MOLECULAR AND CELLULAR BIOLOGY, Nov. 1988, p. 4829–4839
0270-7306/88/114829-11$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 8, No. 11

Patterns of Polyadenylation Site Selection in Gene Constructs Containing Multiple Polyadenylation Signals

ROGER M. DENOME AND CHARLES N. COLE*

Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03756

Received 16 May 1988/Accepted 4 August 1988

We have constructed a series of plasmids containing multiple polyadenylation signals downstream of the herpes simplex virus type 1 (HSV) thymidine kinase (*tk*)-coding region. The signals used were from the simian virus 40 (SV40) late gene, the HSV *tk* gene, and an AATAAA-containing segment of the SV40 early region. This last fragment signals polyadenylation poorly in our constructs and not at all during SV40 infection. All plasmids contained the SV40 origin of replication. Plasmids were transfected into Cos-1 cells; after 48 h, cytoplasmic RNA was isolated and the quantity and 3′-end structure of *tk* mRNAs was analyzed by using S1 nuclease protection assays. In all constructs, all polyadenylation signals were used. Increasing the number of poly(A) signals 3′ to the *tk*-coding region did not affect the total amount of polyadenylated RNA produced, even with the weakest signal. Increasing the distance between two signals caused an increase in the use of the 5′ signal and a decrease in the use of the 3′ signal. Changing the distance between the 5′ cap and first signal did not affect signal use. Analyses of cytoplasmic mRNA stability, nuclear RNA distribution, and transcription in the polyadenylation signal region indicated that the distribution of *tk* RNAs ending at different poly(A) sites was the result of poly(A) signal choice, not other aspects of RNA metabolism. Four possible mechanisms of polyadenylation signal recognition are discussed.

The 3′ ends of most eucaryotic mRNAs consist of 50 to 200 adenylic acid residues [the poly(A) tail] added at a site of endonucleolytic cleavage in the nascent RNA. The signal sequences required for this processing have been defined in several systems. The sequence 5′-AAUAAA-3′ (or one very similar to it) is located 6 to 30 bases 5′ to the polyadenylation site in almost all eucaryotic mRNAs (32) and is required for processing (9, 41). Efficient poly(A) addition also requires a second sequence element located 3′ to the poly(A) site. This sequence is often U rich (14) or G+U rich (6, 7, 11, 24, 35, 36, 41) but seems to have no highly conserved sequence motifs (30). Polyadenylation can be carried out in vitro with HeLa cell nuclear extracts (15, 25, 26, 41), but the biochemistry of this processing reaction is not understood.

Although most studies of polyadenylation have used constructs with only a single poly(A) signal, many genes contain multiple signals. (For reviews see references 19 and 28). Different poly(A) signals in a given transcription unit are often used in a cell-type- or time-dependent manner. In some adenovirus and immunoglobulin heavy-chain genes, poly(A) site choice may determine which of several RNA splicing pathways is used; this determines the coding potential of the mRNA. In other cases (e.g., dihydrofolate reductase [37, 40]), RNAs produced using alternate poly(A) sites contain the same protein-coding region. The function of these different 3′-flanking regions is unknown.

Plasmids and viruses containing two poly(A) signals have been constructed (9, 12, 13, 18, 29, 31). In these studies, both polyadenylation signals were used under at least some conditions. When a signal was present twice, the promoter-proximal signal was usually used more. In most cases, poly(A) site choice was not the major emphasis of the study and other aspects of RNA metabolism (turnover, time of appearance, etc.) were not assayed.

* Corresponding author.

In this study we have addressed several unresolved questions about poly(A) site selection in the presence of multiple poly(A) signals. The simian virus 40 (SV40) late poly(A) signal (SVL), the herpes simplex virus (HSV) thymidine kinase poly(A) signal (TK), and an 88-base-pair (bp) AA TAAA-containing DNA fragment from the SV40 early region (88 signal) were inserted in single or multiple copies 3′ to the HSV *tk*-coding region. The quantity and 3′-end structure of RNAs produced from these constructs during transient expression in Cos-1 monkey kidney cells were assayed by S1 nuclease protection. Although all signals in each construct were used, increasing the number of poly(A) signals from one to four did not significantly affect the amount of RNA produced, regardless of how much RNA was produced from a construct with a single copy of the signal. Increasing the distance between two signals favored the use of the 5′ signal, while changing the distance between the promoter and the first signal had no effect on signal use. In concert with data from experiments assaying various aspects of RNA metabolism, these data are inconsistent with a number of models of poly(A) signal recognition. A model involving scanning from the point of transcription is consistent with our data.

MATERIALS AND METHODS

Plasmids, plasmid DNA, and bacteria. Bacterial growth and transfection (21), minilysate DNA preparation and analysis (4), and large-scale plasmid purification (10) were performed by using standard techniques. All plasmids were propagated in *Escherichia coli* HB101.

Plasmid construction. Plasmid pTK_1 (Fig. 1 and 2) is the same as pTK205R/SV010 described by Cole and Stacy (6). It was produced by BAL 31 nuclease resection of pTK2/SV010 (5) and contains the *tk*-coding region, both sequence elements of the polyadenylation signal, and 62 bp 3′ to the poly(A) site. pTK206R/SV010 (6, 41) contains the *tk*-coding

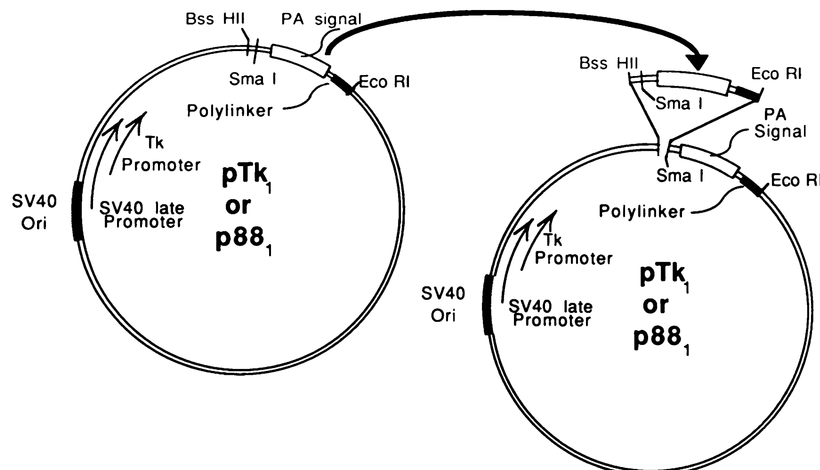

FIG. 1. Construction of plasmids containing multiple polyadenylation signals. Plasmids pTK_1 and $p88_1$ were digested with *Eco*RI, and the ends of the resulting fragments were filled by using Klenow fragment of DNA polymerase I in the presence of deoxynucleoside triphosphates. The DNA was then digested with *Bss*HII, and the fragments containing the polyadenylation signals were isolated from 2% agarose gels. These fragments were ligated to pTK_1 or $p88_1$ that had been digested with *Bss*HII and *Sma*I, producing pTK_2 and $p88_2$. This procedure was repeated, with $p88_2$ used as starting material to produce $p88_4$.

region plus 7 bp 3' to the termination codon. This is followed by a *Bam*HI linker, the polylinker from pSV010, pBR322-derived vector sequences, and the SV40 origin from pSV010 (27). The location, orientation, and number of inserts in the constructs described below were confirmed by restriction enzyme digestion of minilysate DNA. A number of the constructs described here required the ligation of two DNA fragments with termini that are not normally compatible. Before ligation, these termini were made blunt by using Klenow fragment of DNA polymerase I and deoxynucleoside triphosphates.

$pSVL_1$, $pSVL_2$, and $pSVL_4$. SV40 DNA was digested with *Bam*HI and *Bcl*I, and the 237-bp fragment containing the early and late poly(A) sites were isolated from a 2% agarose gel. *Bam*HI- and *Bcl*I-digested DNA fragments have mutually complementary termini. When *Bam*HI-generated ends are ligated to *Bcl*I-generated ends, the DNA is resistant to both enzymes. The 237-bp fragment was polymerized by ligation with T4 DNA ligase and then digested with *Bam*HI and *Bcl*I. The DNA was ligated with *Bam*HI-linearized pTK206R/SV010 to produce constructs containing one ($pSVL_1$), two ($pSVL_2$), and four copies ($pSVL_4$) of the SV40 late poly(A) signal.

$p88_1$. Plasmid pTK206/88/SV010 was described previously (5) and contains the 88-bp *Bst*NI fragment of SV40 (nucleotides 3247 to 3335) inserted with *Bam*HI linkers into the *Bam*HI site of pTK206/SV010. In pTK206/88/SV010, the HSV *tk* promoter and SV40 late promoter transcribe opposite strands of the plasmid. To construct a plasmid in which both promoters transcribe the *tk* gene and with this fragment as a poly(A) signal, the 188-bp *Bss*HII-*Eco*RI fragment of pTK206/88/SV010 (containing the 88-bp fragment, *Bam*HI linkers, and flanking HSV sequences) was isolated from a 2% agarose gel and inserted into *Bss*HII-*Sma*I-cut pTK206R/SV010 to create $p88_1$.

pTK_2, $p88_2$, and $p88_4$. pTK_1 was digested with *Eco*RI, and the ends of the DNA were made blunt. The DNA was then digested with *Bss*HII, and the 206-bp fragment containing the poly(A) signal was isolated from a 2% agarose gel and inserted into *Bss*HII-*Sma*I-cut pTK_1 (Fig. 1) to create pTK_2.

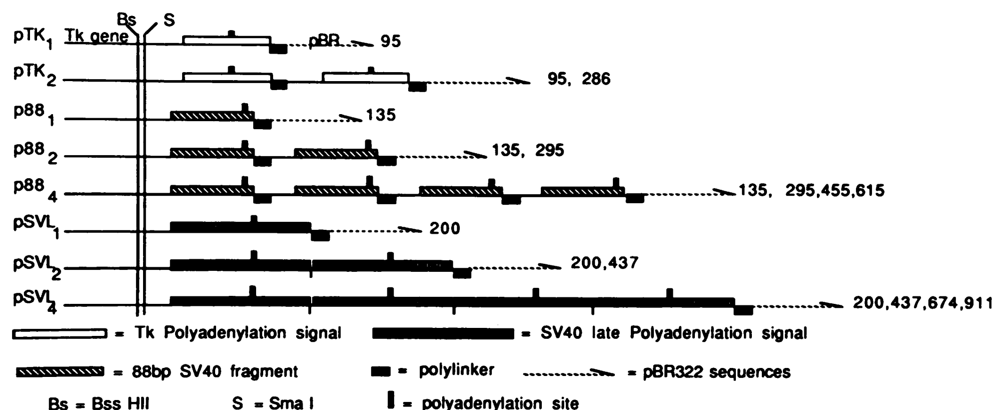

FIG. 2. Diagram of the polyadenylation signal region of constructs containing multiple polyadenylation signals. All constructs have the SV40 origin of replication, pBR322 sequences, and HSV *tk*-coding region as shown in Fig. 1 and were constructed as described in the legend to Fig. 1 and the text. The distance of the predicted polyadenylation sites from the *Bss*HII site (Bs), which was used as the labeling site for most S1 protection assays, is given after each construct. (For constructs containing the TK or 88 signals, the distances given are to the 5'-most polyadenylation site.)

The same procedure, starting with $p88_1$ DNA, was used to make $p88_2$. $p88_4$ was constructed by inserting the *Bss*HII-*Eco*RI fragment of $p88_2$ [containing two 88 poly(A) signals] into *Bss*HII-*Sma*I-cut $p88_2$.

Insertion and deletion constructs. Separate aliquots of $pSVL_1$ were digested with *Aat*II, *Pst*I, or *Bgl*I, and the termini of the DNA fragments were made blunt as described above. The samples were digested with *Cla*I, which cuts in the polylinker region just downstream of the SVL signal. The DNA fragments containing the polyadenylation signal were inserted downstream of the SVL signal in $pSVL_1$ (between the *Xba*I and *Cla*I sites in the polylinker) to produce constructs containing two SVL signals 437 bp (*Aat*II-*Cla*I inserts), 628 bp (*Pst*I-*Cla*I inserts), or 827 bp (*Bgl*I-*Cla*I inserts) apart ($pSVL_{2+437}$, $pSVL_{2+628}$, and $pSVL_{2+827}$, respectively).

Deletions in the *tk*-coding region were made by digesting $pSVL_{2+437}$ with *Mlu*I, *Eco*RV, or *Pst*I. After digestion, DNA was ligated and transfected into *E. coli*, and miniprep DNA was screened for the loss of the 150-bp *Mlu*I fragment ($^{-150}pSVL_{2+437}$), the 104-bp *Eco*RV fragment ($^{-104}pSVL_{2+437}$), and the 840-bp *Pst*I fragment ($^{-840}pSVL_{2+437}$). $^{-83}pSVL_{2+437}$ and $^{-390}pSVL_{2+437}$ were created by ligating the 1,067-bp *Mlu*I-*Bss*HII or 760-bp *Eco*RV-*Bss*HII fragment of $pSVL_1$, respectively, to the 3,174-bp *Bss*HII-*Bgl*II fragment of $pSVL_{2+437}$.

Cos-1 cell transfection, RNA preparation, and S1 nuclease protection assays. Cell culture and transfection conditions for Cos-1 monkey cells were described previously (5). Plasmid DNAs were transfected into Cos-1 cells by using DEAE-dextran–chloroquine phosphate treatment as described previously (20). At 48 h after transfection, cytoplasmic RNA was isolated as described by White et al. (39). S1 nuclease protection analysis was performed as described elsewhere (3, 8). S1-digested samples were subjected to electrophoresis on denaturing polyacrylamide–50% urea gels in TBE buffer (89 mM Tris hydroxide, 89 mM boric acid, 2.5 mM EDTA). Gels were dried and exposed to Kodak XAR-5 film at −80°C (without intensifying screen) for 4 h to 20 days. S1 nuclease protection assays were performed in DNA excess; RNA from mock-transfected cells gave no signal (data not shown). Transcripts from the opposite strand of these plasmids are undetectable, making strand-specific probes unnecessary (41).

Southern blots and nick translation. Plasmid DNA was isolated from Cos-1 cells by the method of Hirt (16). DNA was digested with *Eco*RI, subjected to electrophoresis on 1.2% agarose gels, and transferred to nitrocellulose by the method of Southern (38). Nick-translated pTK206R/SV010 was prepared by the method of Maniatis et al. (22).

Transcription in isolated nuclei. The distribution of transcription complexes on $pSVL_4$ DNA was assayed by the procedure of Babich et al. (2). Nuclei isolated from Cos-1 cells 48 h after $pSVL_4$ transfection were incubated in a solution containing 25% glycerol, 5 mM $MgCl_2$, 12 mM β-mercaptoethanol, 0.02 mM EDTA, 25 mM Tris chloride (pH 8.0), 1 mM ATP, 1 mM GTP, 1 mM CTP, and 500 μCi of [α-^{32}P]UTP (800 Ci/mM) per ml for 15 min at room temperature. The reaction mixture was brought to 1% sodium dodecyl sulfate and 0.5 M NaCl and repeatedly phenol extracted. RNA was precipitated with the addition of 2.5 volumes of cold ethanol. $pSVL_4$ was digested with the appropriate restriction endonucleases, and DNA fragments to be used as probes were isolated after electrophoresis in 8% polyacrylamide gels in TBE buffer. DNA fragments were immobilized on Gene Screen Plus (Du Pont Co., Boston, Mass.) with a slot blot apparatus as described by the manufacturer (Schleicher & Schuell, Inc., Keene, N.H.). ^{32}P-labeled RNA (0.25×10^6 to 1.0×10^6 cpm) was hybridized with the filter for 48 to 72 h. The blots were washed at a final stringency of 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 60°C and autoradiographed as described above.

RESULTS

The plasmids used here contain multiple polyadenylation signals downstream from (3′ to) the HSV *tk*-coding region (Fig. 1 and 2). The SV40 replication origin located upstream of (5′ to) the *tk* promoter and coding region contains the SV40 late promoter; this promoter and the *tk* promoter direct transcription of the *tk* gene. Downstream of the coding region and polyadenylation signal(s) are the polylinker from pSV010 and the pBR322 sequences containing the ampicillin resistance gene.

Three polyadenylation signals were used in these constructs. Plasmids pTK_1 and pTK_2 contained one and two copies, respectively, of the HSV *tk* poly(A) signal. $pSVL_1$, $pSVL_2$ and $pSVL_4$ contained one, two, and four copies, respectively, of the 237-bp *Mbo*I DNA fragment containing the SV40 late signal. $p88_1$, $p88_2$, and $p88_4$ contained one, two, and four copies, respectively, of a 188-bp DNA fragment containing the 88-bp *Bst*NI fragment of SV40, plus 5′-flanking HSV and 3′-flanking polylinker sequences. This last signal contains the sequence 5′-AATAAA-3′ in the *tk* sense orientation. Although this sequence does not signal polyadenylation during SV40 infection, it does act as a weak poly(A) signal in conjunction with the *tk*-coding region (5). The majority of *tk* mRNA produced from this construct is polyadenylated just 3′ to the AAUAAA encoded by this fragment. However, some RNA is processed by using a cryptic site in pBR322-derived sequences (41).

Cytoplasmic RNA was isolated from Cos-1 cells 48 h after transfection with $pSVL_1$, $pSVL_2$, $pSVL_4$, pTK_1, pTK_2, $p88_1$, $p88_2$, or $p88_4$. The 3′ ends of these constructs are shown diagrammatically in Fig. 2; the sizes of the expected S1 nuclease-resistant fragments are listed after each construct. S1 nuclease protection assays with homologous probes (labeled at the *Bss*HII site just upstream of the polyadenylation signal region) were used to determine the relative amount of RNA ending at a given poly(A) signal within a construct (Fig. 3). All the poly(A) sites were used, but there were large differences (up to 20-fold) in the cytoplasmic concentrations of RNAs ending at different poly(A) signals within a given construct. Average data from densitometric scans of autoradiographs from four to seven independent experiments using these constructs are presented in Table 1. In general, more RNA ended at promoter-proximal sites. The only exceptions were with $p88_2$ and $p88_4$, in which cases RNAs ending at the last site in the construct was always more abundant than that ending at the next signal upstream.

The relative amount of *tk* RNA produced from different plasmids was measured by S1 nuclease protection assays using one 3′-end-labeled DNA probe with cytoplasmic RNA from all transfections. This probe (*Bss*HII-cut pTK206R/SV010) is identical to the plasmids used here for only 44 bp (pTK_1 and pTK_2) or 49 bp (all others) 3′ of the *Bss*HII site. All *tk* RNAs protected 44 or 49 bases of this probe in S1 nuclease protection assays, regardless of which polyadenylation site was used. The intensity of the single autoradiographic band was used to measure the amount of *tk* RNA

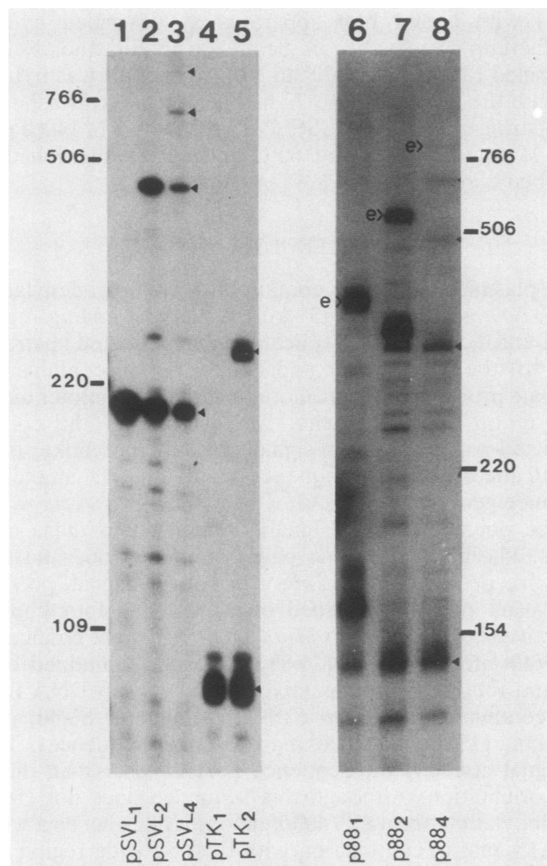

FIG. 3. Analysis of 3′ ends of RNAs produced from plasmids containing multiple polyadenylation signals. Cos-1 cells were transfected with pTK_1, pTK_2, $pSVL_1$, $pSVL_2$, $pSVL_4$, $p88_1$, $p88_2$, and $p88_4$ as described in the text. Cytoplasmic RNA was isolated 48 h later and analyzed with S1 nuclease protection assays by using homologous probes labeled at the *Bss*HII site. After S1 nuclease digestion, samples were run on 5% polyacrylamide–50% urea gels, which were dried and autoradiographed at −80°C. The plasmid used for transfection is listed below each lane, and the arrowheads mark the predicted distance to the polyadenylation site(s). For constructs containing the TK or 88 signals, these distances are to the most promoter-proximal processing site for each signal. The symbol e> indicates the position of RNA polyadenylated in pBR322 sequences, as described in the text. Numbers at the right and left indicate the position of DNA size markers (in bases).

(Fig. 4). Average densitometric data from four to seven independent experiments is shown in Table 2. Cos-1 cells transfected with $pSVL_2$ contained very little more *tk* RNA than cells transfected with $pSVL_1$; $pSVL_4$-transfected cells had slightly less *tk* RNA. The cytoplasmic *tk* RNA concentration in cells transfected with pTK_1 or pTK_2 was about 70% that in $pSVL_1$-transfected cells, but there was little difference between transfections using pTK_1 and those using pTK_2. Similarly, $p88_1$-, $p88_2$-, and $p88_4$-transfected cells had about the same amount of *tk* RNA. This level was about 20% that found in $pSVL_1$-transfected cells. Although there was some variation from experiment to experiment, the general pattern of RNA quantity was very consistent. Increasing the number of signals had little effect on RNA quantity, regardless of signal strength.

The exact site of polyadenylation in RNAs containing an 88 signal seemed to depend on the sequences 3′ to that signal. The poly(A) site in $p88_1$ is displaced ~10 bases 3′ to

TABLE 1. RNA processing at each polyadenylation signal in constructs containing multiple polyadenylation signal[a]

Plasmid	Proportion of total RNA ending at polyadenylation site[b]:			
	1	2	3	4
pTK_1	1.0			
pTK_2	0.66	0.33		
$p88_1$	1.0			
$p88_2$	0.42	0.58		
$p88_4$	0.46	0.34	0.09	0.12
$pSUL_1$	1.0			
$pSUL_2$	0.61	0.39		
$pSUL_4$	0.68	0.21	0.08	0.03

[a] Determined by densitometric analysis of homologous and nonhomologous S1 protection assays (see text and legends to Fig. 3 and 4).

[b] Relative to total RNA produced by the construct. The 5′-most site is site 1. Data for constructs containing the "88" signal do not include RNA ending in pBR322-derived sequences. For constructs in which processing occurs at several closely spaced sites in the signal region, data represent the sum of the processed species produced in that region.

the first site in $p88_2$ or $p88_4$, even though the RNAs up to this first point are identical. Similarly, the second site used in $p88_2$ is displaced relative to the second site in $p88_4$. The processing site for the 88 signal is close to the end of the DNA fragment that was duplicated in these constructs. Therefore, the last signal in each of these plasmids is very close to the flanking polylinker- and pBR322-derived sequences. Others have reported that sequence changes in the vicinity of the polyadenylation site can affect the exact positioning of processing (14, 41). We believe this to be the case here also. We also saw some RNA ending in an area corresponding to vector sequences in constructs containing the 88 signal. This RNA is polyadenylated, but we have not defined this signal further.

Do different *tk* mRNAs have different stabilities? Although all the signals in a construct were used, different RNAs produced from a construct were not present in equal con-

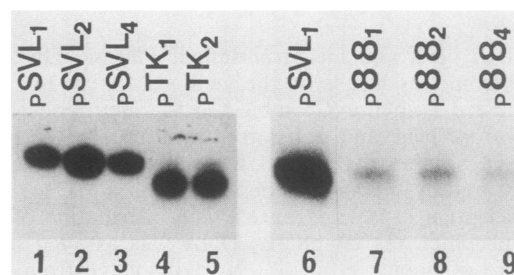

FIG. 4. Analysis of the amount of RNA produced from constructs containing multiple polyadenylation sites. Cos-1 cells were transfected and cytoplasmic RNA was isolated as described in the legend to Fig. 3 and the text. *Bss*HII-cut pTK206R/SV010 was labeled as described in the text and used as a probe for RNAs isolated from all transfections. RNAs, electrophoresis, and autoradiography conditions were as described in the legend to Fig. 3. RNA from plasmids containing the SVL or 88 polyadenylation signal protects a 49-base fragment of this probe, while RNA from plasmids containing the TK signal protects a 44-base fragment. The plasmids used for transfection are listed above the lanes. Lanes 1 through 5 are from a single experiment. Lanes 6 through 9 are from a single experiment, with lane 6 containing the $pSVL_1$ RNA from this experiment for comparison.

TABLE 2. Relative amounts of RNA produced from plasmids containing multiple polyadenylation signals[a]

Plasmid	Amt of RNA produced relative to[b]:	
	$pSVL_1$	Monomer
pTK_1	0.73	1.0
pTK_2	0.63	0.86
$p88_1$	0.21	1.0
$p88_2$	0.19	0.90
$p88_4$	0.22	1.04
$pSVL_1$	1.0	1.0
$pSVL_2$	1.18	1.18
$pSVL_4$	0.76	0.76

[a] Determined by densitometric analysis of homologous and nonhomologous S1 protection assays (see text and legends to Fig. 3 and 4).

[b] Amount of RNA produced by the given construct relative to the amount produced by $pSVL_1$ or the amount produced by the construct containing only one copy of the signal under study (monomer).

centrations. This could reflect differences in polyadenylation site use. An alternative explanation is that all poly(A) signals were used equally, but larger RNAs were much less stable than shorter species. To test this, cells were transfected with $pSVL_4$, pTK_2, or $p88_4$, and, 48 h later, the medium was replaced with medium containing actinomycin D (5 μg/ml) to inhibit transcription. Cytoplasmic RNA was isolated at the time of actinomycin treatment and 2, 5, and 10 h later. The RNA was analyzed by using S1 nuclease protection with homologous probes as described above (Fig. 5). The amount of each RNA species present at each time was determined densitometrically. Table 3 shows the amount remaining at 5 h. There is very little difference in the cytoplasmic stability of any of the RNA species produced at any of the signals used. These small differences cannot explain the >20-fold differences in RNA concentration that were seen (Fig. 3 and Table 1).

Are different polyadenylation signals used at different times? Are shorter RNA species exported from the nucleus with greater efficiency? Large differences in cytoplasmic concentration of different RNAs might be caused by different efficiencies of nucleocytoplasmic RNA transport. There might also be some temporal bias in site usage, dependent on the length of time after transfection, the amount of template present, or some other factor. To test these possibilities, Cos-1 cells were transfected with $pSVL_4$, and cytoplasmic RNA was isolated 5, 24, 48, and 72 h later. Nuclear RNA was isolated at 48 h. The quantity (Fig. 6A) and 3′ end structure (Fig. 6B) of *tk* RNA in these samples were assayed. On the basis of assays of *tk* RNA quantity, the same amount of *tk* RNA from each time point was added to homologous probe S1 protection assays. Autoradiographs were analyzed densitometrically (Table 4).

We could find no *tk* mRNA in the sample isolated at 5 h (Fig. 6A, lane 1). The 24- and 72-h RNAs contained 40 and 70%, respectively, as much *tk* RNA as the 48-h RNA (Table 4). The patterns of poly(A) site use at 24, 48, and 72 h (Fig. 6B, lanes 2 through 4, and Table 4) were almost identical, indicating that there was no bias in signal use over time. Nuclear RNA isolated at 48 h looked very much like cytoplasmic RNA, indicating that all RNAs were transported equally well. These data confirm the results of stability experiments described above. If different-length RNAs had different cytoplasmic stabilities, then the nuclear and cytoplasmic *tk* RNA distributions would be different.

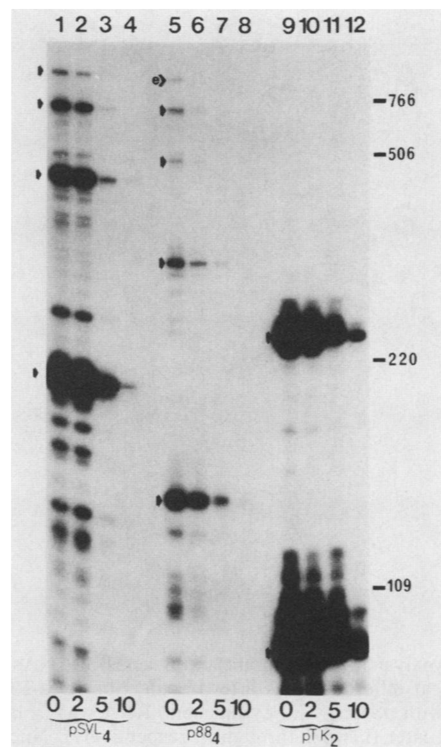

FIG. 5. Analysis of stability of RNAs ending at different polyadenylation sites. Cos-1 cells were transfected with pTK_2, $pSVL_4$, and $p88_4$; 48 h later, growth medium was replaced with medium containing 5 μg of actinomycin D per ml. Cytoplasmic RNA was isolated 0, 2, 5, and 10 h later and analyzed by S1 nuclease protection assays by using probes labeled at the *Bss*HII site. S1 nuclease-digested samples were electrophoresed as described in the legend to Fig. 3. All lanes represent the analysis of 10 μg of cytoplasmic RNA. Plasmids used in transfection and as probes and the time (in hours) after the addition of actinomycin D are given at the bottom. The predicted sizes for RNAs produced from these plasmids are indicated by arrows. The symbol e> indicates the extended transcripts produced from $p88_4$.

Most of the plasmids used here contain duplications. Recombination within or between plasmids during the transfection period could produce plasmids with more or fewer polyadenylation signals than the original plasmid. Plasmid DNAs were isolated from Cos-1 cells 48 h after transfection, and the structure of the plasmids was analyzed by restriction digestion followed by Southern blot analysis (data not

TABLE 3. Stability of RNAs ending at different polyadenylation sites[a]

Plasmid	Amt of RNA remaining after 5 h in medium containing actinomycin D			
	1[b]	2	3	4
$pSVL_4$	0.24	0.16	0.16	0.25
pTK_2	0.20	0.20		
$p88_4$	0.29	0.20	0.24	0.14

[a] White light densitometry was used to analyze the autoradiograph shown in Fig. 5. Multiple exposures were required to ensure that the data were in the linear response range of the X-ray film.

[b] The numerals 1 through 4 refer to *tk* RNAs ending at the first, second, third, and fourth polyadenylation sites, respectively, in the constructs indicated. All values are relative to the amount of each RNA species present at 0 h.

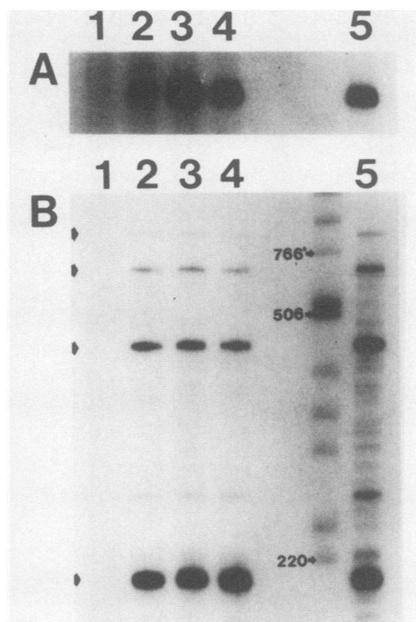

FIG. 6. Analysis of the amount and 3' ends of RNAs produced from $pSVL_4$ at different times after transfection. Cos-1 cells were transfected with $pSVL_4$, and cytoplasmic RNA was isolated 5, 24, 48, and 72 h later (lanes 1 through 4, respectively). Nuclear RNA was isolated at 48 h (lane 5). A 10-μg sample of each RNA was analyzed by S1 nuclease protection by using *Bss*HII-cut pTK206R/SV010 as a probe as described in the legend to Fig. 4. (A) No *tk* RNA was detected at 5 h. (B) S1 nuclease protection analysis of these RNAs using *Bss*HII-cut $pSVL_4$ as probe (as described in the legend to Fig. 3); the amount of RNA used was that required to give signal strength equivalent to the 48-h cytoplasmic RNA. The positions of the polyadenylated species are indicated by the arrows at left. The sizes (in bases) of marker DNAs are indicated.

shown). The structure of the plasmids did not change during the transfection period. Furthermore, the concentration of different plasmids was approximately equal in all transfections (as assayed by the intensity of bands common to all plasmids), indicating that all constructs replicated to the same extent.

Is transcription continuing past the polyadenylation signal region? Unequal use of poly(A) sites could be caused by transcription attenuation in the polyadenylation signal region. To assay this, nuclei were isolated from Cos-1 cells 48 h after transfection with $pSVL_4$. Run-on transcription was performed with [α-^{32}P]UTP. Controls indicated that transcription reinitiation was not occurring (data not shown). RNA was hybridized to filter-bound DNA fragments (Fig. 7A) from the SV40 late promoter region (a), *tk* promoter and 5' noncoding region (b), fragments from within the *tk*-coding region (c and d), and a fragment from sequences 600 bp downstream of the last poly(A) signal (e). Filters were washed and autoradiographed (Fig. 7B). The same amount of ^{32}P-labeled RNA hybridized to each probe, indicating that transcription was equimolar across the region containing polyadenylation signals.

TABLE 4. RNA production by $pSVL_4$ at different times after transfection[a]

RNA type and isolation time (h)	Proportion of total *tk* RNA ending at polyadenylation site[b]:				Relative RNA conc[c]
	1	2	3	4	
Cytoplasmic					
24	0.65	0.24	0.08	0.02	0.39
48	0.67	0.23	0.08	0.02	1.0
72	0.69	0.21	0.09	0.01	0.67
Nuclear					
48	0.57	0.27	0.12	0.04	0.86

[a] Densitometric data from the autoradiographs shown in Fig. 6.
[b] The 5'-most site is site 1.
[c] Cytoplasmic *tk* RNA concentration at 48 h is 1.0.

Does the distance between poly(A) sites influence site use? More distal signals may be used less because they are farther from the promoter or because they are downstream of other signals. If the latter is true, is the distance between signals important? A second SVL signal and various amounts of *tk*-coding region were placed 3' to the one poly(A) signal in $pSVL_1$, creating plasmids with signals 237, 437, 628, and 827 bp apart ($pSVL_2$, $pSVL_{2+437}$, $pSVL_{2+628}$, and $pSVL_{2+827}$, respectively; Fig. 8A and C). Plasmids were transfected into Cos-1 cells, and the *tk* RNAs were analyzed as described above. All these plasmids produced the same amount of RNA as $pSVL_1$ did (data not shown). The first SVL signal was used more and the second SVL signal was used less with increasing distance between the two poly(A) sites (Fig. 8A). After very long exposures, a very small amount of RNA was detected ending at the second signal in $pSVL_{2+827}$ (data not shown).

In a second set of constructs, the distance between the promoter and the first polyadenylation signal of $pSVL_{2+437}$ was decreased by removing 83, 104, 150, 390, or 840 bp from the *tk*-coding region ($^{-83}pSVL_{2+437}$, $^{-104}pSVL_{2+437}$, $^{-150}pSVL_{2+437}$, $^{-390}pSVL_{2+437}$, $^{-840}pSVL_{2+437}$, respectively; Fig. 8B and C). Decreasing the distance between the promoter and the first SVL signal (Fig. 8B) had little effect on the pattern of signal use. Densitometric analysis indicated that the more distal signal in RNAs analyzed in Fig. 8B was used in 5 to 6% of the *tk* RNAs, regardless of the length of the RNA upstream of the first signal.

Taken together, the data described here suggest that the relative use of each polyadenylation signal, not some other aspect of RNA metabolism, determined the relative abundance of different RNA species produced by our constructs. Furthermore, the distance between polyadenylation signals in a construct had a major effect on how much each was used. The length of RNA upstream of the first signal had little affect on poly(A) site use.

DISCUSSION

These studies describe patterns of *tk* mRNA 3'-end processing when the HSV *tk* gene contains multiple polyadenylation signals. The data can be summarized as follows. (i) In all constructs, all polyadenylation signals were used (Fig. 3). (ii) Increasing the number of copies of any signal did not cause an increase in the amount of RNA produced (Fig. 4 and Table 2). Note that the *tk* gene followed by a single SVL signal produced five times as much RNA as the *tk* gene in conjunction with the 88 signal (Fig. 4, lanes 6 through 9). (iii) In general, shorter RNA species were more abundant than longer RNAs made from the same construct. However, in $p88_2$ and $p88_4$, the last 88 signal was used more than the next to the last signal (Fig. 3 and Table 1). (iv) Changing the distance between the promoter and two SVL signals did not affect signal usage (Fig. 8B). Increasing the distance between the two signals increased use of the 5' signal and decreased use of the 3' signal (Fig. 8A).

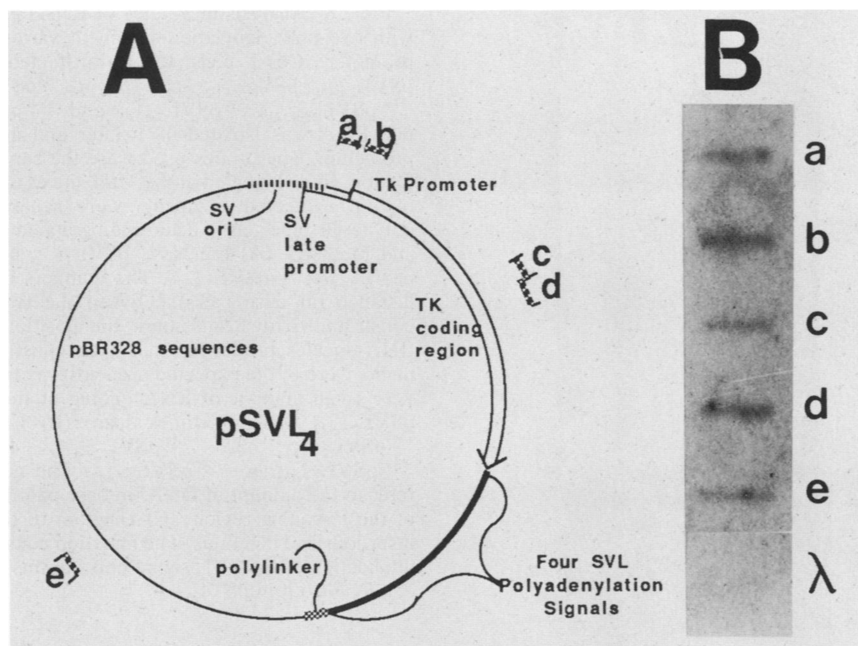

FIG. 7. Analysis of transcription rate in $pSVL_4$. Cos-1 cells were transfected with $pSVL_4$; 48 h later, nuclei was isolated. After extensive washing, the nuclei were allowed to elongate preinitiated transcripts in the presence of [α-^{32}P]UTP and unlabeled ATP, GTP, and CTP. Labeled nuclear RNA was isolated as described in the text and hybridized to a slot blot containing 200 to 250 ng of each of the DNAs indicated in panel A (a through e) and lambda DNA. (B) Hybridization was for 48 h at 42°C in 50% formamide–5× SSC. The blot was washed to a final stringency of 0.1× SSC at 60°C. Autoradiography was for 4 days. The DNAs used as probes were as follows: a, 138-bp *Eco*RI fragment from between the SV40 late promoter and the *tk* promoter; b, 150-bp *Mlu*I fragment from just downstream of the *tk* promoter; c, 147-bp *Bgl*I fragment from within the *tk*-coding region; d, 190-bp *Bgl*I-*Pst*I fragment from within the *tk*-coding region; e, 134-bp *Pst*I-*Bgl*I fragment approximately 600 bp downstream of the last polyadenylation signal in $pSVL_4$; λ, lambda DNA.

The various *tk* RNAs produced in these experiments might differ in aspects of RNA metabolism other than polyadenylation site choice. These metabolic differences could obscure the effects of differential poly(A) site use. Experiments described here indicated that RNAs produced from the same plasmid did not differ significantly in their cytoplasmic stabilities (Fig. 5 and Table 3), rates of transport from the nucleus, or times of synthesis (Fig. 6 and Table 4). Nuclear runoff experiments showed that transcription continued unimpeded past the last polyadenylation signal in $pSVL_4$ (Fig. 7).

Adding a signal to a construct might affect the nuclear stability of the RNAs produced from that plasmid. Although we cannot measure nuclear RNA stability, we believe that our data indicate that differences in RNA abundance that we see were due to differences in synthesis, not nuclear stability. The shortest RNAs produced from $pSVL_1$, $pSVL_2$, and $pSVL_4$ have exactly the same sequence, regardless of the template; this RNA was 50% more abundant when produced from the $pSVL_1$ template than when produced from $pSVL_2$ or $pSVL_4$ (Table 1). Since this RNA should be metabolized at the same rate regardless of its origin, this difference in abundance must reflect a difference in synthetic rate. The abundance of the shortest RNAs produced from constructs containing the TK or 88 signals is similarly affected by the presence of downstream signals.

Our experiments assayed cytoplasmic RNA stability by stopping transcription with actinomycin D. The data indicated that there were no large differences in the stabilities of the different *tk* mRNAs under these conditions. Although treatment with actinomycin D may affect the stability of an RNA, we believe that stability differences of the magnitude required to explain the differences in RNA levels observed in a given transfection would be apparent in our analysis. We saw no differences in stability.

Increasing the number of poly(A) signals did not affect the amount of RNA produced. The RNA was just processed at more sites. When a new signal was added to the 3′ end of a construct, less RNA was processed at the signal immediately 5′ to the new signal; signals further upstream were not affected. The data presented were derived by using constructs containing one, two, or four copies of a signal. We have extended these studies to include $pSVL_3$ (containing three SVL signals), and constructs containing one, two, and three copies of the SV40 early signal (R. M. Denome and C. N. Cole, unpublished results). In all of these cases, a new downstream signal only decreased the use of the signal immediately upstream of it.

Models. Our data can be used to test a number of possible mechanisms of polyadenylation site selection. The actual mechanism of site selection (and any good model) must account for the observations listed above. In addition, any model should take into account the observation that different polyadenylation signals cause the production of different amounts of RNA. Signals could differ in efficiency of recognition or in efficiency of processing (or both). A weak signal could also result in aberrant processing of pre-mRNA such that the products are neither substrates for further processing nor recoverable as polyadenylated RNA. (Although we have not proven that the RNAs produced by using different signals are equally stable, the data in Fig. 5 suggest that all *tk* RNAs in this study have very similar cytoplasmic stabilities.)

The models that we examine are simple; each is consistent with some of our data. To make the first three models consistent with all the data requires the addition of a number

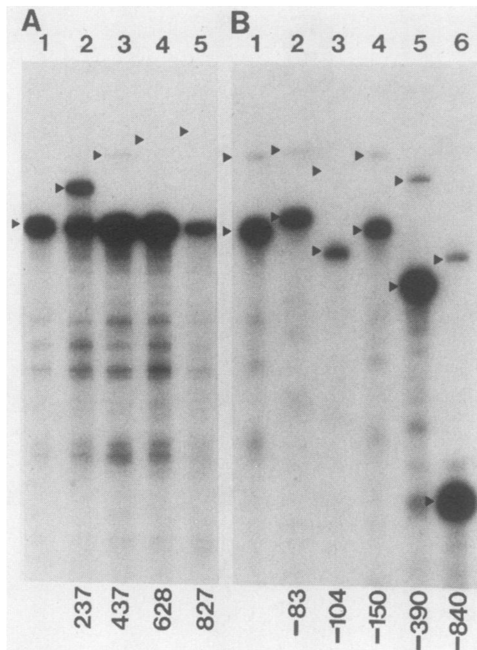

FIG. 8. Analysis of 3′ ends of RNAs produced from constructs with two polyadenylation signals at various distances from the *tk* promoter. Cos-1 cells were transfected with $pSVL_1$, $pSVL_2$, $pSVL_{2+437}$, $pSVL_{2+628}$, $pSVL_{2+827}$, $^{-83}pSVL_{2+437}$, $^{-104}pSVL_{2+437}$, $^{-150}pSVL_{2+437}$, $^{-390}pSVL_{2+437}$, and $^{-840}pSVL_{2+437}$ (C). Cytoplasmic RNA was isolated 48 h later and analyzed by S1 nuclease protection. The probes used were the same DNAs used for transfection, 3′-end-labeled at the *Mlu*I site at the 5′ end of the *tk*-coding region. Protected fragments were separated on 3% denaturing polyacrylamide gels and autoradiographed as described in Materials and Methods. (A) Lanes: 1, $pSVL_1$; 2, $pSVL_2$; 3, $pSLV_{2+437}$; 4, $pSVL_{2+628}$; 5, $pSVL_{2+827}$. The numbers below the lanes give the distance (in base pairs) between the two SVL signals in each construct. Arrowheads mark the positions corresponding to the RNA species produced from each construct. All protected fragments were of the expected size. After extremely long exposures, a very small amount of RNA ending at the second SVL signal in $pSVL_{2+827}$ was sometimes seen. (B) Lanes: 1, $pSVL_{2+437}$; 2, $^{-83}pSVL_{2+437}$; 3, $^{-104}pSVL_{2+437}$; 4, $^{-104}pSVL_{2+437}$; 5, $^{-390}pSVL_{2+437}$; 6, $^{-840}pSVL_{2+437}$. The numbers below the lanes refer to the amount of DNA (in base pairs) deleted from the 5′ end of the *tk*-coding region. (C) Diagram of the size and locations of insertions and deletions. The insertion constructs were produced by duplication of the SVL signal plus various amounts of DNA immediately 5′ to it in $pSVL_1$.

of complex modifications. Although these complexities may exist, we believe that they make these mechanisms less likely. The fourth model, involving 3′-to-5′ scanning from the point of transcription, is consistent with all our data.

(i) The polyadenylation signal is recognized by a soluble factor associating randomly with pre-mRNA. If this were the case, the probability of using a signal would be dependent on its affinity for this factor, the time available for binding, the concentration of the factor, and the processing efficiency of the signal once it is recognized. Since all our constructs contained only one type of signal, all signals in a transcript would have equal affinity for the factor and equal processing efficiency. This could explain why, in general, shorter RNAs were more abundant than their longer counterparts; the first signal transcribed would have a temporal advantage over more distal signals. Increasing the distance between two signals would increase this advantage.

This model does not easily explain some data. If a signal is weak because it is poorly recognized or poorly processed, then increasing the number of signals should increase the probability of producing a productive polyadenylation complex. In either case, the amount of RNA produced should increase when the number of signals is increased. We saw no such increase in the amount of RNA produced when the number of signals was increased.

The last 88 signals in $p88_2$ and $p88_4$ were used more than their next upstream neighbors. This should not happen if random association determines signal use. The possibility that a weak signal results in aberrantly processed, unrecoverable mRNA is consistent with our data on this point. Regardless of why weak signals are weak, this model offers no ready explanation for why adding a signal downstream affects the use of only the nearest upstream signal. If random association governs signal recognition, all signals should be equally affected by addition of new signals.

(ii) The polyadenylation site is recognized by some factor associated with the RNA polymerase. If this factor recognized a signal as it was transcribed, a second signal would never be used, unless signals could be skipped. Additional signals would give those RNAs on which signals were skipped another chance to be processed. This would increase the amount of RNA produced. The frequency of skipping for a given signal would have to depend on the presence of downstream signals, since how often a signal was used depended on whether or not there was another signal downstream of it. This would require that the polymerase-bound

recognition factor somehow detect the presence of a polyadenylation signal downstream of the polymerase, before that downstream signal was transcribed. Although possible, this seems unlikely.

(iii) The polyadenylation site is recognized by a factor scanning from 5′ to 3′ along the RNA, independent of the RNA polymerase. If scanning must begin at the 5′ cap, this model and the previous one are mechanistically similar, except that an RNA could be scanned multiple times. Signal skipping would have to occur to allow downstream signals to be used. With poorly recognized or inefficiently processed signals, increasing the number of signals should increase the amount of RNA produced. Positing that a weak signal results in aberrantly processed pre-mRNA solves this problem but does not explain why adding a signal downstream affects only its nearest upstream neighbor; neither does it explain the pattern of signal use in constructs containing multiple 88 signals.

If scanning begins at any point along the RNA, then the probability of using a signal would depend on the amount of RNA upstream of the signal that can act as the initiation point for scanning. Increasing the distance between two signals would increase the use of the downstream signal. When we increased this distance, the upstream signal was used more and the downstream signal was used less (Fig. 8A). Decreasing the distance from the promoter to the first of several signals would decrease the use of the first signal. Changing this distance had no effects in our constructs (Fig. 8B).

(iv) Sites are recognized by a factor scanning from 3′ to 5′ along the nascent transcript. Scanning would begin only at (or immediately 5′ to) the transcription complex. (The factor might associate with protein-free RNA present only immediately 5′ to the transcription complex.) When the factor reaches a signal, it would form a recognition complex. Signal efficiency would be determined by the efficiency of processing, once recognition occurs. Recognition complexes would abort if they did not cause processing, leaving the RNA available for processing at that or other sites. There are no sequence requirements for scan initiation nor limits on the number of scans that can occur per transcript, other than that imposed by nuclear turnover of unprocessed precursors. This model makes several predictions.

Increasing the number of signals would not increase the amount of RNA produced. At any given time, the only signal available for recognition is the most recently synthesized one. Since recognition is effectively 100% efficient, this 3′-most signal would block scanning factors from reaching signals further upstream. In essence, the maximum number of active signals in a transcript at any time is one. Since there is only one active signal, increasing the number of signals just partitions polyadenylation among several sites.

The frequency of signal recognition would depend on the distance to the next signal downstream. Increasing this distance (i.e., the signal-free 3′-flanking region) would increase the use of the upstream signal by increasing the amount of time available for recognition factors to begin a scan. Changing the distance between the promoter and the first signal would have no effect, since scans beginning in this region would run off the 5′ end of the RNA without encountering a poly(A) signal. Data shown in Fig. 8 are consistent with these predictions.

Adding a signal downstream of several others would affect only the amount of RNA produced from the nearest upstream neighbor of the new signal. In this case, the length of the signal-free 3′-flanking region for this nearest upstream neighbor would be reduced, resulting in less RNA being produced by using that site. The signal-free 3′-flanking region of signals further upstream would be unaltered by the addition of another signal, so the amount of RNA produced from these would be unaffected.

The frequency of signal use would depend on the ability of the signal to cause processing and the distance to the next signal. For constructs containing several very weak signals, the 3′-most signal could be used more than its upstream neighbor, since this last signal has a very long signal-free 3′-flanking region. This explains why the last 88 signal in $p88_2$ and in $p88_4$ was used more than the next signal upstream. With very efficient signals, there would be very little transcript that was not processed or committed to use upstream signals when this last signal was synthesized.

This model fails if signals were skipped because they were recognized poorly, since this would cause an increase in the amount of RNA produced from plasmids containing multiple weak signals. This model also fails if scanning could begin at any point in the transcript, since this would increase the probability of formation of a productive complex when multiple weak signals are present. It is possible that weak signals would result in aberrant processing of preRNAs. This is consistent with our data and is being tested at present.

Other signals. We have not addressed the problem of transcripts that contain two (or more) signals of different strength. Studies in progress assay various combinations of the signals used here. Other investigators have produced plasmids or viruses containing two signals. SV40 mutant 1471 (9) contained two late polyadenylation signals and produced SV40 late RNA with a processing pattern similar to that of $pSVL_2$. Lanoix et al. (18) produced a polyomavirus mutant with two copies of the DNA containing both the early and late polyadenylation signals. Plasmids containing the rat preproinsulin poly(A) signal 5′ to the SV40 late signal (12), the mouse β-major globin signal 5′ to the SV40 late signal (13), and the signals associated with the exons for the secreted and membrane carboxyl termini of mouse μ immunoglobulin (29) have been produced. In all these studies, patterns of polyadenylation were similar to what we observed. In the polyomavirus and immunoglobulin poly(A) signal studies (the only two that assayed RNA quantity), there was no significant change in the amount of RNA produced when the number of signals was doubled. Furthermore, Nishikura and Vuocolo (29) found that the relative order of the signals, not their sequence, caused the difference in signal use.

Several studies have suggested that the switch between membrane and secreted forms of immunoglobulin μ heavy chain might be regulated by polyadenylation signal choice (1, 17, 23, 33, 34). The polyadenylation site that is used to produce RNA encoding the secreted form (μ_s) is contained in RNA removed during splicing of RNA encoding the membrane-associated form (μ_m). Peterson and Perry (31) suggested that competition between the splicing and polyadenylation reactions determines which RNA is produced; slowing one reaction relative to the other would favor production of one form of the RNA. They showed that cell-type-specific regulation of RNA production required that signals for the production of both RNAs be in *cis*. Increasing the length of the intron downstream of the first (μ_s) poly(A) signal caused the production of more μ_s-encoding RNA relative to the amount of μ_m-encoding RNA. Peterson and Perry (31) suggested that this supported the idea that there was competition between splicing and polyadenylation. We see the same pattern when we increase the

distance between two SVL signals in the absence of splicing. Although competition between splicing and poly(A) site choice may be important to cell-type-specific regulation, the distance between the two poly(A) sites has a major effect on the levels of each RNA.

Our studies define a set of rules by which poly(A) sites are chosen in the presence of multiple polyadenylation signals. A model involving 3'-to-5' scanning along the nascent RNA is consistent with our data, whereas several other models are not. We are examining these models further using other signals and conditions.

ACKNOWLEDGMENTS

We thank Michele Chamberlain and Catherine Heath for excellent technical assistance and all the members of our laboratory for helpful discussions. Critical reading of the manuscript by David Auble, C. Robertson McClung, R. J. Patterson, Jeff Silverman, and Betty Werner is greatly appreciated.

This work was supported by Public Health Service grants GM33998 (C.N.C.) and GM11382 (R.M.D.) from the National Institutes of Health.

LITERATURE CITED

1. **Alt, F. W., A. L. W. Bothwell, M. Knapp, E. Siden, E. Mather, M. Koshland, and D. Baltimore.** 1980. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chain is directed by mRNAs that differ at their 3' ends. Cell **20**:293–301.
2. **Babich, A., L. T. Feldman, J. R. Nevins, J. E. Darnell, and C. Weinberger.** 1983. Effect of adenovirus on metabolism of specific host mRNAs transport control and specific translational discrimination. Mol. Cell. Biol. **3**:1212–1221.
3. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
4. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1413–1423.
5. **Cole, C. N., and G. M. Santangelo.** 1983. Analysis in Cos-1 cells of processing and polyadenylation signals by using derivatives of the herpes simplex virus type 1 thymidine kinase gene. Mol. Cell. Biol. **3**:267–279.
6. **Cole, C. N., and T. P. Stacy.** 1985. Identification of sequences in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation. Mol. Cell. Biol. **5**: 2104–2113.
7. **Conway, L., and M. Wickens.** 1985. A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3' termini in frog oocytes. Proc. Natl. Acad. Sci. USA **82**:3949–3953.
8. **Favaloro, J., R. Treissman, and R. Kamen.** 1981. Transcription maps of polyoma virus-specific RNA: analysis by two dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
9. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
10. **Garger, S. J., O. M. Griffith, and L. K. Grill.** 1983. Rapid purification of plasmid DNA by a single centrifugation in a two step cesium chloride-ethidium bromide gradient. Biochem. Biophys. Res. Commun. **117**:835–842.
11. **Gil, A., and N. J. Proudfoot.** 1984. A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. Nature (London) **312**:473–474.
12. **Gruss, P., and G. Khoury.** 1981. Expression of simian virus 40-rat preproinsulin recombinant in monkey kidney cells: use of preproinsulin RNA processing signals. Proc. Natl. Acad. Sci. USA **78**:133–137.
13. **Hamer, D. H., and P. Leder.** 1979. SV40 recombinants carrying a functional RNA splice junction and polyadenylation site from the chromosomal mouse beta-major globin gene. Cell **17**:737–747.
14. **Hart, R. P., M. A. McDevitt, H. Ali, and J. R. Nevins.** 1985. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the simian virus 40 polyadenylation sites. Mol. Cell. Biol. **5**:2975–2983.
15. **Hart, R. P., M. A. McDevitt, and J. R. Nevins.** 1985. Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences. Cell **43**:677–683.
16. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cultures. J. Mol. Biol. **26**:365–369.
17. **Kemp, D. J., G. Morahan, A. F. Cowan, and A. W. Harris.** 1983. Production of RNA for secreted immunoglobulin chains does not require transcription termination 5' to the M exons. Nature (London) **301**:84–86.
18. **Lanoix, J., R. W. Tseng, and N. H. Acheson.** 1986. Duplication of functional polyadenylation signals in polymavirus does not alter efficiency of polyadenylation or transcription termination. J. Virol. **58**:733–742.
19. **Leff, S. E., M. G. Rosenfeld, and R. M. Evans.** 1986. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Annu. Rev. Biochem. **55**:1091–1118.
20. **Luthman, H., and G. Magnusson.** 1983. High efficiency polyoma DNA transfection of chloroquine treated cells. Nucleic Acids Res. **11**:1295–1308.
21. **Mandel, M., and A. Higa.** 1970. Calcium-dependent bacteriophage DNA interaction. J. Mol. Biol. **53**:159–162.
22. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage lambda. Proc. Natl. Acad. Sci. USA **72**:1184–1188.
23. **Mather, E. L., K. J. Nelson, J. Haimovich, and R. P. Perry.** 1984. Mode of regulation of immunoglobulin mu and delta chain expression varies during B-lymphocyte maturation. Cell **36**:329–338.
24. **McLaughlan, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucleic Acids Res. **13**:1347–1368.
25. **Moore, C. L., and P. A. Sharp.** 1984. Site-specific polyadenylation in a cell-free reaction. Cell **36**:581–591.
26. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41**:845–855.
27. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tjian.** 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell **25**:373–384.
28. **Nevins, J. R., and M. C. Wilson.** 1981. Regulation of adenovirus-2 gene expression at the level of transcription termination and RNA processing. Nature (London) **290**:113–119.
29. **Nishikura, K., and G. A. Vuocolo.** 1984. Synthesis of two mRNAs by utilization of alternate polyadenylation sites: expression of SV40-mouse immunoglobulin chain gene recombinants in Cos-1 monkey cells. EMBO J. **3**:689–699.
30. **Nussinov, R.** 1986. Sequence signals which may be required for efficient formation of mRNA 3' termini. Nucleic Acids Res. **14**: 3557–3571.
31. **Peterson, M. L., and R. P. Perry.** 1986. Regulation of production of μ_m and μ_s mRNA requires linkage of the poly(A) addition sites and is dependent on the length of the μ_s-μ_m intron. Proc. Natl. Acad. Sci. USA **83**:8883–8887.
32. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3' noncoding region sequences in eukaryotic messenger RNA. Nature (London) **263**: 211–214.
33. **Rogers, J., P. Early, C. Carter, K. Calame, M. Bond, L. Hood, and R. Wall.** 1980. Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin μ chain. Cell **20**:99–106.
34. **Rogers, J., N. Fasel, and R. Wall.** 1985. A novel RNA in which the 5' end is generated by cleavage at the poly(A) site of immunoglobulin heavy-chain secreted mRNA. Mol. Cell. Biol. **6**:4749–4752.
35. **Sadofsky, M., and J. C. Alwine.** 1984. Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. Mol. Cell. Biol. **4**:1460–1468.
36. **Sadofsky, M., S. Connelly, J. L. Manley, and J. C. Alwine.** 1985. Identification of a sequence element on the 3' side of AAUAAA which is necessary for simian virus 40 late mRNA 3' end

processing. Mol. Cell. Biol. **5**:2713–2719.

37. **Setzer, D. R., M. McGrogan, and R. T. Schimke.** 1982. Nucleotide sequence surrounding multiple polyadenylation sites in mouse dihydrofolate reductase gene. J. Biol. Chem. **257**:5143–5147.
38. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
39. **White, R. T., P. Berg, and L. P. Villarreal.** 1982. Simian virus 40-rabbit-globin recombinants lacking late mRNA splice sites express cytoplasmic RNAs with altered structures. J. Virol. **42**:262–274.
40. **Yen, J.-Y. J., and R. E. Kellems.** 1987. Independent 5′- and 3′-end determination of multiple dihydrofolate reductase transcripts. Mol. Cell. Biol. **7**:3732–3739.
41. **Zhang, F., R. M. Denome, and C. N. Cole.** 1986. Fine-structure analysis of the processing and polyadenylation region of the herpes simple virus type 1 thymidine kinase gene by using linker scanning, internal deletion, and insertion mutations. Mol. Cell. Biol. **6**:4611–4623.