

Relative Roles of Signals Upstream of AAUAAA and Promoter Proximity in Regulation of Human Immunodeficiency Virus Type 1 mRNA 3' End Formation

JAMES D. DEZAZZO,† JEANNINE M. SCOTT, AND MICHAEL J. IMPERIALE*

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0620

Received 2 July 1992/Returned for modification 3 August 1992/Accepted 17 September 1992

At least two mechanisms have been implicated in regulating poly(A) site use in human immunodeficiency virus type 1 (HIV-1): inhibition of basal signals within 500 nucleotides (nt) of the cap site, leading to specific suppression of the 5' poly(A) site, and stimulation of basal signals by long terminal repeat U3 sequences, leading to specific activation of the 3' poly(A) site. We determined the relative contributions of these mechanisms in a HeLa cell transcription/processing reaction and by transient transfection analysis. In vitro, the efficiency of basal signals is equivalent close to (270 nt) and far from (1,080 nt) the promoter and is stimulated at least 30-fold in both positions by upstream U3 sequences. In vivo, U3 sequences also enhance processing at both positions. There are two additional effects when the poly(A) site is close to the cap site: at least a 15-fold reduction in total RNA levels and a 5-fold decrease in relative levels of RNA processed at the HIV-1 site in constructs containing U3. Both effects are overcome by insertion of upstream splicing signals in an orientation-dependent manner. Splicing appears to influence poly(A)⁺ RNA levels by two distinct mechanisms: stabilizing nuclear transcripts and directly stimulating 3' end formation. It is proposed that upstream elements play major roles in regulating poly(A) site choice and in controlling the subsequent fate of polyadenylated RNA. The impact of these findings on mechanisms of mRNA biogenesis in the HIV-1 provirus is discussed.

During infection, the RNA genome of the retrovirus is transcribed into double-stranded DNA, which upon insertion into the host chromosome can become a transcriptionally active provirus (26). The provirus includes long terminal repeats (LTRs), each of which is divided into three functionally delineated regions: U3, R, and U5 (Fig. 1). The transcription start site is at the U3/R junction, and the polyadenylation [poly(A)] site is at the R/U5 junction. Each LTR also encodes signals controlling transcription initiation and core elements required for 3' end processing. The core processing elements include the AAUAAA signal located 10 to 30 nucleotides (nt) upstream of the poly(A) site and a GU-rich signal located 5 to 50 nt past the site (2). Such terminally redundant transcription and 3' end processing signals pose special regulatory problems to retroviruses. Promoter elements must be recognized only in the 5' LTR, and 3' end processing signals must be used efficiently only in the 3' LTR to ensure adequate levels of viral gene products. Indeed, in cells infected with human immunodeficiency virus type 1 (HIV-1), most if not all detectable transcripts originate from the 5' promoter and are processed at the 3' poly(A) site (4).

Given these regulatory constraints, a provirus must have mechanisms for discriminating promoter and 3' end processing elements in one LTR from those in the other. In the case of promoter control, transcription elongation from the 5' LTR is believed to disrupt stable assembly of transcription initiation complexes in the 3' LTR (7, 12). In the case of 3' end processing control, there are at least two different

regulatory strategies, depending on the configuration of core processing signals in the LTR (for reviews, see references 14 and 19). In most avian retroviruses and in human T-cell leukemia virus type I, the AAUAAA signal is encoded in the U3 region. Consequently, in these viruses a complete set of core poly(A) signals is transcribed only from the 3' LTR, allowing exclusive activation of the 3' poly(A) site. In most other retroviruses, including HIV-1, however, the AAUAAA signal is encoded in the R region (Fig. 1). Here, a complete set of core signals is transcribed from both LTRs, presenting a choice of processing sites in the primary transcript. Therefore, in retroviruses such as HIV-1, mechanisms presumably exist to minimize use of the signals in the 5' LTR and to maximize their use in the 3' LTR.

Mechanisms controlling poly(A) site selection in HIV-1 have been extensively investigated with the use of both plasmid and viral systems (3, 4, 9, 24, 25, 27; for a review, see reference 14). It is clear from these studies that U3 sequences stimulate poly(A) site use in the 3' LTR. On both circular plasmids and linear viral chromosomes, these sequences increase the relative use of either the HIV-1 site or a heterologous site at least 20-fold (3, 9, 21, 25). The mechanisms controlling poly(A) site use in the 5' LTR are not clear, however. By extrapolation, studies of processing at the 3' LTR suggest that the core poly(A) signals in the 5' LTR are intrinsically weak because they lack U3 sequences. Complicating this issue are two studies which report that poly(A) site-to-cap site distances of less than 500 nt suppress the use of core processing signals (4, 27). In addition, attempts to study HIV-1 RNA processing close to the promoter in vivo have been hampered by the observation that steady-state RNA levels decline when poly(A) signals are placed close to the promoter (3, 4, 27). Although the

* Corresponding author.

† Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724.

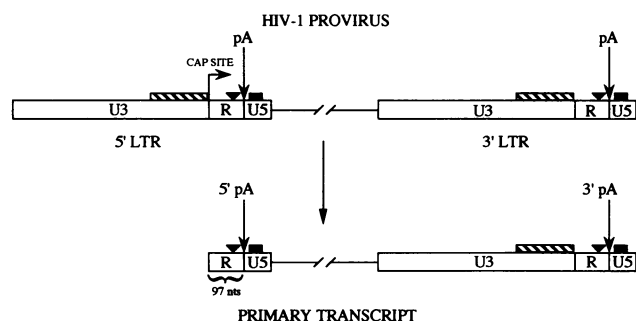


FIG. 1. Structure of the HIV-1 provirus. The HIV-1 provirus includes LTRs and is transcribed into a pre-mRNA with terminally redundant R and U5 sequences. The cap site and poly(A) (pA) sites are indicated. Solid triangles and boxes flanking each poly(A) site (arrow) represent the core AAUAAA and downstream processing signals, respectively. Hatched boxes represent regulatory processing signals in U3.

basis for this decrease is not known, it is not attributable to changes in cytoplasmic stability and is observed in a variety of systems with both retroviral and nonretroviral poly(A) signals. Without an explanation, this effect necessarily limits the conclusions of these studies. For example, decreases in levels of RNA transcripts may be due to posttranscriptional events other than decreased 3' end processing. Recently, it has been reported that T7 precursors in which the poly(A) site is close to the cap site are processed efficiently in vitro (24).

To further address these regulatory issues, we have studied the combined regulatory roles of promoter proximity and U3 sequences by developing in vitro and in vivo strategies that provide for the first time a detailed view of processing events close to the promoter. Our results demonstrate a strong positive role for upstream sequences, and less of a role for promoter proximity, in regulating 3' end processing. We present evidence that the general decrease in RNA levels close to the promoter is due to nuclear posttranscriptional events other than 3' end processing. This conclusion is supported by the finding that poly(A)⁺ mRNA levels in HIV-1 can be stimulated by the insertion of splicing signals. We discuss the impact of these findings on regulation of RNA levels in the HIV-1 provirus.

MATERIALS AND METHODS

Reagents. Enzymes and molecular linkers were purchased from New England Biolabs, Inc., Bethesda Research Laboratories, Inc., and Boehringer Mannheim Biochemicals, Inc. Actinomycin D (mannitol) was purchased from Sigma Chemical Co., and radiochemicals were purchased from Amersham Corp.

Transfections and RNA isolation. The human 293 cell line (11) was maintained in Dulbecco's modified Eagle medium supplemented with 10% fetal calf serum and used in all experiments. Transfections were performed as previously described (9), with 15 μ g of assay plasmid and 5 μ g of control plasmid. Total RNA was isolated 48 h later by the acid guanidinium thiocyanate-phenol-chloroform extraction method of Chomczynski and Sacchi (6).

In vitro transcription processing reactions. Nuclear extracts were prepared by the method of Dignam et al. (10) and dialyzed in the presence of 0.5 mM phenylmethylsulfonyl fluoride. Conditions for processing reactions are described in

detail elsewhere (28). Briefly, circular or linear plasmid DNA was incubated for 2 h at 30°C in a reaction mixture containing 50% extract, 3% polyvinyl alcohol, 2 mM MgCl₂, 20 mM creatine phosphate, and 0.6 mM each ATP, CTP, UTP, and GTP. Reactions were terminated by adding 150 μ l of ETS (10 mM Tris-HCl [pH 8.0], 10 mM EDTA, 0.5% sodium dodecyl sulfate), standardized by adding 5 μ g of control RNA, and extracted once with phenol-chloroform. The reaction products were isolated by ethanol precipitation with 15 μ g of yeast RNA as the carrier and analyzed by S1 nuclease mapping as described below. To ensure that these reactions were carried out under conditions of factor excess, we performed control transcription/processing reactions with template concentrations ranging from 0.04 to 3.0 μ g per reaction. The amount of 3' end processing in all four constructs increased linearly with template concentration up to 0.3 μ g. The in vitro reactions were therefore performed under nonsaturating conditions with equimolar amounts of DNA template.

Plasmid constructions. All plasmids were grown in *Escherichia coli* DH5 α , screened by restriction enzyme mapping, purified by the Qiagen plasmid preparation kit, and quantified by reaction with diphenylamine. All constructs used in this study include a 4-bp mutation in the *tar* region in R, previously shown to block transcriptional activity of the LTR sequences but not RNA processing (9, 16). Plasmids pEH9tar-L3 and pEHX Δ U3-L3 were used in cloning and have been described previously (9). Splicing signals flanking the second leader exon of the adenovirus major late transcription unit were obtained from pMXSVL (17).

To facilitate sequential cloning steps, a plasmid cassette was first constructed by inserting a 435-nt *Sst*I-*Pvu*II fragment spanning the promoter, cap site, and first 30 nt of the adenovirus type 2 major late transcription unit (nt 5634 to 6069) into the *Sst*I and *Sma*I sites of pGEM3Zf(+) (Promega). This plasmid, pZMLP, allowed insertion of splicing signals into a *Bam*HI site 33 nt downstream of the cap site and poly(A) signals into an *Xba*I site 36 nt downstream of the cap site. To construct pU3C-L3, an 805-nt *Sca*I-*Xba*I fragment from pEH9tar-L3 was inserted into the *Xba*I site of pZMLP via a *Sca*I-*Xba*I linkage. To add the poly(A) sites in p Δ U3C-L3, an 807-nt *Xba*I fragment from pEHX Δ U3-L3 was inserted into the *Xba*I site of pZMLP, and then a *Bgl*II-*Sal*I fragment was replaced with the corresponding fragment from pEH9tar-L3 to place the *tar* mutation in the HIV-1 sequences.

To make pU3F-L3 and p Δ U3F-L3, the *Xba*I fragment in pEH9tar-L3 was replaced with the *Xba*I fragments from pU3C-L3 and p Δ U3C-L3, respectively. For construction of prvsU3-L3, psvU3-L3, prvs Δ U3-L3, and psv Δ U3-L3, a 211-nt *Eco*RI-*Bam*HI fragment from plasmid pMXSVL, consisting of two exons flanking a 120-nt intron, was inserted in both orientations into the *Bam*HI site of pU3C-L3 and p Δ U3C-L3 via an *Eco*RI-*Bgl*II linkage. To construct p5'U3-L3, an 89-nt *Eco*RI-*Hind*III fragment from pMXSVL, spanning the 5' exon and first 39 nt of the intron, was inserted into the *Bam*HI site of pZMLP via a *Bgl*II linkage. Finally, to construct p3'U3-L3, a 122-nt *Hind*III-*Bam*HI fragment from pMXSVL, spanning the distal 81 nt of the intron and adjoining 3' exon, was inserted into the *Bam*HI site of pZMLP via a *Hind*III-*Bgl*II linkage.

S1 nuclease mapping. Poly(A)⁺ RNA from each transfection or transcription/processing reaction was purified by oligo(dT) chromatography and coprecipitated with 25 μ g of yeast RNA as the carrier and 10 to 20 ng of probe (5×10^5 to 10×10^5 cpm) made from the corresponding plasmid. Probes

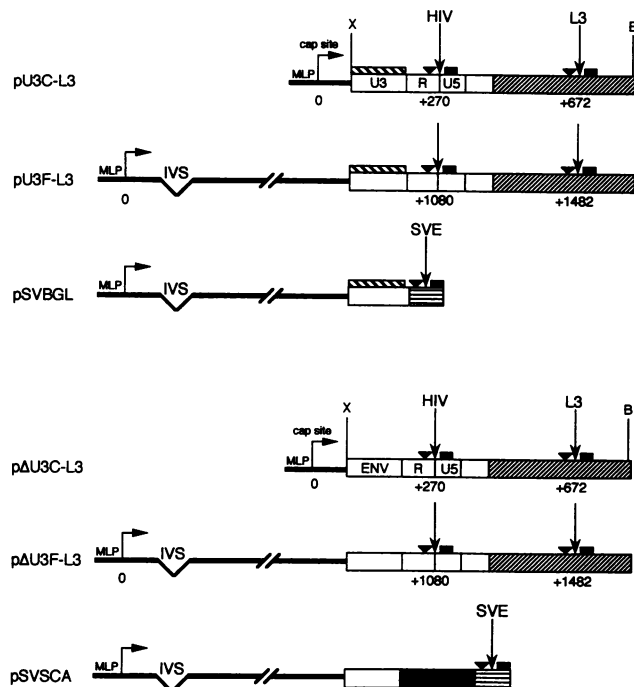


FIG. 2. Tandem poly(A) site constructs and reference plasmids. All transcription units are driven by the adenovirus major late promoter (MLP). Cap site-to-poly(A) site distances (in nucleotides) are indicated below the poly(A) sites (HIV and L3) in each tandem construct. The body of the adenovirus E1A gene, which contains an intron (IVS), is present between the promoter and LTR sequences in the far constructs. In the reference plasmids, the tandem sites are replaced by the simian virus 40 early poly(A) site (SVE). Solid triangles and boxes are defined in Fig. 1. U3 regulatory signals, present in the top three constructs, are replaced by HIV envelope sequences (ENV) in the bottom three constructs. Abbreviations for restriction sites: X, *Xba*I; B, *Bam*HI.

used to detect 3' end processing were 3' end labeled at the *Xba*I site and extend to the *Bam*HI site. Probes used to detect splicing were 5' end labeled at the *Bgl*II site and extend to sites upstream of the cap site: a *Stu*I site in the Δ U3-L3 series or a *Xho*I site in the U3-L3 series. The pellet was resuspended in 50 μ l of hybridization buffer [80% formamide, 10 mM piperazine-*N,N'*-bis(2-ethanesulfonic acid) (pH 6.5), 0.4 M NaCl], heated to 90°C for 3 min, and incubated at 52°C overnight. Following hybridization, each reaction mixture was diluted with 450 μ l of ice-cold S1 digestion buffer (0.4 M NaCl, 30 mM sodium acetate [pH 4.5], 1 mM ZnSO₄, 10 μ g of denatured salmon sperm DNA, 250 U of S1 nuclease). The samples were incubated at 25°C for 2 h, extracted with phenol-chloroform, and precipitated with 5 μ g of carrier RNA. Protected probe fragments were separated on a 6% polyacrylamide-8 M urea gel and exposed to Kodak XAR-5 or XRP film at -70°C. The resolved products were quantified by excising them from the gel and counting them in a liquid scintillation counter.

RESULTS

We began this study by constructing plasmids pU3F-L3 and p Δ U3F-L3 (Fig. 2), in which the HIV-1 poly(A) site with its core processing signals is positioned 1,080 nt from the cap site. Upstream of the R region in pU3F-L3 are 140 nt of U3 sequence, shown previously to stimulate HIV-1 RNA proc-

essing (3, 9, 21, 25). In p Δ U3F-L3, these sequences are replaced by a segment from the HIV-1 envelope region which is identical in size and has no effect on 3' end processing (9). To study the effects of promoter proximity on HIV-1 RNA processing, we generated constructs pU3C-L3 and p Δ U3C-L3 (Fig. 2), in which the HIV-1 site is positioned 270 nt downstream of the same promoter. The far and close sets of plasmids were then assayed by two different methods: in vitro transcription/processing reactions in a HeLa cell nuclear extract (28) and transient transfection into the human 293 cell line (11). These cells express factors that efficiently drive transcription from the major late promoter used in these constructs (1). In all constructs, the HIV-1 site is followed by the adenovirus L3 poly(A) site as a default site. Thus, both methods rely on *cis* competition between the HIV-1 site and the L3 site, with the results expressed as ratios of processing at the two sites.

HIV-1 processing signals are not suppressed within 300 nt of the cap site in vitro. We first incubated the four constructs in a HeLa cell nuclear extract, using conditions which support transcription and RNA processing in a coupled reaction. Transcripts in these reactions are properly initiated by RNA polymerase II and are accurately 3' end processed (28). Since this is a soluble system in which a test plasmid incubates for only 2 h, we reasoned that it might provide greater sensitivity in detecting short, processed transcripts. In addition, this assay allows us to directly assess the effects of U3 sequences and promoter proximity on 3' end processing, since the poly(A)⁺ RNA levels will not reflect other possible events dependent on an intact nucleus. Following incubation of each plasmid, poly(A)⁺ RNA was isolated and analyzed by S1 nuclease mapping of the 3' ends. The products (Fig. 3A) were subsequently quantified by band excision and scintillation counting. The values are displayed in Table 1. In both the far and close constructs, U3 sequences stimulate the relative level of HIV-1 RNA approximately 30-fold. Moreover, the total level of HIV-1 and L3 RNA, compared with that of control RNA, does not decrease in the close constructs (Fig. 3A; compare lanes 3 and 4 or 5 and 6).

To determine whether these results reflect regulation by U3 sequences at the level of 3' end processing, we performed the following control experiments. First, we evaluated the stability of HIV-1 and L3 transcripts in vitro. It could be argued, for example, that relative to the close constructs, 3' end processing in the far constructs is actually much more efficient, but the resulting poly(A)⁺ transcripts are much less stable. To test this possibility, we added α -amanitin to each transcription/processing reaction after 2 h and examined the amount of poly(A)⁺ HIV-1 and L3 transcripts remaining 2 h later (Fig. 3B). There is no significant turnover of any RNA during the course of the reaction, indicating that the levels of HIV-1 and L3 poly(A)⁺ mRNAs seen in vitro are not being influenced by differential mRNA stability. Second, we tested whether the processed transcripts originating from the close constructs in vitro were, in fact, short transcripts. It was possible that processing at poly(A) sites close to the promoter was inefficient and that we were actually detecting the 3' ends of primary transcripts that had circled the plasmid. To test this possibility, we linearized pU3C-L3 at a *Sca*I site 1,948 bp downstream of the L3 poly(A) site and compared processing from this template with that of its circular counterpart. As shown in Fig. 3C, there is no decrease in processing at the HIV-1 site in the linearized template; therefore, 3' end processing does occur at the proximal poly(A) site. The same result is

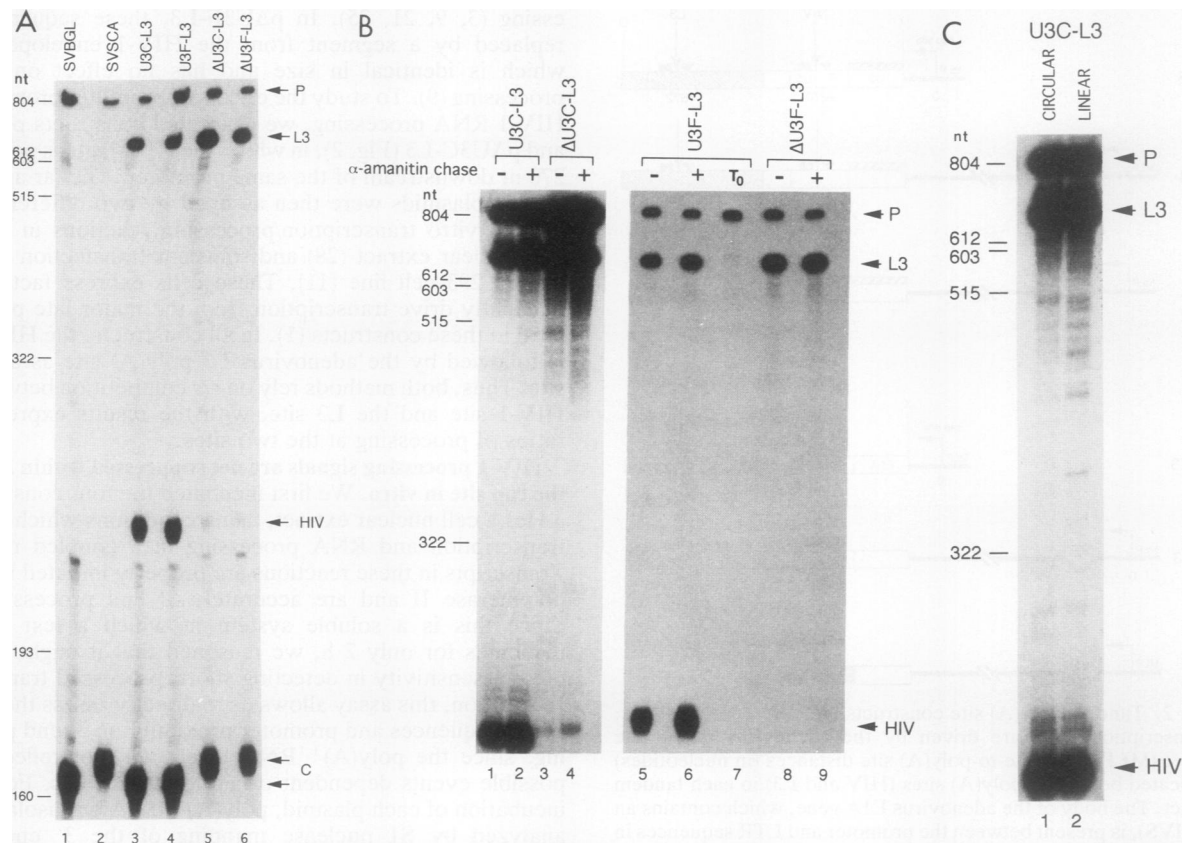


FIG. 3. S1 nuclease analysis of poly(A)⁺ RNA from *in vitro* transcription/processing reactions. Locations are indicated for probes (P) and for S1 products representing control RNA processed at the simian virus 40 site (C) and assay RNA processed at the HIV-1 and L3 poly(A) sites. (A) Processing patterns of close versus far constructs. Tandem poly(A) site constructs were incubated for 2 h in HeLa cell nuclear extracts and assayed by S1 nuclease protection. Poly(A)⁺ reactions included control RNA from pSVBGL (lanes 3 and 4) and pSVSCA (lanes 5 and 6). S1 analysis of control RNA alone is shown in lanes 1 and 2. (B) Stability of transcripts from tandem poly(A) site constructs. After 2 h, reactions were either terminated (–) or continued for 2 h in the presence of 2 mM α -amanitin (+). In lane 7, α -amanitin was added at the start of the incubation period. (C) Processing patterns of circular versus linear DNA templates. pU3C-L3 was incubated for 2 h as either a linear or supercoiled template, and poly(A)⁺ RNA was assayed as described above.

obtained with p Δ U3C-L3 or with a plasmid linearized just 5' of the major late promoter (data not shown).

Poly(A)⁺ RNA levels *in vivo* are regulated in multiple ways. Having obtained evidence that U3 sequences, and not promoter proximity, are the primary determinant of HIV-1 poly(A) site selection *in vitro*, we next tested the relative contribution of these factors *in vivo*. We transfected each of the four constructs with a control plasmid into the 293 cell line, isolated RNA 48 h later, and quantified the amount of HIV-1 and L3 poly(A)⁺ mRNA by S1 nuclease mapping (Fig. 4) and scintillation counting of excised bands (Table 2). In agreement with the *in vitro* results presented above and *in vivo* results from previous studies (3, 4, 9, 21, 25, 27), U3 sequences stimulate the relative levels of HIV-1 poly(A)⁺ RNA *in vivo*. Furthermore, we noted two significant differences between the far and close constructs *in vivo* that are not observed *in vitro*. First, whereas *in vitro* there was no difference in the total amount of poly(A)⁺ mRNA made by the corresponding far and close constructs, *in vivo* there is at least a 15-fold decrease in total poly(A)⁺ mRNA made from the close constructs relative to the control (compare pU3C-L3 and pU3F-L3 or p Δ U3C-L3 and p Δ U3F-L3). Second, whereas *in vitro* U3 sequences enhance the relative level of HIV-1 mRNA to the same degree in the close and far

constructs, *in vivo* they increase the relative level of HIV-1 mRNA in the far construct about fivefold more than in the close constructs (compare pU3C-L3 and pU3F-L3).

As a control, we tested whether these changes could be accounted for by differential stability of the poly(A)⁺ products. After blocking transcription from each plasmid 48 h after transfection by the addition of actinomycin D, we analyzed the poly(A)⁺ mRNA remaining at various times afterward by S1 nuclease mapping. The half-lives of the HIV-1 and L3 transcripts, relative to that of control RNA, are essentially the same in all the close and far constructs (data not shown). Thus, the 15-fold changes in the steady-state mRNA levels *in vivo* cannot be explained by differential stability of the transcripts.

Splicing signals have stimulatory effects on poly(A)⁺ RNA levels *in vivo*. What could explain why a change in mRNA levels between the close and far constructs occurs *in vivo* but not *in vitro*? One obvious difference between the close and far constructs is their genetic structure: introns are present in the far constructs but are absent in the close constructs (Fig. 2). Several recent studies have revealed that introns can dramatically influence steady-state mRNA levels by modulating multiple steps in RNA metabolism, including polyadenylation (5, 13, 22).

TABLE 1. Quantification of in vitro data

Construct	HIV/L3 ^a
pU3C-L3	1.3 ± 0.3
pU3F-L3	1.6 ± 0.5
pΔU3C-L3	0.04 ± 0.01
pΔU3F-L3	0.05 ± 0.01

^a The amount of RNA processed at the HIV poly(A) site over that processed at the L3 poly(A) site. Results are the averages of four experiments ± standard deviations.

To test for such effects in our system, we inserted a 220-nt splicing cassette into each close construct between the cap site and HIV-1 poly(A) site to make prvsΔU3-L3 (Fig. 5A) and prvsU3-L3 (Fig. 6A) and examined whether it could confer the processing phenotype of the far constructs. To control for nonspecific spacing effects, we also inserted the splicing cassette in the reverse orientation to make psviΔU3-L3 (Fig. 5A) and psviU3-L3 (Fig. 6A). The S1 analysis of poly(A)⁺ mRNA derived from transfection of the ΔU3-L3 series is shown in Fig. 5B; quantitative data are shown in Table 2. Inserting the splicing cassette in the correct orientation raises the total amount of poly(A)⁺ mRNA in prvsΔU3-L3 to nearly the same level as in pΔU3F-L3. By contrast, no such increase is observed upon inserting the splicing cassette in the reverse orientation. The half-lives of HIV-1 and of L3 RNAs relative to that of control mRNA are equivalent for prvsΔU3-L3 and psviΔU3-L3, as determined by an actinomycin D chase (data not shown). Therefore, the 20-fold stimulation of mRNA production by the splicing cassette is not due to selective stabilization of steady-state transcripts. Moreover, this stimulation is not a result of simply increasing the distance of the HIV-1 site from the promoter. To test for splicing of the poly(A)⁺ mRNAs in prvsΔU3-L3, a 5'-end-labeled probe was used to map use of the splice acceptor site by S1 nuclease analysis. As shown in Fig. 5C, efficient splicing occurs in prvsΔU3-L3 but is not detectable in psviΔU3-L3. The analysis of the ΔU3-L3 set of constructs therefore reveals that in an orientation-dependent manner, a functional splicing cassette greatly increases the steady-state level of poly(A)⁺ mRNA.

The S1 analysis of poly(A)⁺ mRNA derived from transient transfection of the U3-L3 set of constructs is shown in Fig. 6B and Table 2. Similar to its effect in the ΔU3-L3 plasmid series, a splicing cassette in the correct orientation increases

TABLE 2. Quantification of transfection data

Construct	Amt of HIV/L3 RNA processed ^a	HIV + L3/control ^b
pU3C-L3	6.2 ± 1.8	0.09 ± 0.01
pU3F-L3	35.2 ± 10.7	3.2 ± 0.6
prvsU3-L3	32.8 ± 8.2	1.8 ± 0.2
psviU3-L3	5.2 ± 2.8	0.05 ± 0.02
p5'U3-L3	3.1 ± 1.8	0.03 ± 0.01
p3'U3-L3	15.1 ± 5.1	0.2 ± 0.03
pΔU3C-L3	0.12 ± 0.03	0.17 ± 0.01
pΔU3F-L3	0.12 ± 0.02	2.6 ± 0.39
prvsΔU3-L3	0.21 ± 0.05	2.2 ± 0.5
psviΔU3-L3	0.11 ± 0.03	0.07 ± 0.03

^a See Table 1 for definition. Results are averages of at least three experiments for each construct ± standard deviations.

^b Total amount of RNA processed at both sites, normalized to the control value.

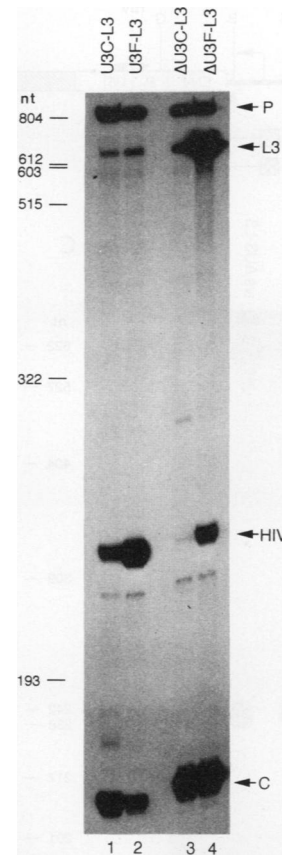


FIG. 4. S1 nuclease analysis of poly(A)⁺ RNA from transfections. Fifteen micrograms of each assay plasmid was cotransfected into the 293 cell line with 5 μg of either pSVBGL (lanes 1 and 2) or pSVSCA (lanes 3 and 4). Abbreviations are the same as for Fig. 3.

the steady-state mRNA levels in prvsU3-L3 to those seen with the far construct (compare prvsU3-L3 with pU3F-L3). In addition to this general increase, the splicing cassette in prvsU3-L3 has a specific stimulatory effect: it increases the relative abundance of HIV-1 mRNA over L3 mRNA back to the 35-fold preference seen in pU3F-L3. Therefore, in the U3-L3 series, the splicing cassette stimulates both the total level of processing and the distribution of processing between the two sites. These effects are not due to differential stability of the HIV-1 and L3 steady-state transcripts (data not shown). Thus, the changes in HIV-1 mRNA levels mediated by the splicing cassette appear to reflect regulation of nuclear RNA metabolism, not cytoplasmic stability.

To localize the stimulatory regions of the splicing cassette more precisely, we tested the ability of either the 5' or 3' splice site alone to enhance HIV-1 mRNA levels in the U3-L3 series. (Fig. 6B, lanes 5 and 6). The 5' splice site alone does not stimulate total or relative levels of HIV-1 mRNA, whereas the 3' splice site stimulates both about threefold (compare p5'U3-L3 and p3'U3-L3 in Table 2). When 5'-end-labeled probes were used to map use of the splice acceptor site, we found that splicing occurs with high efficiency in prvsU3-L3 but is not detectable in psviU3-L3 or in p5'U3-L3 (Fig. 6C). In p3'U3-L3, however, splicing appears to occur at an intermediate level, presumably through use of a cryptic splice donor which is not mapped in this assay. The partial stimulation of HIV-1 mRNA in

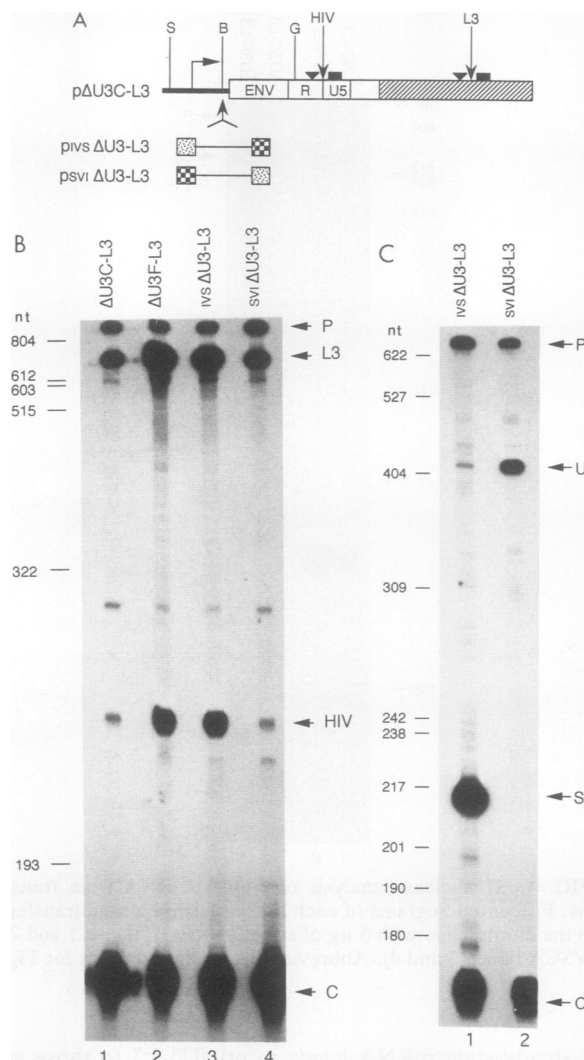


FIG. 5. Effects of splicing signal addition to pΔU3C-L3. (A) Diagram showing the insertion into pΔU3C-L3 of a 211-nt intron in the correct (ivs) and incorrect (svi) orientations. Abbreviations for restriction sites: S, *StuI*; B, *BamHI*; G, *BgIII*. (B) S1 nuclease analysis of poly(A)⁺ RNA from transfection of ΔU3-L3 series constructs. Assay conditions and abbreviations are the same as for Fig. 4. (C) S1 nuclease analysis of poly(A)⁺ RNA to detect splicing of transcripts. Locations are indicated for probes (P) and for S1 products representing control (C), spliced (S), and unspliced (U) transcripts.

p3'U3-L3 may therefore reflect partial stimulation by a weakly spliced intron. Thus, insertion of a functional splicing cassette upstream of the HIV-1 site in the close constructs switches the processing phenotype to that of the corresponding far construct in an orientation-dependent manner.

DISCUSSION

A number of studies of HIV-1 processing have led to the general agreement that U3 sequences enhance the processing efficiency of HIV-1 core elements in the 3' LTR, but there is contention as to their overall regulatory significance (3, 4, 9, 21, 24, 25, 27; for a review, see reference 14). This debate arises primarily because of two reports that poly(A)

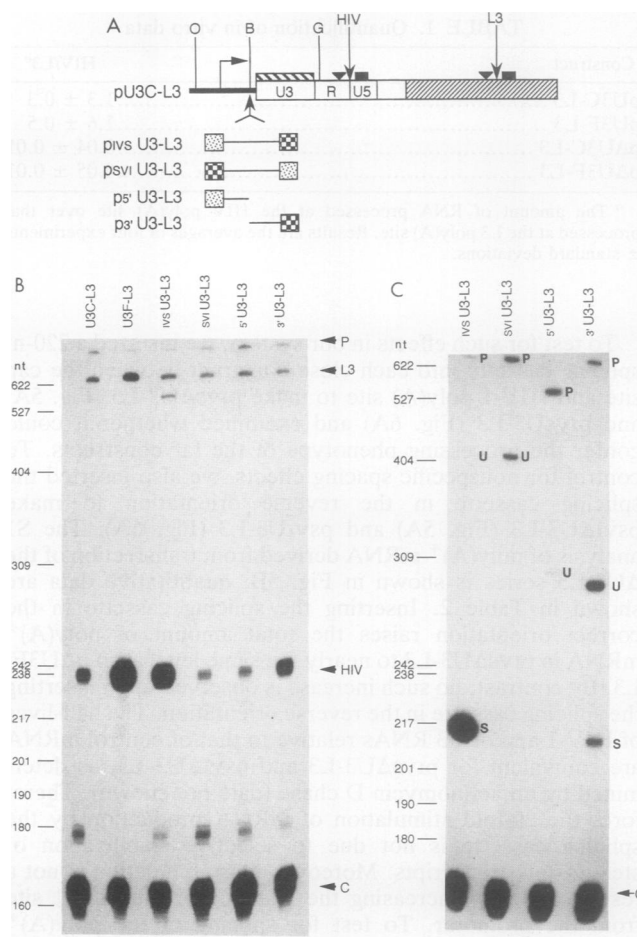


FIG. 6. Effects of splicing signal addition to pU3C-L3. (A) Diagram showing the insertion into pU3C-L3 of a 211-nt intron in the correct (ivs) and incorrect (svi) orientations and of 5' or 3' splicing signals alone. Abbreviations for restriction sites: O, *XhoI*; B, *BamHI*; G, *BgIII*. (B) S1 nuclease analysis of poly(A)⁺ RNA from transfection of U3-L3 series constructs. Assay conditions and abbreviations are the same as for Fig. 4. (C) S1 nuclease analysis of poly(A)⁺ RNA to detect splicing of transcripts. Abbreviations are the same as for Fig. 5C.

site-cap site proximity may suppress processing in the 5' LTR (4, 27). Here we have described in vitro and in vivo experiments in which high levels of poly(A)⁺ RNA are seen close to the promoter, allowing us to directly assess regulatory mechanisms operating within 500 nt of the cap site. Employing an in vitro transcription/processing system, we find that poly(A) site-cap site proximity has little or no effect either on the distribution of processing at the two sites or on the total level of processing. U3 sequences, however, stimulate HIV-1 processing efficiency about 30-fold at both positions, suggesting that they are the primary regulatory determinant in vitro. Valsamakis et al. have also recently shown that T7 precursor RNAs containing U3 sequences are processed more efficiently in vitro than are those without U3 sequences (24). This demonstration of U3 function in vitro should greatly facilitate analyses of the exact role of U3 sequences in the processing reaction.

In transfection studies, U3 sequences also stimulate processing both close to and far from the promoter, but

in the close constructs, we detect decreases in both steady-state RNA levels and the stimulatory effect of U3 processing signals. Steady-state decreases have been noted in other studies but have not been explained (3, 4, 27). A plausible explanation is suggested by several clues obtained in this study: (i) they are not caused by differential stability of steady-state RNA; (ii) they are not due to a general inhibition of transcription because, *in vitro*, poly(A)⁺ RNA levels in the close constructs accumulate to levels seen with the far constructs; and (iii) they cannot be exclusively explained by changes in transcription or by poly(A) site-cap site proximity because an upstream intron restores levels to those seen in the far constructs in an orientation-dependent manner. On the basis of these data, we propose that transcripts made from the close constructs undergo 3' end processing but are turned over rapidly in the nucleus. We suspect that splicing signals boost the levels of poly(A)⁺ RNA by at least two mechanisms. First, stimulation of total poly(A)⁺ mRNA levels in the Δ U3-L3 and U3-L3 series may reflect a general coupling of splicing, 3' end processing, and nuclear export (5, 13, 22). This hypothesis is supported by the correlation of increased splicing efficiency with increased steady-state mRNA levels in the U3-L3 series. By such mechanisms, splicing signals may rescue close transcripts from nuclear degradation pathways and allow them to access the steady-state pool. Second, their stimulation of HIV-1 RNA levels in the U3-L3 series suggests that splicing signals may specifically stimulate 3' end processing, in support of the exon definition model proposed by Berget and coworkers (18, 20). The requirement of U3 sequences for stimulation suggests that terminal exons are better defined with certain classes of poly(A) signals or with relatively strong signals. In summary, the *in vivo* studies confirm the conclusions of the *in vitro* studies and lead to the additional finding that upstream splicing signals can dramatically influence poly(A)⁺ RNA levels, perhaps by multiple mechanisms.

Our ability to detect efficient HIV-1 processing 270 nt from the promoter *in vitro* and within 500 nt *in vivo* contrasts with the inability of Weichs *et al.* (27) to detect HIV-1 processing within this range. This difference is not due to our use of a different promoter, as the effect that they saw was promoter independent (27). However, while we measured processing at both sites in a *cis* competition assay, these investigators measured processing at the downstream competitor site only; processing at the HIV-1 site was inferred. Our results also contrast with those of Cherrington and Ganem (4), who reported a gradual decrease in relative efficiency of core processing signals within 500 nt of the cap site. They also observed a decrease in the absolute level of RNA at these distances. Our experiments indicate that these investigators may not have detected all of the HIV-1 3' end processing events, since posttranscriptional nuclear events other than 3' end processing may modulate poly(A)⁺ RNA levels when processing signals are close to the promoter. Therefore, there is no direct evidence that poly(A) site-cap site proximity directly inhibits HIV-1 processing. Nevertheless, we cannot completely discount such an effect of promoter proximity, since we have not examined processing at the HIV-1 site closer than 270 nt from the cap site, even though this is well within the 400- to 500-nt distance within which proximity effects have been reported in HIV-1 (4, 27) and other viral systems (15, 23).

How do these findings relate to the control of mRNA levels in the HIV-1 provirus? At the 3' LTR, processing is stimulated by upstream U3 signals and perhaps also by upstream introns. By mediating spliceosome assembly, the

introns may protect the processed RNA from nuclear degradation and direct it to pathways leading to the steady-state pool. In the case of unspliced viral mRNAs, the interaction of the virally encoded Rev protein with the Rev response element provides a similar means for directing nuclear RNA to the steady-state pool (reviewed in reference 7). It may be a combination of these mechanisms in HIV-1 which allows highly efficient expression of mRNAs encoding viral proteins.

At the 5' LTR, the absolute strength of the core poly(A) signals has not been determined. Other investigators have used relative strengths of such core signals, measured in *cis* competition assays, to argue that such core sequences are efficiently processed and therefore need to be suppressed in the 5' LTR (4, 27). We feel, however, that one cannot make statements about the absolute strength of a poly(A) site on the basis of such assays, since the relative strength varies depending on the competitor. Indeed, even in constructs containing only a single known poly(A) site, there is competition with cryptic sites on the plasmid (9). Given the limitations of these experimental models, we believe that the absolute efficiency of core processing elements cannot be reliably determined in a context outside of the HIV-1 provirus. It is clear, however, that HIV-1 core processing signals are consistently stimulated 30-fold by U3 sequences. Moreover, in this study, we have shown *in vivo* that the combined effect of upstream splicing and U3 signals can shift the relative use of the HIV-1 site from approximately 0.1 ($p\Delta$ U3C-L3) to greater than 30 (*prvs*U3-L3). Thus, the issue remains as to whether competition actually occurs between the poly(A) sites in the 5' and 3' LTRs on the provirus. In this regard, when an HIV-1 poly(A) site without U3 sequences is placed upstream of a second HIV-1 site with U3 sequences, similar to the situation on the provirus, use of the second site predominates (9). In conclusion, studies of 3' end processing at the HIV-1 poly(A) site are sketching a complex regulatory picture that suggests an interplay of the various steps of mRNA biogenesis in modulating the steady-state output of poly(A)⁺ mRNA. We expect that this information will be relevant to cellular as well as viral genes.

ACKNOWLEDGMENTS

We thank Rich Jove, Oveta Fuller, and members of the Imperiale laboratory for stimulating discussions; Jay Kilpatrick for technical assistance and for reviewing the manuscript; Sue Berget for providing the splicing cassette vector pMXSVL; and Jim Alwine for communicating results prior to publication.

This work was supported by PHS grant GM34902 from NIH. J.M.S. is supported in part by NIH training grant GM07544, and M.J.I. is supported in part by Faculty Research Award FRA-388 from the American Cancer Society.

REFERENCES

1. Aiello, L., R. Guilfoyle, K. Huebner, and R. Weinmann. 1979. Adenovirus 5 DNA sequences present and RNA sequences transcribed in transformed human embryo kidney cells (HEK-Ad-5 or 293). *Virology* **94**:460-469.
2. Bohnlein, S., J. Hauber, and B. R. Cullen. 1989. Identification of a U5-specific sequence required for efficient polyadenylation with the human immunodeficiency virus long terminal repeat. *J. Virol.* **63**:421-424.
3. Brown, P. H., L. S. Tiley, and B. R. Cullen. 1991. Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. *J. Virol.* **65**:3340-3343.
4. Cherrington, J., and D. Ganem. 1992. Regulation of polyadenylation in human immunodeficiency virus (HIV): contributions of promoter proximity and upstream sequences. *EMBO J.* **11**:

- 1513–1524.
5. **Chiou, H. C., C. Dabrowski, and J. C. Alwine.** 1991. Simian virus 40 late mRNA leader sequences involved in augmenting mRNA accumulation via multiple mechanisms, including increased polyadenylation efficiency. *J. Virol.* **65**:6677–6685.
 6. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156–159.
 7. **Cullen, B. R., P. T. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian retroviruses—implications for the promoter insertion model of leukaemogenesis. *Nature (London)* **307**:241–245.
 8. **Cullen, B. R., and M. H. Malim.** 1991. The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators. *Trends Biochem. Sci.* **16**:346–350.
 9. **DeZazzo, J. D., J. E. Kilpatrick, and M. J. Imperiale.** 1991. Involvement of long terminal repeat U3 sequences overlapping the transcription control region in human immunodeficiency virus type 1 mRNA 3' end formation. *Mol. Cell. Biol.* **11**:1624–1630.
 10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
 11. **Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**:59–74.
 12. **Herman, S. A., and J. M. Coffin.** 1986. Differential transcription from the long terminal repeats of integrated avian leukosis virus DNA. *J. Virol.* **60**:497–505.
 13. **Huang, M. T., and C. M. Gorman.** 1990. Intervening sequences increase efficiency of RNA 3' processing and accumulation of cytoplasmic RNA. *Nucleic Acids Res.* **18**:937–947.
 14. **Imperiale, M. J., and J. D. DeZazzo.** 1991. Poly(A) site choice in retroelements: déjà vu all over again? *New Biol.* **3**:531–537.
 15. **Iwasaki, K., and H. M. Temin.** 1990. The efficiency of RNA 3'-end formation is determined by the distance between the cap site and the poly(A) site in spleen necrosis virus. *Genes Dev.* **4**:2299–2307.
 16. **Kliwer, S., J. Garcia, L. Pearson, E. Soultanakis, A. Dasgupta, and R. Gaynor.** 1989. Multiple transcriptional regulatory domains in the human immunodeficiency virus type 1 long terminal repeat are involved in basal and E1A/E1B-induced promoter activity. *J. Virol.* **63**:4616–4625.
 17. **Niwa, M., and S. M. Berget.** 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.* **4**:1552–1559.
 18. **Niwa, M., and S. M. Berget.** 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.* **5**:2086–2095.
 19. **Proudfoot, N.** 1991. Poly(A) signals. *Cell* **64**:671–674.
 20. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**:84–94.
 21. **Russnak, R., and D. Ganem.** 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes Dev.* **4**:764–776.
 22. **Ryu, W. S., and J. E. Mertz.** 1989. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J. Virol.* **63**:4386–4394.
 23. **Sanfacon, H., and T. Hohn.** 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. *Nature (London)* **346**:81–84.
 24. **Valsamakis, A., N. Sheek, and J. C. Alwine.** 1992. Elements upstream of the AAUAAA within the human immunodeficiency virus polyadenylation signal are required for efficient polyadenylation in vitro. *Mol. Cell. Biol.* **12**:3699–3705.
 25. **Valsamakis, A., S. Zeichner, S. Carswell, and J. C. Alwine.** 1991. The human immunodeficiency virus type 1 polyadenylation signal: a 3' long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. *Proc. Natl. Acad. Sci. USA* **88**:2108–2112.
 26. **Varmus, H., and R. Swanstrom.** 1984. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), *RNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 27. **Weichs an der Glon, C., J. Monks, and N. J. Proudfoot.** 1991. Occlusion of the HIV poly(A) site. *Genes Dev.* **5**:244–253.
 28. **Wilson-Gunn, S. I., J. E. Kilpatrick, and M. J. Imperiale.** 1992. Regulated adenovirus mRNA 3' end formation in a coupled in vitro transcription/processing system. *J. Virol.* **66**:5418–5424.