Intervening sequences increase efficiency of RNA 3′ processing and accumulation of cytoplasmic RNA

Manley T.F.Huang+ and Cornelia M.Gorman*
Department of Cell Genetics, Genentech, Inc., 460 Pt San Bruno Blvd., CA 94080, USA

Received October 12, 1989; Revised and Accepted December 11, 1989

ABSTRACT

Two expression vectors were constructed that differ only in the presence (+) or absence (−) of an intervening sequence (IVS) in their 5′-untranslated leaders. Transient transfection into four mammalian cell lines resulted in higher levels of the indicator protein (CAT) from the IVS(+) vector (6 to 50-fold). Cytoplasmic RNA concentrations in 293s and HeLa cell lines corresponded directly to resultant protein levels; measurements in 293s cells of transcription initiation and elongation, steady-state total nuclear RNA, and cytoplasmic RNA stability, were equivalent for the two vectors. Surprisingly, the amount of poly(A)+ nuclear RNA was greater from the IVS(+) vector. Since this difference matches the ratio seen with polyadenylated cytoplasmic RNA, our results imply that splicing is coupled to a polyadenylation/transport pathway.

INTRODUCTION

Although precise roles for the intervening sequences (IVS) of eucaryotic genes remain undefined, RNA splicing is clearly a functionally diverse process. Among its effects are the ability to arrange exons encoding different functional domains (for review see 1); the assembly of developmentally specific mRNAs via alternative splicing (2); and the control of tissue-specific gene expression such as the P element transposase gene (3) and transformer (*tra*) locus in Drosophila (4). Mechanistically, several studies have characterized the nuclear components required for spliceosome assembly and the temporal sequence of splicing for RNA polymerase II transcripts (for review see 5).

Much less, however, is known about the regulatory roles of splicing. Early studies of gene expression utilized DNA viruses such as SV40 and adenovirus to study splicing and as models for eucaryotic gene expression. Soon after the discovery of splicing (6, 7), several experiments with SV40 virus indicated that splicing was a prerequisite for the production of stable cytoplasmic 16S mRNA (8, 9, 10), 19s mRNA (11), and small t-antigen mRNA (12). In the case of an unspliceable 16s message, transcription appeared to be unaltered (13), and the deficiency in the accumulation of SV40 late RNA sequences could be rescued by the addition of a functional heterologous splice from the β-globin gene (14, 15). However, viral mutants have been identified that can express the E1A gene of adenovirus (16) and the 19s gene of SV40 (17) even though there are deletions within each gene which include part of the IVS and span the 3′-splice sequence. These mutants produced viral titres significantly below those of wild type. Unfortunately, no conclusions could be made about the functional role of splicing during wild type infection other than that splicing is not always essential for the production of functional transcripts.

Due to the apparent requirement for splicing, some early expression vectors were used to express a variety of genes by inserting cDNAs downstream of the splice signals of the SV40 late region (18, 19). Although protein synthesis occurred in these instances, no attempts were made to determine the functional role of splicing by directly comparing vectors that contained or lacked the exact IVS.

In contrast, many cDNAs were subsequently expressed in cell culture from vectors which excluded a splice [an SV40-preproinsulin hybrid (9), DHFR (20), polyoma middle T antigen (21), chicken thymidine kinase (TK) (22), and apolipoprotein D (23)]. Additionally, cellular genes for histone (24), interferon (13), and heat shock proteins (25, 26) were cloned and found to be devoid of introns. Although these genes provide examples where splicing is apparently unnecessary for expression, Callis *et al.* (27) have determined that the addition of a splice to the 5′-untranslated leader can significantly increase expression from an unspliced parental vector. In that study, the first intron of the maize alcohol dehydrogenase gene was added to the leader of either adh, CAT, neo, or luciferase mRNA; expression increased from 5 to 170-fold in maize cells, with protein synthesis directly related to levels of cytoplasmic RNA. In addition, Buchman and Berg (28) have recently demonstrated that splicing leads to an increase in the levels of both nuclear and cytoplasmic poly(A)+ RNA, and they suggested that (i) transcription initiation, (ii) nuclear RNA stability, or (iii) cleavage and polyadenylation could be responsible for such an effect.

In order to directly examine how splicing enhances expression and affects mRNA processing, we have constructed a set of vectors which differ only in the presence or absence of an IVS in their 5′-untranslated leaders. These plasmids were individually transfected into four mammalian cell lines from which protein

* To whom correspondence should be addressed

+ Present address: GenPharm, Inc , 180 Kimball Way, South San Francisco, CA 94080, USA

levels (CAT) (29) or the cellular distribution of mRNA were determined. Depending upon cell type, between 6 to 50-fold more CAT activity was obtained from the vector containing a splice. Our analysis of RNA maturation shows that transcription initiation and cytoplasmic stability were the same for IVS(+) and IVS(−) vectors, yet there is a marked increase in the accumulation of polyadenylated nuclear RNA from the IVS(+) vector. Since accumulation of steady-state total nuclear RNA for each of these vectors is equal, the difference between expression of the IVS(+) and IVS(−) vectors is focused at the point of polyadenylation. The increase in expression most likely arises from a postranscriptional mechanism which favors the 3′-processing and transport of IVS(+) mRNA.

MATERIALS AND METHODS

Cell lines. Two human cell lines used in this study were a subclone of 293 human embryonic kidney cells (30) adapted for growth in suspension in Joklik's media referred to herein as 293s (received from B.Stillman); and HeLa. African green monkey kidney cells, CV-1, and Chinese hamster ovary cells, CHO, were two other mammalian cell lines employed. All cells were maintained by serial passage in F12:DMEM (Gibco) 50:50 with added glutamine and without antibiotics.

Construction of bacterial plasmids. Two plasmids were constructed that express the bacterial gene encoding chloramphenicol acetyl-transferase (CAT). Both were pML-based plasmids containing the SV40 virus early promoter with its enhancer, beginning with the 72 base pair repeats and continuing through to position +6 with respect to the early-early 1 start sites (31). Immediately following the promoter, in plasmid pML.SIS, a portion of the 5′-untranslated leader from the adenovirus-major-late (AML) region was inserted which contains nucleotides +14 to +200 with respect to the CAP site. This region contains all but the first 13 nucleotides of the first exon of the tripartite leader, plus nucleotides +42 to +200 of the first intervening sequence (IVS). A synthetic oligonucleotide was inserted which merges with the adenovirus intron to provide a functional splice acceptor sequence derived from an IgG variable region (32). This hybrid 5′-regulatory region was used to express the CAT gene, terminating with the late polyadenylation signal from SV40 virus.

The second vector, pML.I-.CAT, was identical to pML.SIS.CAT except for the precise removal of the IVS as it would occur in cDNA. Plasmid pML.SIS.CAT was cut at two restriction sites bordering the intron (*ClaI* and *PstI*) to remove a 292 bp fragment. A matching synthetic oligonucleotide linker was inserted, the sequence of which lacks the 230 nt IVS.

A third vector was used as an internal control for transfection efficiency. Plasmid pRSV.hGH expresses human growth hormone under the control of the Rous sarcoma virus long terminal repeat (33).

Transfections and transient assays. DNAs were transfected into mammalian cells by the calcium phosphate precipitation method (34) as modified by Gorman *et al.* (35). Each 60 millimeter dish was treated with 0.5 ml of precipitate containing 2.5 μg of CAT expression vector and 100 ng of RSV.hGH DNA. For the preparation of CAT RNA, 100 mm dishes were used with twice the amount of precipitate. At approximately 48 hr following transfection, a sample of the media was removed for the assay of human growth hormone (Hybritech) and the cells harvested for CAT assays by the ^{3}H-sodium acetate method modified by Nordeen *et al.* (36). Multiple timepoints were taken to assure that measurements were within the linear range. The relative levels of CAT activity were adjusted to reflect the basal levels of hGH from each sample.

RNA purification. Cytoplasmic and nuclear RNAs were obtained from 293s and HeLa cells. The isolation of RNA from 293s cells was performed from between 24−48 hr after transfection. In each experiment, relative transfection efficiency was monitored by the expression of human growth hormone (hGH) from a cotransfected vector, pRSV.hGH. Separate dishes were harvested from the same transfection to measure CAT activity. Cells were lysed in a solution containing Triton-X followed by Dounce homogenization (37) and staining by trypan blue to ensure complete lysis Nuclei were obtained by low speed centrifugation at 2,000 RPM for 3 min at 4°C in an HSB-4 rotor. The supernatant containing cytoplasmic RNA was removed, and the nuclei thoroughly resuspended in lysis buffer Centrifugation was repeated for both nuclear and cytoplasmic fractions and the cytoplasmic lysate carefully withdrawn to avoid cellular contamination. Cytoplasmic RNA was purified by mixing the cytoplasmic lysate with an equal volume of buffer containing 7 M urea, 0.35 M NaCl, 10 mM Tris-HCl (pH 7.5), 10 mM EDTA, and 1%SDS, followed by extracting once with phenol (saturated with Tris-HCl, pH 8.0, and containing 0.1% hydroxyquinaldine); once with phenol:chloroform (50:50); and once with chloroform. The RNA was precipitated with 2.5 volumes of ethanol at −20°C.

For nuclear RNA, the nuclear pellet was resuspended in 5 ml of 2 M LiCl, 4 M urea using polytron disruption (15 seconds), and placed on ice for over 2 hr. The RNA was pelleted by centrifugation for 1 hr at 8,000 RPM in an HSB-4 rotor. The RNA pellet was resuspended in 5 ml of TES (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 0.1% SDS) plus 200 μg/ml proteinase K (BM) and incubated for 10 min at 37°C. The RNA was extracted with phenol and chloroform as described above and precipitated by the addition of 0.1 volumes of 3 M NaOAc and 2.5 volumes of ethanol. After precipitation at −20°C, both nuclear and cytoplasmic RNAs were pelleted and resuspended in 400 μl TE/0.3 M NaOAc and precipitated again in ethanol on dry ice or at −20°C. After pelleting in a microfuge, the RNA was resuspended in 80 μl distilled water and mixed with 120 μl HSB (0.5 M NaCl, 50 mM $MgCl_2$) and then treated with DNAseI at 50 μg/ml for 10 min at 37°C. This was followed by the addition of 0.2x volumes of 5x SET (50 mM Tris-HCl, pH 7.5; 0.5% SDS; and 25 mM EDTA) and proteinase K to 200 μg/ml and incubation at 42°C for 10 min. The RNA was extracted once with phenol/chloroform (50:50) and precipitated in ethanol at −20°C. An OD 260:280 ratio of greater than 1.9 was routinely obtained.

The procedure for isolating RNA from HeLa cells was identical to that for 293s, except that the cells were lysed with NP-40 24 hr post-transfection, and the nuclei purified by centrifugation through a 25% sucrose layer (38).

Nuclear run-on analysis. Nuclei from 293s cells cotransfected with pML.SIS.CAT and pRSV.hGH, pML.I-.CAT and pRSV.hGH, or pRSV.hGH alone were isolated as described above (37). Cells were scraped from 100 mm plates 16 hr after transfection and lysed by Dounce homogenization (18 strokes, type B plunger in Triton-X lysis buffer). Approximately 2×10^7

nuclei were stored in 25 μl storage buffer (50 mM Tris-Cl, pH 8.3; 5 mM $MgCl_2$; 0.1 mM EDTA; 40% glycerol) in liquid nitrogen prior to labeling. Nuclear run-on conditions were essentially those of Linial *et al.* (39) with some modifications. Thawed nuclei were mixed with 10 μl 5x run-on buffer (5x consists of 25 mM Tris-Cl, pH 8.0; 12.5 mM $MgCl_2$; 750 mM KCl), 12.5 μl 50% glycerol, 1.3 μl 10 mM rATP, GTP, and CTP, 1 mCi dry ^{32}P-UTP (1:1 mixture of >3,000 Ci/mM and >600 Ci/mM specific activity), and 1 μl RNAsin (Promega, 40 U/μl). Nuclei were incubated for 20 min at 27°C. Next, 12.5 μl 30 mM $MgCl_2$, 75 mM $CaCl_2$ plus 3 μl DNAseI (Worthington DPRF, 2.5 mg/ml) were added and incubated for 10 min at 37°C. This was followed by the addition of 5μl yeast tRNA (20 mg/ml) 18 μl 5X SET (5% SDS; 25 mM EDTA; 50 mM Tris-HCl, pH 7.5), and 1μl proteinase K (Boehringer Mannheim, 20 mg/ml) and incubation for 30 min and 37°C. The reaction mixture was extracted once with phenol/chloroform (1:1) and the interphase back-extracted twice with TES (10 mM Tris-Cl, pH 8.0; 1 mM EDTA; 0.1% SDS). The pooled aqueous phases were extracted once with chloroform and precipitated in ethanol in the presence of 2 M ammonium acetate. The RNA was centrifuged, rinsed once with 70% ethanol, and dried under vacuum. The pellet was resuspended in 50 μl TES and the RNA precipitated by the addition of 50 μl ice-cold TCA (10% TCA, 60 mM sodium pyrophosphate) and incubation on ice for 5 min. RNA was pelleted by centrifugation in a microfuge for 5 min. The pellet was rinsed twice (5% TCA, 30 mM sodium pyrophosphate), resuspended in 200 μl 0.1 M sodium acetate, and precipitated in ethanol on dry ice. The subsequent RNA pellet was resuspended in 90μl TE plus 10 μl 1 M NaOH and kept on ice for 7 min, after which 33 μl 1 M HEPES was added. RNA was precipitated in ethanol a final time, and resuspended in 250 μl distilled water treated with diethylpyrocarbonate (0.1%). Approximately 1.5 to 3.0×10^7 CPMs/reaction were obtained from this procedure. Total counts were adjusted to approximately 1×10^7 CPM/ml of hybridization buffer (2 ml per filter) and hybridized to filters containing approximately 3 μg single-stranded RNA probes for greater than 12 hr at 42°C. Filters were rinsed twice in 2x SSC (30 min/rinse at 70°C) and then for 1 hr in 0.1×SSC at 70°C (three changes of wash, 20 min each). Single-stranded RNA probes were generated as either SP6 or T7 transcripts from the entire CAT coding region or 285 bp of human growth hormone coding sequence subcloned in pGEM-4. The RNAs were denatured in glyoxal (40) and fixed to nitrocellulose filters using a Schleicher and Schuall slot-blot apparatus.

S1 Analysis. Four synthetic oligonucleotide probes were prepared whose sequences are homologous to the 5'-untranslated region of the CAT expression vectors. The four probes vary in length, depending upon the extent of homology which begins at +1 with respect to the CAP site. All probes contain identical 3'-tails which were designed to mismatch the plasmid sequences upstream of the CAP site, and to minimize any intrastrand homology that would permit hairpin formation.

S1 digest reactions were performed by coprecipitating 5−10 μg of total nuclear RNA with 50 to 500 fmole of ^{32}P-end-labeled probe (100 counts per second total). The RNA/probe pellet was resuspended in 24 μl DW and dissolved by shaking for 10−15 min on a VWR multitube vortexor. Six μl of 5x annealing buffer was added (0.2 M Pipes, pH 6.4; 2 M NaCl; 5 mM EDTA) and the probe annealed to RNA for 2 hr at 50°C. Three hundred units of S1 nuclease (BRL) were added in 250 μl digest buffer (30 mM Na-acetate, pH 5.0; 250 mM NaCl; 1 mM $ZnSO_4$; 5% glycerol) and incubated for 2 hr at 28°C. The digestion was stopped by the addition of phenol/chloroform (50:50) and the aqueous phase precipitated directly with 2.5 volumes of ethanol after the addition of 20 μg of carrier tRNA. The nucleic acid was pelleted and resuspended in gel loading buffer consisting of 80% formamide, 0.1 M NaOH, and 0.02% bromophenol blue and xylene cyanol dyes. The pellet was dissolved by vortexing, as above, followed by denaturation at 100°C for 2 min. Samples were then electrophoresed in 10% acrylamide, 8 M urea gels (BRL sequencing manual) in 0.5x TBE buffer under 30 mA current for 30 min Gels were fixed for 15 min in 6.25% acetic acid, 20% methanol and dried onto filter paper before exposure to Kodak XAR-5 film.

Primer extension analysis. Three oligonucleotide primers were used to examine both cytoplasmic and nuclear RNA. Two were specific for separate regions of the CAT coding sequence, designed to prime cDNA synthesis from either the 5'- or 3'-ends of the CAT gene. The 5'-end primer was a gift of R. Tjian and contained the sequence: 5' GCCATTGGGATATATCAAC-GGTGG-3'. The 3'-end primer contained the sequence: 5'-GGCGTAGCACCAGGCGTTTAAG GGC-3'. As an internal control, a β-actin primer was obtained from E.Jakobovits. Its sequence was derived from the rat actin gene, but was sufficient to prime the synthesis of human β-actin: 5'CGGCCCACGAT-GGAGGGGA AGACGGCCCGGGGAGC-3'. Approximately 20 μg of cytoplasmic RNA or 5−10 μg of nuclear RNA plus 20 μg of carrier tRNA were precipitated with 50−500 fmole of ^{32}P-end-labeled primer, and the reactions performed as described (33). The reactions were terminated by phenol/chloroform (50:50) extraction and the nucleic acids precipitated in ethanol. The samples were prepared for electrophoresis, as above, and analyzed in 6% acrylamide, 8 M urea gels.

Determination of relative RNA stability. The stability of cytoplasmic CAT RNA was determined by quantitative primer extension analysis of RNA purified at various times after the inhibition of RNA synthesis by actinomycin C (Sigma). 5μg/ml of actinomycin C was first shown to inhibit the uptake of ^{3}H-uridine within 1 hr after its addition to 293s cells by measuring the incorporation of TCA precipitable ^{3}H-uridine. Subsequently, cytoplasmic RNA was prepared from cells transfected with either pMLSIS.CAT or pML.I-.CAT by NP-40 lysis (38) and then phenol extraction as described above. Primer extension reactions were performed as described earlier using either the CAT primer from the 5'-coding region, or a primer for the HSP70 mRNA derived from the sequence reported by Hunt and Morimoto (25): 5'- GGATCCGCGAGAAGAGCTCGGTCC-3' (a gift of Rein Streiker). Extension products were electrophoresed in 6% acrylamide, 8 M urea gels.

RESULTS

CAT Expression Vectors Containing Adenovirus-Major-Late (AML) Leader Sequences. The vectors used in this study are both pML-based plasmids. They contain the SV40 virus early promoter region with its enhancer, the bacterial CAT gene, and the SV40 virus late polyadenylation signal (see Materials and Methods and Fig. 1). They differ in the amount of the adenovirus-major-late (AML) leader sequence present between the SV40

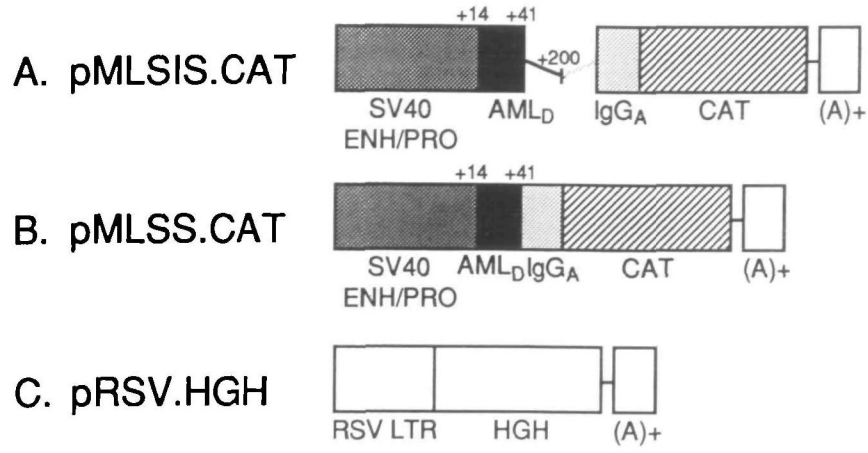

Figure 1. Linear maps of IVS(+) and IVS(−) expression plasmids The two experimental expression vectors both contain the SV40 early region enhancer and promoter, and leader sequence up through position +6 with respect to the SV40 early-early 1 start sites The untranslated leader of pMLSIS CAT (A) contains a hybrid splice derived from the region +14 to +200 of the adenovirus type 5 tripartite leader fused to the 3′-splice sequence of an IgG variable region (32) This adenovirus sequence contains the 5′-splice sequence of the first exon plus 149 bases of adjacent intervening sequence The splice is followed by the bacterial CAT gene and SV40 late polyadenylation signal pML I- CAT [or pML SS CAT] (B) is identical to pMLSIS CAT except for the precise deletion of the intervening sequence pRSV hGH (C) contains the human growth hormone gene linked to the long terminal repeat of Rous sarcoma virus, for use as an internal control marker for transfection efficiency

promoter and the CAT coding sequence. The composite tripartite leader (TPL) of adenovirus contains three exons and two IVS which are coordinately spliced and linked to each of the different late genes of the virus (21). Plasmid pMLSIS CAT (Fig. 1A) contains a continuous region of the AML leader from position +14 to +200 with respect to the CAP site of the AML transcripts. It includes all but the first 13 bases of the first exon as well as bases +42 to +200 of the first IVS of the TPL. This sequence provides a functional 5′-splice sequence. Fused to this leader is a hybrid intron/3′-splice sequence derived from an IgG variable region (32). This plasmid is also referred to as the IVS(+) vector and its RNA as IVS(+).

In plasmid pML.I-.CAT (Fig 1B) the IVS has been deleted. A restriction fragment containing the IVS from pMLSIS.CAT was replaced by a synthetic oligonucleotide linker that connects the consensus 5′- and 3′-splice sequences precisely as it would occur in spliced cDNA. This plasmid is also referred to as the IVS(−) vector and its RNA as IVS(−).

In order to compensate for differences in transfection efficiency, a marker plasmid was cotransfected in each experiment. pRSV.hGH, encoding human growth hormone (hGH) (Fig. 1C), was added at 4% the concentration of CAT DNA, and the supernatants of each plate assayed for secreted hGH activity. CAT assays from each plate were performed and the relative levels adjusted for differences in the basal activity of hGH.

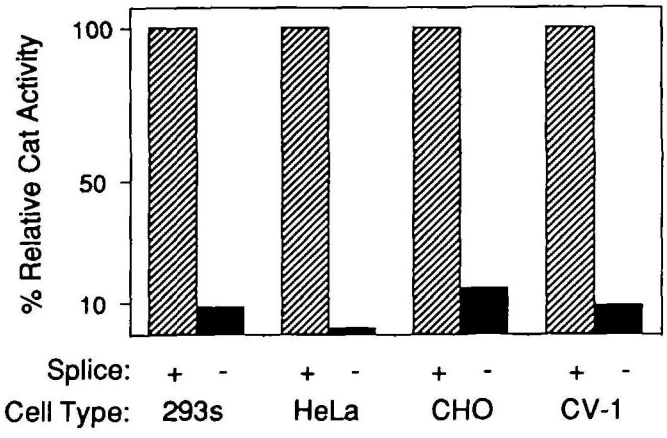

Figure 2. Relative CAT expression of IVS(+) vs IVS(−) vectors in mammalian cell lines Plasmids pMLSIS.CAT ('+' splice) and pML I- CAT '−' splice) were cotransfected with the control plasmid, pRSV hGH, into four mammalian cell lines 293s, HeLa, CHO, and CV-1 Cell extracts were prepared from between 40−48 hr after transfection and assayed for CAT activity Histograms depict the relative activity of CAT based upon 100% activity of pMLSIS CAT within each cell type. Bars represent an average of two to four experiments The ratio of human growth hormone (hGH) expression between the IVS(+) and IVS(−) vectors ranged from 0.8 to 2 0 in HeLa cells (four expereiments), 1 2 to 2 2 in 293s cells (four experiments 0 8 to 1.5 in CV-1 cells (2 two experiments), and 0 8 to 1 4 in CHO cells (three experiments)

Transient CAT expression from spliced vs. unspliced mRNA. The relative levels of CAT expressed from plasmids pMLSIS.CAT and pML.I-.CAT were determined in 293s, HeLa, CHO, and CV-1 cells (Fig. 2). Levels of CAT activity are expressed as a percentage of CAT obtained from pMLSIS.CAT for each cell line and have been normalized for differences in transfection efficiency (see Materials and Methods). Therefore, comparison can only be made between the two vectors within a particular cell type. In all cases, the presence of the IVS in the 5′-leader (pMLSIS.CAT) led to significantly greater CAT activity (between 6−50 fold). The most pronounced effect occurred in HeLa cells where pMLSIS.CAT produced an adjusted average of 50−fold more CAT protein.

Analysis of steady-state RNA in 293s and HeLa cell lines. Due to the general difference in CAT expression between pMLSIS.CAT and pML.I-.CAT, we analyzed the levels of CAT-specific RNA in both the cytoplasm and nucleus of 293s and HeLa cells. Total cytoplasmic RNA was quantitated by primer extension reactions (Fig. 3) and compared to levels of CAT activity in duplicate transfections. An oligonucleotide primer was prepared that initiates synthesis by reverse transcriptase from the 5′-end

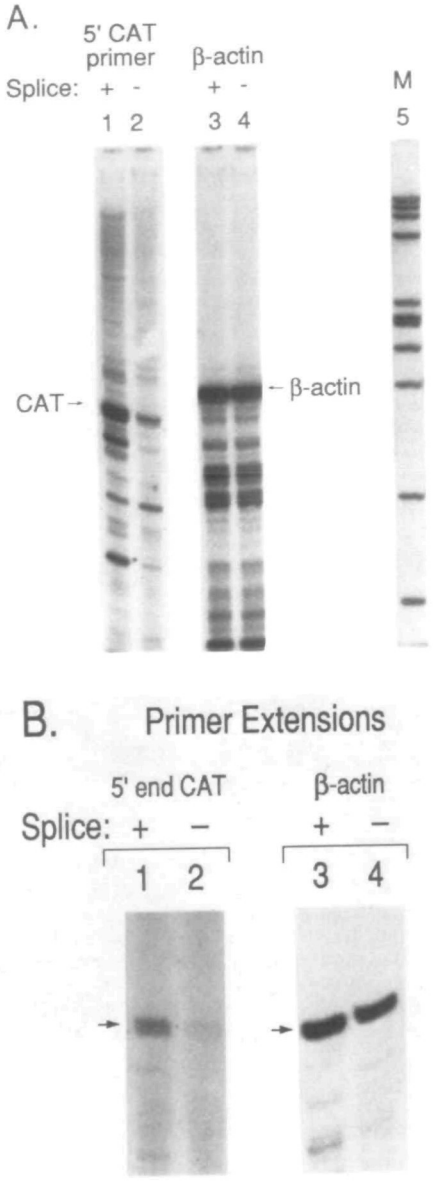

Figure 3. Quantitative primer extension analysis of CAT-specific RNA 293s (A) and HeLa (B) cells were transfected with either pMLSIS CAT ('+' slice) or pML I- CAT ('−' splice) DNA. Approximately 18 hr after transfection, cytoplasmic RNA was harvested and total RNA samples were annealed to 5'-radiolabeled oligonucleotide primers specific for either CAT (20 μg RNA, lanes 1 and 2) or cellular β-actin (10 μg RNA, lanes 3 and 4) coding sequences The primers were extended using avian myeloblastosis virus reverse transcriptase and electrophoresed through 6% polyacrylamide, 8 M urea denaturing gels Lane 5 (M) contains 5'-end-labeled phi-X174/*Hae*III DNA markers. The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1 5 for 293s cells, and 0 8 for HeLa cells The uncorrected ratio of CAT activity was 20-fold for 293s and 50-fold for HeLa cells. For the quantitative estimate the major CAT specific band was analyzed by densitometry

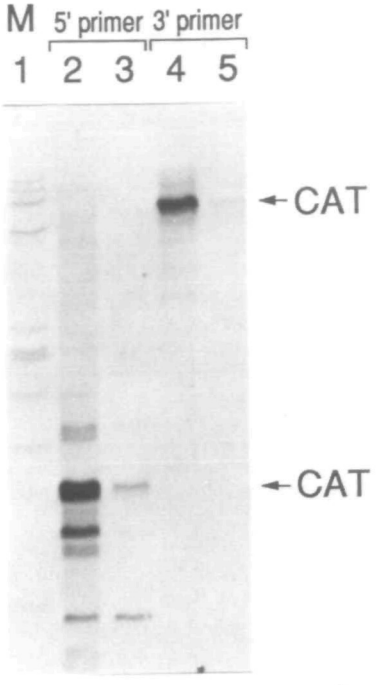

Figure 4. Comparison of IVS(+) vs IVS(−) poly(A)$^+$ RNA using primers specific for the 5'- and 3'-coding regions of CAT 293s cells were transfected and processed for cytoplasmic RNA as in Fig 3 Poly(A)$^+$ mRNA was isolated using oligo(dT) cellulose and divided into equal fractions for primer extensions Two primers were used which hybridize to opposite ends of the CAT coding sequence and generate extension products of 155 (5'-end) and 823 nucleotides (3'-end) Lane 1· phi-X174/*Hae*III DNA markers Lanes 2 and 4 RNA from cells transfected with pMLSIS CAT ('+' splice) Lanes 3 and 5 RNA from cells transfected with pML I- CAT ('−' splice) Arrows point to CAT-specific bands of the expected lengths for each primer The ratio of hGH expression between the IVS(+) and IVS(−) vector was 0 8

of the CAT coding sequence at nucleotide +165 relative to the CAP site of a fully processed message. The ratio of IVS(+) to IVS(−) RNA was approximately 15 to 20-fold for 293s cells as determined by scanning densitometry (Fig. 3A, lanes 1 and 2). In this experiment, the ratio of CAT activity for the two vectors was 20-fold. Therefore, the levels of CAT activity corresponded directly to the amount of CAT mRNA present in the cytoplasm, and the difference between the IVS(+) and IVS(−) messages was not due to altered rates of translation. Lanes 3 and 4 (Fig. 3A) contain primer extensions of duplicate RNA samples using a β-actin primer to ensure accurate comparison of RNA samples (note that henceforth, values for CAT activity and RNA concentrations are not adjusted for differences in transfection efficiencies; however, the relative levels of hGH expression are noted in the figure legends). The same analysis of HeLa cytoplasmic RNA is shown in Fig. 3B. Whereas the ability to transfect HeLas and monitor RNA from transient transfections is much less efficient than for 293 cells, we observed the same qualitative effect. Note the high ratio of IVS(+) to IVS(−) RNA (lanes 1 and 2) with the IVS(−) RNA virtually undetectable. In this instance, the ratio of IVS(+) to IVS(−) CAT activity was 50:1.

Further analysis, using poly(A)$^+$ selected cytoplasmic RNA from 293s cells, was performed to compare the above 5'-CAT primer to a second primer that initiates from the 3'-coding region at position +700 (Fig. 4). The ratio of IVS(+) to IVS(−) RNA from the 5'-region was within 2-fold of that of the 3'-region (by scanning densitometry). We thus concluded that cytoplasmic RNA contains appropriately spliced and polyadenylated RNA, and that the 5'-primer could be used reliably for the analysis of total cytoplasmic RNA samples. Moreover, it is fully processed, poly(A)$^+$ RNA that is directly related to CAT activity levels.

Stability of cytoplasmic RNA. Even though the spliced RNA from pMLSIS.CAT has the same primary structure as that from

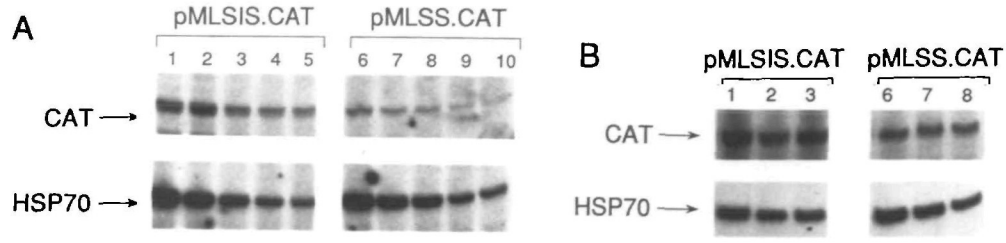

Figure 5. Cytoplasmic stability of IVS(+) vs IVS(−) mRNA The relative stability of IVS(+) (pMLSIS.CAT) vs. IVS(−) (pML I-.CAT) RNA was examined in transcriptionally arrested 293s cells over the course of 8 5 hr. First, 293s cells were treated with actinomycin C approximately 18 hr after transfection At subsequent time points, total cytoplasmic RNA was harvested and analyzed by primer extension. Individual reactions using primers for either the transfected CAT gene (top line in each panel, identical primer as in Figs 3 and 4) or the endogenous cellular gene, heat shock protein HSP70 (bottom line in each panel) are shown (A) Long time course RNAs were harvested immediately upon the addition of actinomycin C (lanes 1 and 6) and at several times thereafter 1 hr (lanes 2 and 7), 3 5 hr (lanes 3 and 8), 6 hr (lanes 4 and 9), and 8 5 hr (lanes 5 and 10). 20 μg of IVS(+) RNA and 30 μg of IVS(−) were annealed to the 5′ CAT primer, 10 μg/each were annealed to the HSP70 primer The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1 0 throughout the time course of the experiment (B) Short time course. The same conditions were applied as in (A) with RNA harvested at closer time points after the addition of actinomycin C 3 hr (lanes 1 and 6), 4 hr (lanes 2 and 7), and 5 hr (lanes 3 and 8) 10 μg of IVS(+) RNA and 30 μg of IVS(−) RNA were annealed to the CAT primer, 10 μg/each were annealed to the HSP70 primer The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1 7 (lanes 1 and 6), 1.3 (lanes 2 and 7), and 1 0 (lanes 3 and 8)

pML.I-.CAT, the preferential accumulation of the spliced message could partially have been due to differences in RNA stability. During processing and transport, the messages might have developed unequal half-lives by forming different secondary structures or complexing with separate proteins. We therefore quantitated the rates of RNA decay by blocking RNA synthesis with actinomycin C and measuring the amount of CAT-specific RNA remaining over the course of 8.5 hr by quantitative primer extension analysis. As an internal control, heat shock protein HSP70 RNA was also examined. Fig. 5A shows the amounts of both CAT- and HSP70-specific RNAs from cells harvested at time zero (addition of actinomycin C, lane 0), at 1 hr (after the blockage of ^{3}H-uridine uptake, lane 1), and at 2.5 hr intervals thereafter (lanes 2−4). The relative stability of IVS(+) (left panel) and IVS(−) (right panel) messages was approximately the same. To more precisely measure mRNA stability, a shorter time course, shown in Fig. 5B, consists of RNA harvested at 3, 4, and 5 hr after the addition of actinomycin C. These results are representative of three separate experiments from which we conclude there is no difference in the relative stability of the two messages.

Analysis of transcription initiation in 293s nuclei. A number of genes have been shown to contain intervening sequences which increase transcription initiation (41, 42, 43, 44). Such a sequence has been identified within the first intron of the AML TPL and responds in an enhancer-like manner in the presence of adenovirus in COS cells (45). Since this sequence is contained within the leader of pMLSIS.CAT, it is possible that it provides enhancer activity in pMLSIS.CAT, and thus it deletion was reducing CAT expression from pML.I-.CAT. We therefore directly compared the rates of transcription initiation from pMLSIS.CAT and pML.I-.CAT by nuclear run-on reactions from cells cotransfected with RSV.hGH (Fig. 6). Anti-sense transcripts from the entire CAT coding region or a 285 bp fragment of human growth hormone were generated by SP6 or T7 RNA polymerase and fixed to nitrocellulose filters. A parallel reaction labeled with ^{32}P-CTP was analysed to confirm that the transcripts were almost entirely full length (data not shown). Isolated nuclei from transfected 293s cells were used to prepare nascent RNA transcripts labeled in vitro with ^{32}P-UTP. Hybridization of the labeled transcripts to the anti-sense fragments indicated that initiation of transcription was approximately the same for both

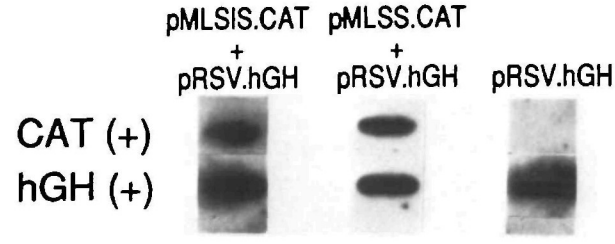

Figure 6. Nuclear run-on analysis of IVS(+) vs IVS(−) vectors 293s cells were cotransfected with pRSV hGH and either pMLSIS CAT ('+' splice) or pML I- CAT ('−' splice) Isolated nuclei were used to prepare nascent ^{32}P-labeled transcripts Separate extracts were hybridized to duplicate sets of nitrocellulose filters that contained anti-sense RNAs generated by SP6 or T7 RNA polymerase Headers indicate plasmid DNAs used in the transfections Left-hand margin indicates the mRNA identified by hybridization The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1 0

vectors when normalized to levels of hGH control RNA (the pattern of hybridization remained constant with a doubling of antisense RNA to the filters, data not shown). 293s cells transfected with hGH alone produced no transcripts that hybridize to the anti-sense CAT transcript. In addition, since the entire coding region of CAT represents the latter two-thirds of the transcript, it is unlikely that the rates of elongation of the two vectors were significantly different.

The AML sequence fails to increase CAT production when positioned upstream of the promoter. We examined the ability of the AML sequences present in the splice to act as a transcriptional activator when placed immediately upstream of the SV40 enhancer in vector pML.I-.CAT (named pML.I-.CAT/5′-IVS). The AML sequence present in pMLSIS.CAT was removed by two restriction enzymes whose sites immediately border the sequence. The overhanging ends were made blunt-ended with Klenow, and then subcloned into the SphI site just 5′ of SV40 enhancer in both the (+) and (−) orientations (relative to wild type adenovirus). Upon transfection into either 293s, HeLa, CV1, and CHO cells, the levels of CAT activity from pML.I-.CAT/5′-IVS (either orientation) were less than or equal to those from the parental vector, pML.I-.CAT (data not shown). These results were consistent with the nuclear run-on analysis

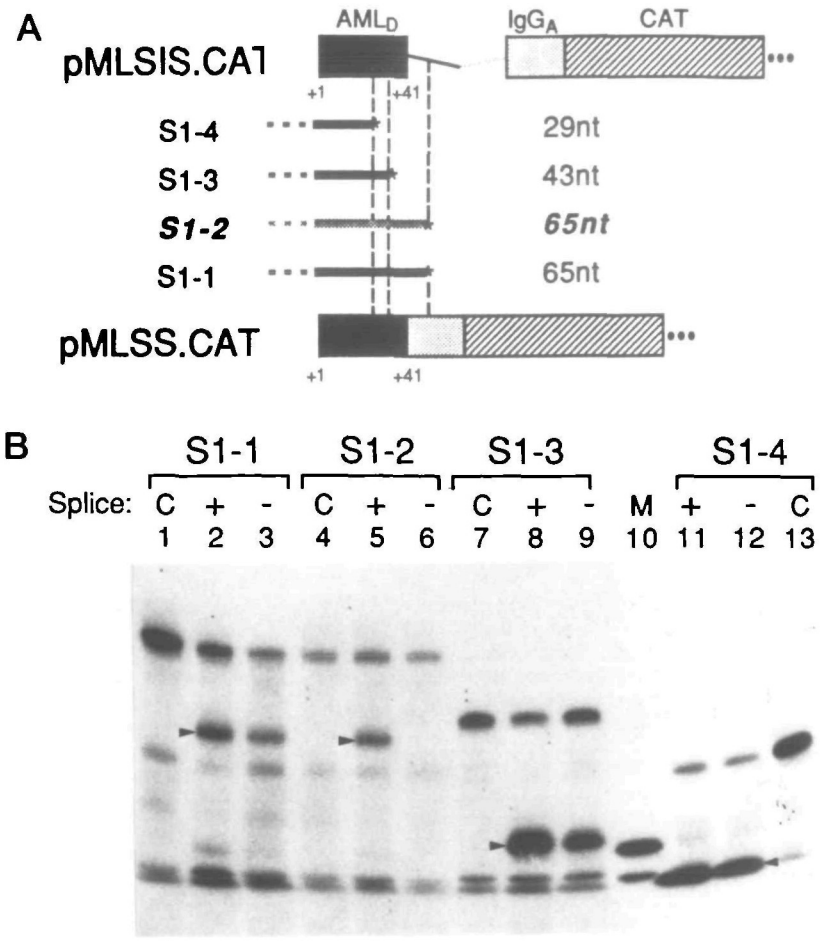

Figure 7A. Relative structure and hybridization pattern of single-stranded oligonucleotide S1 probes. Top and bottom figures represent CAT expression plasmids from Fig 1 In between lie four single-stranded oligonucleotide probes aligned in the anti-sense (3′-5′) orientation All probes contain an identical 3′-overhanging tail, but extend their homology to various regions of the RNA upstream or downstream of the 5′-splice junction The site of end-labeling is indicated by* The size of the protected fragment is shown to the right of each probe Probes S1−1, 3, and 4 are collinear to spliced RNA S1−2 contains 12 nucleotides homologous to the IVS and will only be protected from S1 nuclease digestion by RNA from pMLSIS CAT ('+' splice)

Figure 7B. S1 nuclease protection analysis of IVS(+) and IVS(−) nuclear RNA Individual S1 probes (top line, described in Fig 7A) were annealed to equal amounts of total nuclear RNA samples from transfected 293s cells· pMLSIS CAT ('+' splice), pML I- CAT ('−' splice), or mock-transfected controls (C) After treatment with S1 nuclease, the protected DNA fragments were electrophoresed through 10% acrylamide/8 M urea denaturing gels Arrows depict S1 protected bands of the expected size for each probe Lane M contains markers of 38 and 29 nucleotides in length The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 2 0.

and together indicate that the IVS does not contain sequences which activate transcription.

Steady-state nuclear RNA in 293s and HeLa cells. The nuclear run-on data were corroborated by performing an S1 nuclease analysis of the steady-state levels of nuclear RNA. We examined the 5′-leader region in the vicinity of the splice junction using the four synthetic oligonucleotide probes shown in Fig. 7A. All probes contain an identical 3′-overhanging tail but extend their homology to varying regions of the RNA either upstream or downstream of the IVS. Probe S1−1 hybridizes to the leader from +1 up through the first exon, and then 12 nucleotides (nt) to the 3′-side of the IVS. This probe lacks sequence homology to the IVS and will identify spliced mRNA. Probe S1−2 hybridizes to the same exon 1 sequences as S1−1, but continues across the 5′-splice junction, 12 nt into the IVS. This probe identifies only that mRNA which contains an IVS. Probes S1−3 and S1−4 hybridize to shorter regions of the leader prior to the IVS.

Nuclear RNA was isolated from 293s cells transfected with either pMLSIS.CAT or pML.I-.CAT and compared to RNA from untransfected cells using the above probes. In addition, primer extension reactions using a β-actin primer were performed to confirm the analysis of equivalent amounts of RNA (data not shown). Each of the four S1 probes yielded protected bands specific for transfected cells and having the predicted size (indicated by arrows in each lane, Fig. 7B). The band intensities indicate the presence of approximately 2-fold more of the IVS(+) versus IVS(−) mRNA regardless of whether the probe hybridized downstream (S1−1) or upstream (S1−3, 4) of the splice junction. In this particular experiment, the 2-fold difference in RNA concentration can be accounted for by the 2-fold greater transfection efficiency for pMLSIS.CAT (determined by hGH internal control). Probes S1−2 (lane 5) identifies unspliced mRNA from plasmid pMLSIS.CAT. HeLa cell nuclear RNA was also analyzed using the probe S1−1 (Fig. 8B, lanes 4−6). The CAT-specific bands in the early leader region of the IVS(+) and IVS(−) vectors were present in equivalent amounts. Due

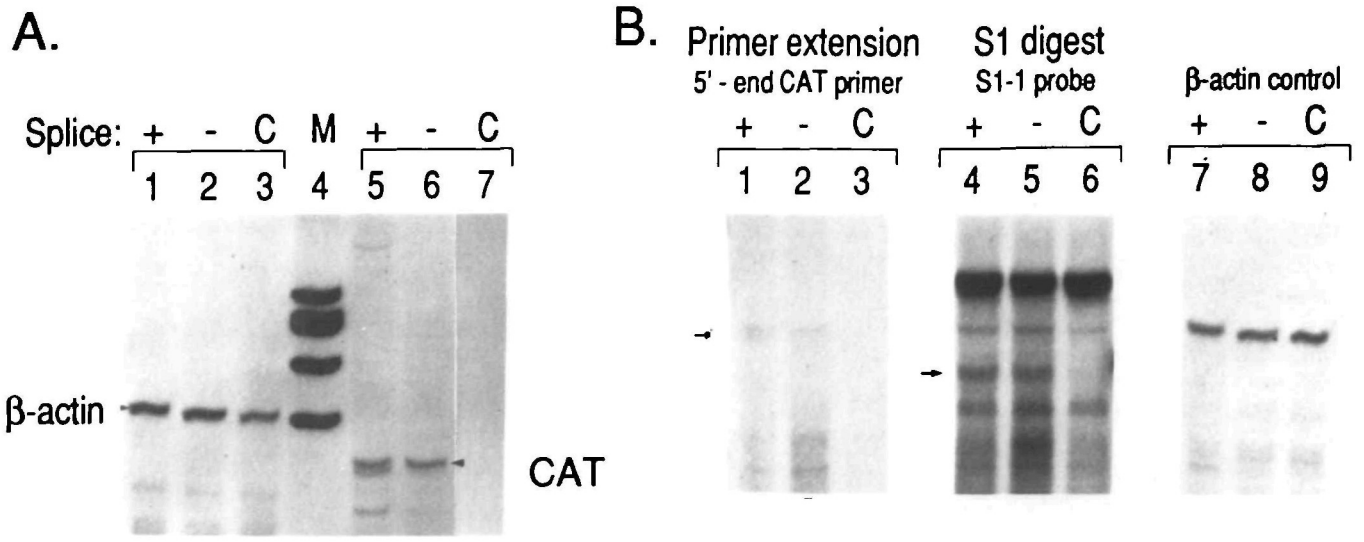

Figure 8. Analysis of total nuclear RNA samples from IVS(+) and IVS(−) vectors. (A) 293s cells Total nuclear RNA samples were examined for IVS(+) and IVS(−) CAT RNA by primer extension reactions Lanes 1−3 β-actin internal controls, and lanes 5−7 CAT-specific extension products RNA samples were from pMLSIS CAT ('+' splice), pML I- CAT ('−' splice), and mock-transfected cells, 'C' Lane M phi-X/*Hae*III DNA markers The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1 5 (B) HeLa cells Total nuclear RNA samples were prepared and analyzed by primer extension or S1 nuclease protection reactions Lanes 1−3 (CAT) and 7−9 (β-actin) are from primer extension reactions, legends are the same as for panel (A) Lanes 4−6 depict S1 nuclease reactions using probe S1−1 (see Fig. 7 for details) Arrows point to CAT-specific bands of the expected size for each primer or probe The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 0 8

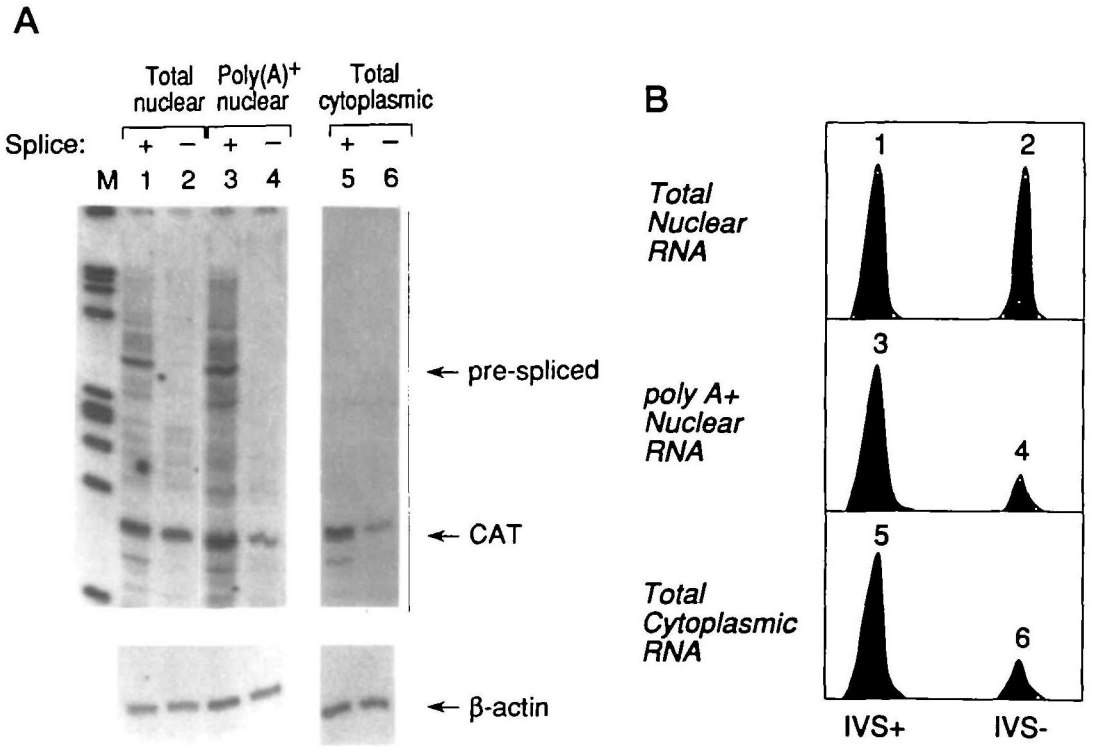

Figure 9. Primer extension analysis of IVS(+) vs IVS(−) RNA RNA from a single pool of transfected 293s cells was isolated as in Experimental Procedures for total nuclear and cytoplasmic fractions 24 hours after transfection Nuclear RNA· The relative amount of spliced vs. unspliced CAT RNA was determined for total RNA (lanes 1 and 2) using 10 percent of total nuclear RNA, each, for CAT and β-actin primer extension reactions The remaining RNA was selected for poly(A)+ by oligo-dT cellulose, which led to approximately 3% of the original amount of total RNA Of this poly(A)+ RNA, 33% was used for the control β-actin primer extension reaction, and 66% was used for the CAT primer extension reaction (lanes 3 and 4). Cytoplasmic RNA 20 μg (CAT primer) or μg (β-actin primer) of total cytoplasmic RNA was used for primer extension reactions for IVS(+) (lane 5) and IVS(−) (lane 6) vectors Lanes 5 and 6 derive from a separately prepared gel, electrophoresed for approximately the same time as the gel containing lanes 1−4, and M The relative amount of RNA used in each assay was determined from duplicate samples by primer extension with the β-actin primer (shown below in each lane). Lane M contains 5'-end-labeled phi-X174/*Hae*III DNA markers The ratio of hGH expression between the IVS(+) and IVS(−) vectors was 1.25 B Scanning densitometry analysis of CAT mRNA primer extension products CAT-specific bands marked by the arrow in panel A were quantitated by scanning densitometry. The ratio of IVS(+) to IVS(−) RNA was: 13 1 for total nuclear (lanes 1 and 2): 7.3 1 for poly (A)+ nuclear (lanes 3 and 4), and 6 8.1 for total cytoplasmic RNA

to the difference in transfection efficiencies, the signal to noise ratio was much lower in HeLa cells than in 293s cells, i.e., the background bands seen above and below the CAT specific protected fragment (Fig. 8B) were also seen in the HeLa control lane and in 293s cells (Fig. 7B, lanes 1 and 2) after longer exposure (data not shown). In addition, we believe that nuclear RNA preparations were not contaminated by cytoplasmic RNA since the presence of any significant amount would shift the ratio of IVS(+) to IVS(−) RNA away from the 1:1 ratios routinely observed, and towards the much higher ratios repeatedly seen in cytoplasmic analyses (see Fig.'s 3 and 4).

Since the S1 probes were of relatively short length, we determined the steady-state levels of longer nuclear transcripts by primer extension analysis (Fig. 8). Using the same 5′-CAT primers as in Fig. 3, we have determined that the relative amounts of total RNA from the 5′-coding region of the message were the same in both 293s (Fig. 8A) and HeLa cells (Fig. 8B). Any significant pausing of transcription within this region would have created a difference in steady-state RNA in this assay. This result, plus the nuclear run-on data, argue that the two vectors are identical with respect to the rates of initiation of transcription and elongation, and steady-state concentrations of total nuclear RNA.

Effect of splicing on the accumulation of poly(A)+ nuclear RNA. Fig. 9 compares the relative amounts of nuclear and cytoplasmic RNAs produced from IVS(+) and IVS(−) vectors in the same set of transfected 293s cells. Twenty-four hours after transfection, total nuclear and cytoplasmic RNA samples were harvested. Approximately 15% of each total nuclear RNA sample was set aside for primer extension analyses. The remaining RNA was selected for poly(A)+ and divided (2:1) for hybridization with either the CAT or β-actin primers, respectively. Total cytoplasmic RNA was analyzed as in Fig. 3. Results of the primer reactions are shown in Fig. 9A. lanes 1 and 2 (total nuclear RNA); lanes 3 and 4 (polyA+ nuclear RNA); lanes 5 and 6 (total cytoplasmic RNA). An arrow points to the reaction product corresponding to correctly initiated and spliced IVS(+) to IVS(−) RNA (which is equivalent to IVS-RNA). This band in each of the six lanes was analyzed by scanning densitometry as shown in Fig. 9B. By comparing the ratio of IVS(+) to IVS(−) RNA in each set of lanes, it is clear that: (i) both messages reach the same steady-state level in the nucleus prior to polyadenylation (lanes 1 and 2); (ii) the relative amount of polyadenylated IVS(+) increases to 7-fold while still in the nucleus (lanes 3 and 4); and (iii) the relative amount of steady state RNA in the cytoplasm (lanes 5 and 6) is virtually identical to that of polyadenylated nuclear RNA. The latter two points are in complete agreement with those described by Buchman and Berg (28), thus supporting the generality of this effect in two different systems. Thus there is a selective transition of (IVS+) mRNA from the total nuclear RNA pool into a steady-state population of poly(A)+ RNA. Most significant is the fact that even though splicing of SIS.CAT RNA is not required for polyadenylation to occur (Fig. 9, lane 3), its polyadenylated form exceeds that from pML.I-.CAT by 7-fold. These results imply that the ability of this message to be spliced, and not necessarily the process of splicing, is coupled to the accumulation of polyadenylated RNA.

DISCUSSION

Our experiments determined that RNA splicing can dramatically increase protein synthesis in mammalian cells. To characterize this phenomenon, we have compared RNA processing and protein synthesis from two CAT vectors which differ only in the presence or absence of a splice in their 5′-untranslated leader (pMLSIS.CAT and pML.I-.CAT, respectively). The splicing sequence consists of a 5′-exon and flanking IVS from the adenovirus tripartite leader fused to a 3′-splice sequence reported for an IgG variable region. Under transient transfection conditions, four mammalian cell lines expressed from between 6 to 50-fold more protein from the IVS(+) message, with the ratio of CAT activity directly proportional to the amount of steady-state cytoplasmic RNA. Although the bacterial gene encoding CAT was employed as a recorder gene in these studies, we emphasize that the effect of positioning a splice within the 5′-untranslated leader was not unique to this gene. Its presence has also led to significant and reproducibly higher levels of expression from eucaryotic genes such as Factor VIII (46), tissue plasminogen activator, and relaxin (latter two, unpublished results Gorman *et al.*).

We have considered four mechanisms that could lead to the differential CAT expression obtained above: Differences in (i) transcription initiation; (ii) transcription elongation, (iii) cytoplasmic RNA stability; and (iv) nuclear RNA processing/transport. Recent works have described a classical enhancer activity for the IVS of the mouse creatine kinase gene (44) as well as transcriptional activation by the IVS of a transgenic mouse gene (47). Although adenovirus sequences contained within the leader of PMLSIS.CAT have previously been shown to enhance transcription initiation from the AML promoter in the presence of virus in COS cells (45), our studies failed to reveal an effect on transcription due to the presence of the IVS. First, nuclear run-on analysis indicated that transcription initiation was the same for both the IVS(+) and IVS(−) vectors; since the bound probe consisted of a single-stranded RNA template containing the entire antisense CAT coding region, it is also unlikely that elongation was significantly affected. Second, the adenovirus sequence present within the splice has no enhancer capability when placed upstream of the SV40 promoter region in either orientation. Therefore, transcriptional effects did not account for the increased expression from pMLSIS.CAT.

Vectors either containing or lacking a splice in the 5′-untranslated leader produced equivalent amounts of steady-state total nuclear RNA. Most surprising, however, was the correlation between splicing and the relative amount of poly(A)$^+$ nuclear RNA. The vector pML.I-.CAT produced 7-fold less poly(A)$^+$ nuclear RNA even though precursor RNA from pMLSIS.CAT could be polyadenylated without first being spliced. Thus, even though the completion of splicing was not required for polyadenylation, there was a connection between the ability of the transcript to be spliced and the accumulation of poly(A)$^+$ RNA.

The effect of splicing characterized in this study is probably not specific for pMLSIS.CAT. To our knowledge, Villarreal and White (10) were the first to report that a splicing mutant (SV40) failed to produce correctly polyadenylated nuclear RNA. In that study, a deletion of the viral DNA removed the entire splice sequence for the 19s late RNA as well as the 3′ portion of the splice sequence for the 16s late RNA. Their most sensitive analysis of cytoplasmic RNA indicated that no 19s message was produced, and only a very minor percentage of 16s RNA appeared relative to wild type infection. Other investigators have demonstrated the ability to increase the expression of non-spliceable genes by the addition of a functional heterologous intron: the beta globin IVS (9, 15, 28), and the maize adh-1 IVS

(27). More specifically, our observation that an IVS(+) vector produced a greater amount of poly(A)+ nuclear RNA is identical to that of Buchman and Berg (28). They compared the effects of rabbit β-globin IVS1 or IVS2 on the production of RNA from recombinant SV40 virus vectors. As in our experiments, they found that the ratio of IVS(+) to IVS(−) poly(A)$^+$ RNA strongly favored the IVS(+) RNA in the nucleus, and that the same ratio developed in both the nucleus and cytoplasm of CV1 cells. Our results are thus consistent with many reports showing increased expression from IVS(+) vectors.

Several mechanisms could account for increased expression from IVS(+) vs. IVS(−) RNA: (1) transcription through the polyadenylation region differs between the two vectors; (2) RNA from both vectors is polyadenylated at the same rate, but IVS(+) RNA is more stable and/or transported more rapidly; or (3) IVS(+) RNA is polyadenylated more efficiently than IVS(−) RNA, but the rate of transport is the same once polyadenylation has occurred. Whereas we have not examined these possibilities directly, we feel our results support the third possibility. First, steady-state nuclear RNA levels and transcription initiation were shown to be the same for the two vectors Since the only difference between the two vectors is the presence of the IVS in pMLSIS.CAT, it seems unlikely that transcription through the polyadenylation sites would be dramatically different from that through the upstream region. Second, since the stability of both RNAs is the same in the cytoplasm, the determining event occurs prior to its appearance outside the nucleus. Third, not only can IVS(+) RNA be polyadenylated prior to splicing, but the presence of the IVS leads to a 7-fold accumulation of fully processed poly(A)+ RNA in the nucleus. Since this ratio precisely matches the ratio found in the cytoplasm, transport rates need not have been affected. Thus, we favor the hypothesis that the accumulation of polyadenylated IVS(+) RNA occurs by a post transcriptional mechanism, perhaps through some dual function of sn- or hnRNPs (48−52).

Mechanisms that allow for the transport of unspliced messages, such as the histone and interferon RNAs, may act in a highly selective manner. The processing of histone genes, which are non-polyadenylated, has been studied using sea urchin (see 41 for review), human and mouse extracts (53). A model has been proposed in which U7 RNA base-pairs to the 3′-terminal cleavage site and participates in the synthesis of appropriate 3′-ends. In other words, even though a message is unspliced and non-polyadenylated, it may be efficiently transported upon forming an appropriate 3′-terminus.

Other unspliced, but polyadenylated, genes may possess very strong poly(A) signals that allow them to enter the transport pathway without splicing. The existence of such a pathway would explain the results of Gross *et al.* (22) who showed the expression levels of the chicken TK gene to be independent of the number of native IVS present. Even more interesting is the activity of the nonspliced, HSP70 mRNA during adenovirus infection; HSP70 mRNA is actively transported along with the late family of adenovirus messages (54) even though most cellular messages are retained in the nucleus (55).

It is clear that the nuclear matrix can regulate the transport of several different forms of mRNA. It must be able to discriminate among RNAs based upon specific RNA processing signals and/or the association of particular ribonucleoprotein particles. Apparently splicing can regulate expression by the ability of intervening sequences to sequester multifunctional RNP complexes that interact to polyadenylate, splice, and transport pre-mRNA to

ACKNOWLEDGEMENTS

We would like to thank Robert Tjian, Rein Strijker, and Edward Jakobovits for supplying us with oligonucleotide primers; R. Strijker for valuable discussions during the course of this study; and William Wood and Robert DuBridge for critical reviews of the manuscript. We are also very grateful to Aya Jakobovits for technical advice on nuclear run-on conditions, and to the Clinical Assays Department at Genentech for performing hGH assays All figures were composed by Kerrie Andow, and the final manuscript prepared by Socorro Cuisia.

REFERENCES

1 Leff,S.E and Rosenfeld,M G. (1986) *Ann Rev Biochem.* **55**, 1091−1117
2 Breitbart,R E and Nadal-Girard,B (1987) *Cell* **49**, 793−803
3 Laski,F A , Rio,D C and Rubin,G M (1986) *Cell* **44**, 7−19
4 Boggs,R.T , Gregor,P , Idriss,S , Belote,J M and McKeown,M (1987) *Cell* **50**, 739−747
5 Sharp,P A (1987) *Science* **235**, 766−771
6 Berget,S M., Moore,C. and Sharp,P A (1977) *Proc. Natl Acad Sci USA* **74**, 3171−3175.
7 Chow,L T , Gelinas,R E , Broker,T R and Roberts,R J (1977) *Cell*, **12**, 1−8
8 Campos,R and Villarreal,L P (1982) *Virology* **119**, 1−11
9 Gruss,P., Lai,C -J , Dhar,R and Khoury,G (1979) *Proc Natl Acad. Sci USA* **76**, 4317−4321
10 Villarreal,L P and White,R T (1983) *Mol Cell Biol* **3**, 1381−1388
11 Villarreal,L P and Carr,S (1982) *Mol. Cell Biol.* **2**, 1550−1557
12 Khoury,G , Gruss,P , Dhar,R and Lai,C -J (1979) *Cell* **18**, 85−92
13 Lai,C -J and Khoury,G. (1979) *Proc Natl Acad Sci USA* **76**, 71−75
14 Gruss,P. and Khoury,G (1980) *Nature* **286**, 634−637
15 Hamer,D and Leder,P. (1979) *Cell* **18**, 1249−1302
16 Carlock,L and Jones,N.C (1981) *Nature* **294**, 572−574
17 Ghosh,P , Roy,P , Barkan,A , Mertz,J E , Weissman,S M and Lebowitz,P (1981) *Proc. Natl Acad Sci USA* **78**, 1386−1390
18 Mulligan,R C and Berg,P (1980) *Science* **209**, 1422−1427
19 Mulligan,R C , Howard,B H and Berg,P (1979) *Nature* **277**, 108−114
20 Gasser,C S , Simonsen,C C , Schilling,J W and Schimke,R T (1982) *Proc Natl Acad Sci. USA* **79**, 6522−6526
21 Treisman,R , Novak,U and Kamen,R (1981) *Nature* **292**, 595−600
22 Gross,M K , Kainz,M S and Merrill,G F (1987) *Mol. Cell Biol* **7**, 4576−4581
23 Drayna,D , Fielding,C , McLean,J , Baer,B , Castro,G , Chen,E , Comstock,L , Henzel,W., Kohr,W , Rhee,L , Won,K. and Lawn,R (1986) *J Biol. Chem* **261**, 16535−16539
24 Kedes,L.H (1979) *Ann Rev Biochem.* **48**, 837−880
25. Hunt,C and Morimoto,R I (1985) *Proc Natl Acad Sci USA* **82**, 6455−6459
26 Wu,B , Hunt,C and Morimoto,R (1985) *Mol. Cell Biol* **5**, 330−341
27. Callis,J , Fromm,M and Walbot,V. (1987) *Genes and Devel.* **1**, 1183−1200
28 Buchman,A.R. and Berg,P (1988) *Mol Cell Biol.* **8**, 4395−4405.
29 Gorman,C M , Moffat,L F and Howard,B (1982) *Mol. Cell Biol* **2**, 1044−1051
30 Graham,F L , Smiley,J., Russell,W C and Nairu,R (1977) *J Gen Virol* **36**, 59−77.
31 Reddy,V B , Ghosh,P K., Lebowitz,P., Piatak,M and Weissman,S M (1979) *J Virol.* **30**, 279−296.
32 Bothwell,A L.M , Paskind,M., Reth,M , Imanishi-Kari,T , Rajewsky,K and Baltimore,D (1981) *Cell* **24**, 625−637.
33 Gorman,C M , Gies,D., McCray,G and Huang,M. (1989) *Virology* **171**, 377−385.
34 Graham,F L and van der Eb,A.J. (1973) *Virology*, **52**, 456−467
35 Gorman,C , Padmanathan,R. and Howard,B. (1983) *Science* **221**, 551−553
36 Nordeen,S.K , Green,P P ,III and Folkes,M (1987) *DNA* **6**, 173−178
37 Mory,Y Y. and Gefter,M (1977) *Nucl Acids Res* **4**, 1739−1757
38 Maniatis,T , Fritsch,E and Sambrook,J (1982)?? Cold Spring Harbor Laboratory, Cold Spring Harbor,N.Y.
39 Lund,M., Gunderson,M.G and Groudine,? (1985) *Science* **230**, 1126−1132
40 Thomas,P. (1983) *Meth. Enzymol.* **100**, 255−266.
41 Banerji,J , Olson,L and Schaffner,W (1983) *Cell* **33**, 729−740
42 Birnstiel,M L., Busslinger,M. and Strub,K. (1985) *Cell* **41**, 349−359
43 Gillies,S.D., Morrison,S L , Oi,V T and Tonegawa,S (1983) *Cell* **33**, 717−728

44 Sternberg,E , Spizz,G , Perry,W M., Vizard,D , Weil,T. and Olsen,E N (1988) *Mol Cell. Biol.* **8**, 2896–2909
45. Mansour,S L , Grodzicker,T. and Tjian,R. (1986) *Mol. Cell. Biol.* **6**, 2684–2694.
46 Huang,M T F , Eaton,D and Gorman,C M (1987) In Miller,J.H and Calos,M P. CSHL (ed), Gene Transfer Vectors from Mammalian Cells, pp. 64–70
47 Brinster,R L., Allen,J M , Behringer,R R , Gelinas,R E and Palmiter,R.D (1988) *Proc Natl Acad. Sci. USA* **85**, 836–840
48 Gilmartin,G M , McDevitt,M.A. and Nevins,J R. (1988) *Genes and Devel.* **2**, 578–587
49 Moore,C L and Sharp,P.A. (1984) *Cell* **36**, 581–591
50 Moore,C L and Sharp,P A (1985) *Cell* **41**, 845–855
51 Moore,C L ,Chen,J. and Whoriskey,J (1988) *EMBO J* **7**, 3159–3169
52 Wilusz,J , Feig,D I and Shenk,T (1988) *Mol Cell. Biol* **8**, 4477–4483
53. Mowry,K L. and Steitz,J A. (1987) *Science* **238**, 1682–1687.
54 Kao,H -T and Nevins,J R. (1983) *Mol Cell Biol* **3**, 2058–2065
55 Babich,A , Feldman,L T , Nevins,J R , Darnell,J E and Weinberger,C. (1983) *Mol. Cell Biol.* **3**, 1212–1221

Note added in proof: pML I- CAT and pML.SS CAT are the same vector