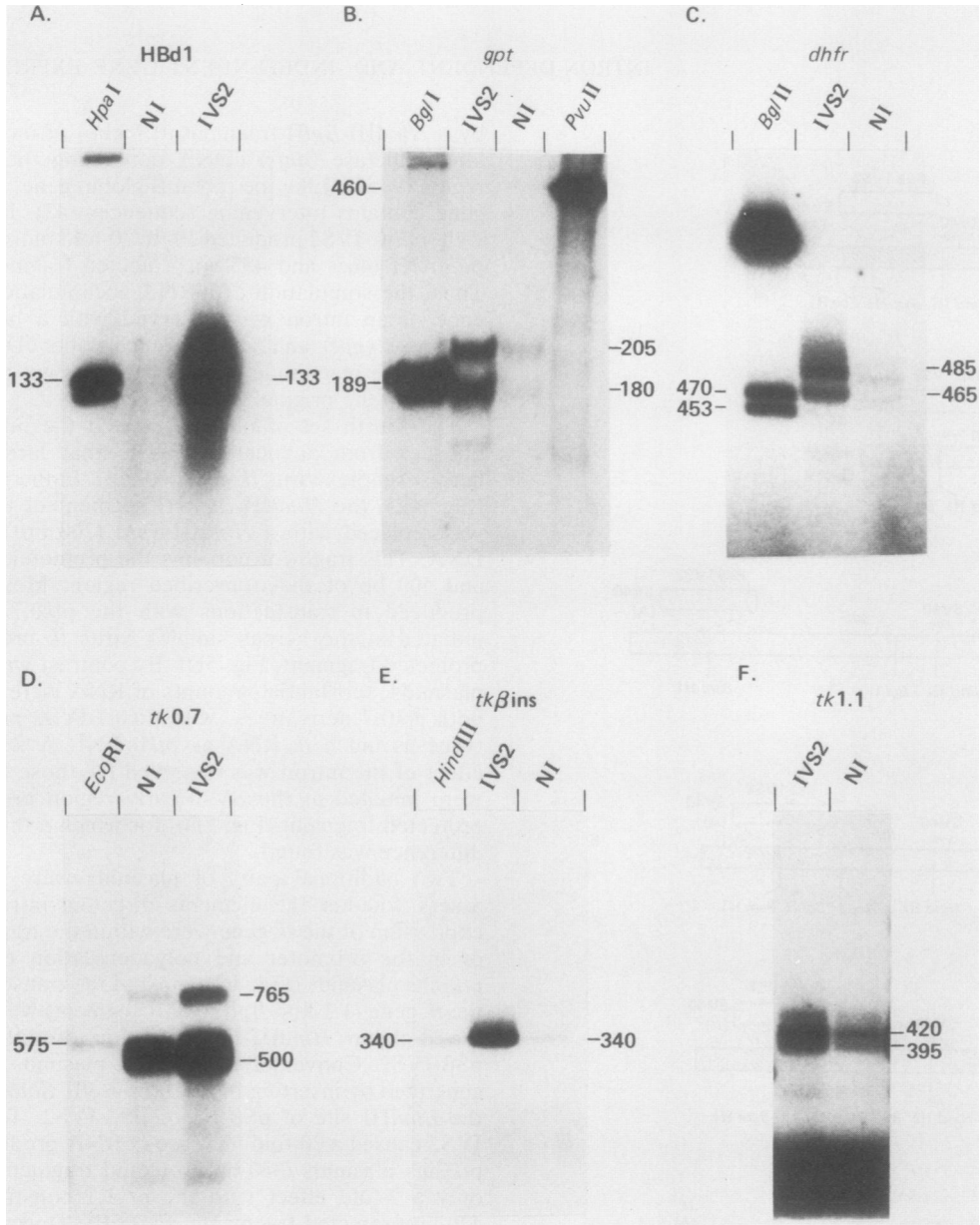


FIG. 5. A survey of different transcribed sequences for intron-dependent expression. Poly(A)$^+$ RNA was isolated from COS cells 48 h after transfection with the pHB*dl*, p*gpt*, p*dhfr*, p*tk*0.7, p*tk*βins, and p*tk*1.1 plasmids. RNAs were analyzed by S1 nuclease assays by using end-labeled DNA probes, described below, and the products were fractionated by gel electrophoresis. Marker fragments produced by digestion of the probes with the indicated enzyme are shown in the leftmost lane of each panel. Numbers on the right of each panel show the size of the fragments protected from nuclease digestion. RNA from 5×10^5 transfected COS cells was used in all the analyses, but different autoradiogram exposures are shown for the different classes of constructions to best illustrate the effect of IVS2. (A) RNAs produced by pHB*dl*-NI and pHB*dl*-IVS2 were annealed to a DNA probe that was derived from pSβ-NI DNA, that was 3′-^{32}P-labeled at the *Bam*HI site (Fig. 4A), and that extended beyond the SV40 late region polyadenylation site. The products of the S1 nuclease assays were separated by electrophoresis in a 4% polyacrylamide-urea gel. (B) RNAs from p*gpt*-NI and p*gpt*-IVS2 were annealed to a probe derived from p*gpt*-NI DNA that was 5′-^{32}P-labeled at the *Bgl*II site (Fig. 4C) and which extended beyond the SV40 early region promoter. The products of the S1 nuclease assays were separated by electrophoresis in a 4% polyacrylamide-urea gel. The 180- and 205-base protected fragments were generated by RNAs that were initiated at the downstream and upstream components, respectively, of the SV40 early region promoter (7). (C) RNAs from p*dhfr*-NI and p*dhfr*-IVS2 were annealed to a probe that was derived from *pdhfr*-NI DNA that was 5′-^{32}P-labeled at the *Acc*I site (Fig. 4D) and which extended beyond the SV40 early region promoter. The products of the S1 nuclease assays were separated by electrophoresis in a 1.5% alkaline agarose gel. The 485- and 465-base protected fragments were produced by RNAs that were initiated at the upstream and downstream components, respectively, of the SV40 early region promoter. (D) RNAs from p*tk*0.7-NI and p*tk*0.7-IVS2 were annealed to a probe that was derived from p*tk*0.7-NI DNA, that was 5′-^{32}P-labeled at the *Bam*HI site (Fig. 4E), and that extended beyond the SV40 early region promoter. The products of the S1 nuclease assays were separated by electrophoresis in a 1.5% alkaline agarose gel. The 765- and 500-base protected fragments were generated by RNAs that were initiated at the SV40 early region and herpes *tk* promoters, respectively. (E) RNAs from p*tk*βins-NI and p*tk*βins-IVS2 were annealed to a probe that was derived from pSβ-NI DNA and that was 5′-^{32}P-labeled at the *Bam*HI site (Fig. 4A). The products of the S1 nuclease assays were separated by electrophoresis in a 1.5% agarose gel. Because of the structure of the p*tk*βins plasmids, the homology of the probe with p*tk*βins RNA extended only up to the *Hin*dIII site, producing a protected fragment of 340 bp. (F) RNAs from p*tk*1.1-NI and p*tk*1.1-IVS2 were annealed to a probe that derived from pSβ-NI DNA, that was 5′-^{32}P-labeled at the *Bam*HI site (Fig. 4A), and that extended beyond the SV40 early region promoter. The products of the S1 nuclease assays were separated by electrophoresis in a 1.5% agarose gel. The 395- and 420-bp protected fragments were generated by RNAs that were initiated at the downstream and upstream components, respectively, of the SV40 early region promoter.

effect was revealed while studying the expression of a β-globin cDNA segment fused at its 5′ and 3′ ends to SV40 promoter and polyadenylation signals, respectively (Fig. 1A). Only barely detectable levels of RNA were produced from such an intronless gene, but the same gene containing either the small (IVS1) or large (IVS2) β-globin intron in the appropriate location produced very substantial (400-fold greater) quantities of β-globin RNA. Other intervening sequences, the SV40 early region introns, or hybrid introns containing 5′ and 3′ splice sites from different intervening sequences were also effective in promoting β-globin mRNA formation (data not shown). Our experiments on the effect of modifications of the structure and location of the β-globin IVS2 on RNA production led to the following generalization. The integrity of the junction sequences that are needed for splicing the intron were considerably more important than either the position of the intron in the transcription unit or the nucleotide sequence between the 5′ and 3′ splice junctions. Where splicing was completely blocked, as when the 3′ splice junction was eliminated, RNA production was greatly reduced. However, when only the 5′ splice junction was missing, alternate splicing between cryptic 5′ splice sites and the 3′ splice junction permitted substantial levels of β-globin RNA production.

We are left with the question of what causes intron-dependent gene expression. The size of the effect of IVS2 in the SVIV recombinants rivals that found for eucaryotic promoter sequences. It was striking that the IVS2 dependence applied to both nuclear and cytoplasmic RNAs (Fig. 1C), demonstrating that the lack of splicing does not just block RNA transport. One explanation for this result might be that sequences within IVS2 control transcription initiation. Indeed, introns can contain sequences that apparently regulate transcription, such as the enhancer elements of immunoglobulin genes (2, 15, 34, 46). Although we have not tested this point directly, several observations appear to argue against this as the basis for our results. First, it seems unlikely that transcriptional control elements would reside in all the introns that we tested and would happen to coincide with the sequences that are required for splicing of IVS2. Also, unlike transcriptional enhancer sequences, IVS2 failed to show an effect when it was placed downstream of the polyadenylation site. Several promoters, the herpes simplex virus *tk* and the SV40 early region downstream and upstream promoters, all exhibited intron-dependent expression, even though these promoters show different dependences on transcriptional enhancer sequences (8). Finally, the behavior we observed was quite similar to that previously found by Hamer et al. (25) in studies of other recombinant SV40 viruses. From an analysis of pulse-labeled RNA, Hamer and colleagues (25) concluded that transcription is not affected by the presence of introns in these constructions. On the other hand, Brinster et al. (5) have recently reported that introns increase the transcription rates of genes that are introduced into transgenic mice. However, no effect of introns was observed by these investigators when the same constructions were examined in animal cells grown in culture, in direct contrast to our results. Thus, the mechanistic basis for the results that we observed and those of Brinster et al. (5) may be different.

It is curious, although not entirely unexpected, that sequences from the transcribed region of the herpes simplex virus *tk* gene could substitute for the intron requirement in our constructions. This raises the issue of whether the intron and *tk* sequences accomplish their effects by similar or different mechanisms. The sequences of *tk* that rescue stable RNA production have not been precisely defined. In preliminary experiments, we found that both the *Bgl*II-*Sph*I fragment (443 bp) and the *Sph*I-*Sma*I fragment (720 bp) of *tk* (Fig. 4) promote significant intron-independent expression. So, there does not appear to be a unique sequence within the *tk* gene that confers this property.

A plausible explanation for our results is that the intron and the transcribed sequences of *tk* exert their effect at a stage after transcription initiation. It is conceivable that this might involve steps that control transcription elongation, cleavage and polyadenylation of the transcript, or turnover of the nascent RNA. It is possible that cleavage and polyadenylation do not occur, in some cases, unless the transcript is also spliced. Although some similarities have been noted between the mechanisms of polyadenylation and RNA splicing, no direct connection has been observed in studies of these reactions in vitro (27, 31, 36, 49). If the block in stable RNA production is caused by premature termination of transcription, one might imagine that a splicing complex, or spliceosome (6, 13, 19), serves a function similar to that of a stalled ribosome in procaryotic transcription attenuators (61), allowing transcription to continue. On the other hand, if the problem in RNA synthesis is caused by degradation of the primary transcript by nucleases, then the spliceosome (or other ribonucleoproteins recruited by the spliceosome) might encapsulate the RNA and protect it from destruction. Related models can be envisioned for sequences within the transcribed region of *tk*, but without invoking the splicing process. Thus, one of the constituents of the splicing system might recognize *tk* RNA sequences without assembling an active spliceosome, or a different factor that is not involved in splicing might specifically bind *tk* RNA and have the same effect.

From a practical point of view, these results have several important implications. It is clear that significant levels of expression can be achieved with vectors that lack intervening sequences, such as p*dhfr*-NI and p*gpt*-NI (Fig. 5). However, the amount of expression is substantially increased (10- to 20-fold) when an intron is included in the construction. Furthermore, some sequences, such as the β-globin cDNA, do in fact show a virtual requirement (400-fold) for the presence of an intervening sequence. In general, it appears that inclusion of an intron is a wise strategy for maximizing gene expression in animal cells. In that connection, we have not observed an instance in which the presence of β-globin IVS2 causes a decrease in RNA production. It remains possible, though, that a particular intron will not be equally effective in all cell types, for example, if it is recognized or spliced inefficiently. Nor is it clear that a given intron will be able to fully exert its effects regardless of the nature of the foreign sequences being expressed or the relative position of the intron within the transcribed region (20, 47; Buchman, Ph.D. thesis). An alternative approach might be to include in the construction sequences from a gene that lacks introns, such as herpes simplex virus *tk*, or to include both an intron and *tk*-type sequences. At present there is no simple answer to this problem, but generally effective strategies may emerge as more is learned about the basis for the phenomena we have described.

ACKNOWLEDGMENTS

We thank R. Brinster and R. Palmiter for communicating results prior to publication.

This study was supported by Public Health Service grant GM-13235 from the National Institutes of Health to P.B.

LITERATURE CITED

1. **Alestrom, P., G. Akusjarvi, M. Perricaudet, M. B. Mathews, D. F. Klessig, and U. Pettersson.** 1980. The gene for polypeptide IX of adenovirus 2 and its unspliced messenger RNA. Cell **19:** 671–681.
2. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33:**729–740.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 sequences. Cell **27:** 299–308.
4. **Berk, A. J., and P. Sharp.** 1978. Spliced early RNAs of simian virus 40. Proc. Natl. Acad. Sci. USA **75:**1274–1278.
5. **Brinster, R. L., J. M. Allen, R. R. Behringer, R. E. Gelinas, and R. D. Palmiter.** 1988. Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. USA **85:**836–840.
6. **Brody, E., and J. Abelson.** 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science **228:**963–967.
7. **Buchman, A. R., and P. Berg.** 1984. Unusual regulation of simian virus 40 early-region transcription in genomes containing two origins of DNA replication. Mol. Cell. Biol. **4:**1915–1928.
8. **Buchman, A. R., M. Fromm, and P. Berg.** 1984. Complex regulation of simian virus 40 early-region transcription from different overlapping promoters. Mol. Cell. Biol. **4:**1900–1914.
9. **Chang, A. C. Y., J. H. Nunberg, R. J. Kaufman, H. A. Erlich, R. T. Schimke, and S. N. Cohen.** 1978. Phenotypic expression in *E. coli* of a cDNA sequence coding for mouse dihydrofolate reductase. Nature (London) **275:**617–624.
10. **Crowley, C. W., C. Liu, and A. D. Levinson.** 1983. Plasmid-directed synthesis of hepatitis B surface antigen in monkey cells. Mol. Cell. Biol. **3:**44–55.
11. **Early, P., J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, and L. Hood.** 1980. Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. Cell **20:**313–319.
12. **Efstratiadis, A., K. C. Kafatos, and T. Maniatis.** 1977. The primary structure of rabbit β-globin mRNA as determined from cloned cDNA. Cell **10:**571–585.
13. **Frendewey, D., and W. Keller.** 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-sn RNPs and specific intron sequences. Cell **42:**355–367.
14. **Fromm, M., and P. Berg.** 1983. SV40 early- and late-region promoter functions are enhanced by the 72-base-pair repeat inserted at distant locations and inverted orientations. Mol. Cell. Biol. **3:**991–999.
15. **Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33:**717–728.
16. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23:**175–182.
17. **Goff, S. P., and P. Berg.** 1976. Construction of hybrid SV40 and lambda phage DNA segments and their propagation in cultured monkey cells. Cell **9:**695–705.
18. **Goff, S. P., and P. Berg.** 1979. Construction, propagation and expression of simian virus 40 recombinant genomes containing the *Escherichia coli* gene for thymidine kinase and a *Saccharomyces cerevisiae* gene for tyrosine transfer RNA. J. Mol. Biol. **133:**359–383.
19. **Grabowski, P. J., S. R. Seiler, and P. A. Sharp.** 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell **42:**345–353.
20. **Greenspan, D. S., and S. M. Weissman.** 1985. Synthesis of predominantly unspliced cytoplasmic RNAs by chimeric herpes simplex virus type 1 thymidine kinase-human β-globin genes. Mol. Cell. Biol. **5:**1894–1900.
21. **Gruss, P., and G. Khoury.** 1981. Expression of simian virus 40-rat preproinsulin recombinants in monkey kidney cells: use of preproinsulin RNA processing signals. Proc. Natl. Acad. Sci. USA **78:**133–137.
22. **Gruss, P., C. J. Lai, R. Dhar, and G. Khoury.** 1979. Splicing as a requirement for the biogenesis of functional 16S mRNA of simian virus 40. Proc. Natl. Acad. Sci. USA **76:**4317–4321.
23. **Hamer, D. H., and P. Leder.** 1979. SV40 recombinants carrying a functional RNA splice junction and polyadenylation site from the chromosomal mouse β^{maj} globin gene. Cell **17:**737–747.
24. **Hamer, D. H., and P. Leder.** 1979. Splicing and the formation of stable RNA. Cell **18:**1299–1302.
25. **Hamer, D. H., K. D. Smith, S. H. Boyer, and P. Leder.** 1979. SV40 recombinants carrying rabbit β-globin gene coding sequences. Cell **17:**725–735.
26. **Hardison, R. C., E. T. Butler, E. Lacy, T. Maniatis, N. Rosenthal, and A. Efstratiadis.** 1979. The structure and transcription of four linked rabbit β-like globin genes. Cell **18:**1285–1297.
27. **Hashimoto, C., and J. A. Steitz.** 1986. A small ribonucleoprotein associates with the AAUAAA polyadenylation signal in vitro. Cell **45:**581–591.
28. **Kedes, L. H.** 1978. Histone genes and histone messengers. Annu. Rev. Biochem. **48:**837–870.
29. **Leff, S. E., R. M. Evans, and M. G. Rosenfeld.** 1987. Splice commitment dictates neuron-specific alternative RNA processing on calcitonin/CGRP gene expression. Cell **48:**517–524.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Maniatis, T., and R. Reed.** 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. Nature (London) **325:**673–678.
32. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific cleavages. Methods Enzymol. **65:**499–580.
33. **McKnight, S. L.** 1980. The nucleotide sequence of the herpes thymidine kinase gene. Nucleic Acids Res. **8:**5949–5966.
34. **Mercola, M., X. Wang, J. Olsen, and K. Calame.** 1983. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. Science **221:**663–665.
35. **Mertz, J. E., and P. Berg.** 1974. Defective simian virus genomes: isolation and growth of individual clones. Virology **62:** 112–124.
36. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41:**845–855.
37. **Moreau, P., R. Hen, B. Wasylyk, R. Everrett, M. P. Gaub, and P. Chambon.** 1981. The SV40 72 bp repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. **9:**6047–6069.
38. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10:**459–472.
39. **Mulligan, R. C., and P. Berg.** 1981. Factors governing the expression of a bacterial gene in mammalian cells. Mol. Cell. Biol. **1:**449–459.
40. **Mulligan, R. C., and P. Berg.** 1981. Selection for animal cells that express the *Escherichia coli* gene for xanthine-guanine phosphoribosyl transferase. Proc. Natl. Acad. Sci. USA **78:** 2072–2076.
41. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit β-globin in cultured monkey kidney cells following infection with a SV40-rabbit β-globin recombinant genome. Nature (London) **277:**108–114.
42. **Nagata, S., N. Mantei, and C. Weissmann.** 1980. The structure of one of the eight or more distinct chromosomal genes for interferon alpha. Nature (London) **287:**401–408.
43. **Nunberg, J. H., R. J. Kaufman, A. C. Y. Chang, S. N. Cohen, and R. T. Schimke.** 1980. Structure and genome organization of the mouse dihydrofolate reductase gene. Cell **19:**355–364.
44. **Padget, R. A., M. M. Konarska, M. Aebi, H. Hornig, C. Weissmann, and P. A. Sharp.** 1985. Nonconsensus branch-site sequences in the in vitro splicing of mutant rabbit β-globin genes. Proc. Natl. Acad. Sci. USA **82:**8349–8353.
45. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus and viral DNA. J. Virol. **31:**360–369.
46. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33:**741–748.
47. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46:**681–690.

48. **Rozek, C. E., and N. Davidson.** 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. Cell **32**:32–44.
49. **Sharp, P. A.** 1987. Splicing of messenger RNA precursors. Science **235**:766–771.
50. **Simonsen, C. C., and A. Levinson.** 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. Proc. Natl. Acad. Sci. USA **80**:2495–2499.
51. **Subramani, S., R. C. Mulligan, and P. Berg.** 1981. Expression of the mouse dihydrofolate reductase complementary deoxyribonucleic acid in simian virus 40 vectors. Mol. Cell. Biol. **1**:854–864.
52. **Tonegawa, S.** 1983. Somatic generation of antibody diversity. Nature (London) **302**:575–581.
53. **Treisman, R., S. H. Orkin, and T. Maniatis.** 1983. Specific transcription and RNA splicing defects in five cloned β-thalassemia genes. Nature (London) **302**:591–596.
54. **Treisman, R., J. Novak, J. Favaloro, and R. Kamen.** 1981. Transformation of rat cells by an altered polyoma virus genome expressing only the middle-T protein. Nature (London) **292**: 595–600.
55. **Villarreal, L. P., R. T. White, and P. Berg.** 1979. Mutational alterations within the simian virus 40 leader segment generate altered 16S and 19S mRNAs. J. Virol. **29**:209–219.
56. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **78**:1441–1445.
57. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure. Nucleic Acids Res. **75**:1175–1193.
58. **White, R. T., P. Berg, and L. P. Villarreal.** 1982. Simian virus 40-rabbit β-globin recombinants lacking late mRNA splice sites express cytoplasmic RNAs with altered structures. J. Virol. **42**: 262–274.
59. **Wieringa, B., E. Hofer, and C. Weissmann.** 1984. A minimal intron length but no specific internal sequence is required for splicing the large rabbit β-globin intron. Cell **37**:915–925.
60. **Wieringa, B., F. Meyer, J. Reiser, and C. Weissmann.** 1983. Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene. Nature (London) **301**:38–43.
61. **Yanofsky, C.** 1981. Attenuation in the control of expression of bacterial operons. Nature (London) **289**:751–758.
62. **Zeitlin, S., and A. Efstratiadis.** 1984. In vivo splicing products of the rabbit β-globin pre-mRNA. Cell **39**:589–602.
63. **Ziff, E. B.** 1980. Transcription and RNA processing by the DNA tumor viruses. Nature (London) **287**:491–499.