The EMBO Journal vol.11 no.4 pp.1513-1524, 1992

Regulation of polyadenylation in human immunodeficiency virus (HIV): contributions of promoter proximity and upstream sequences

Julie Cherrington and Don Ganem

Howard Hughes Medical Institute, Departments of Microbiology and Immunology and Medicine, University of California Medical Center, San Francisco, CA 94143-0502, USA

Communicated by W.Keller

Retroviruses synthesize a terminally redundant genomic RNA that contains canonical polyadenylation signals at both ends. Production of this RNA requires that the 5′ copy of these signals be ignored, while the 3′ copy must be utilized. Two models have been presented for how this occurs in the human immunodeficiency virus, HIV: (i) the core HIV poly(A) signals (AAUAAA and a downstream GU-rich element) might be inefficient and require supplementation by activating sequences found only at the 3′ end of the RNA; or (ii) cap site proximity might actively suppress polyadenylation at the 5′ site. We have examined both possibilities in HIV-infected cells and in cells transfected with a variety of model constructs. We find that infected cells harbor few or no detectable products of 5′ polyadenylation; however, the core HIV processing signals can mediate processing fairly efficiently (65%) when positioned at the 3′ end of heterologous transcripts. While this efficiency can be further increased to >95% by inclusion of upstream sequences from the viral U3 region, the absence of these U3 signals is insufficient by itself to account for 5′ signal bypass. By contrast, the efficiency of these core elements is greatly suppressed when they are positioned within ~450 nucleotides of the cap site. This distance-related suppression can be modestly diminished by insertion of U3 sequences between the cap site and HIV poly(A) signal. We suggest that the primary determinant of 5′ poly(A) site bypass is cap site proximity; the absence of U3 sequences at this position contributes secondarily to the bypass by enhancing the sensitivity of the pA signal to the suppressive effects of promoter proximity.

Key words: HIV/polyadenylation

Introduction

The 3′ ends of virtually all eukaryotic mRNAs are formed by post-transcriptional processing rather than by direct termination. Formation of a mature 3′ end requires an endonucleolytic cleavage of a larger pre-mRNA molecule followed by the addition of ~200 adenosine residues to the 3′ end. Extensive studies have shown that an intact hexanucleotide sequence, AAUAAA, is required for accurate cleavage and polyadenylation both *in vivo* (Proudfoot and Brownlee, 1976; Fitzgerald and Shenk, 1981; Montell *et al.*, 1983; Higgs *et al.*, 1983; Gil and Proudfoot, 1984; Wickens and Stephenson, 1984) and *in vitro* (Hart *et al.*, 1985; Zarkower *et al.*, 1986; Conway and Wickens, 1987). Cleavage and polyadenylation of the pre-mRNA molecule occurs at a specific site 10–30 nucleotides (nt) downstream of this consensus poly(A) signal. An additional, less well defined U-rich or G/U rich sequence 10–30 nt downstream of the cleavage site is also required for the accurate and efficient polyadenylation of many, but not all, mRNAs (Gil and Proudfoot, 1984; McDevitt *et al.*, 1984; Conway and Wickens, 1985; Hart *et al.*, 1985; Sadofsky and Alwine, 1985; McLaughan *et al.*, 1985; Green and Hart, 1988; Zarkower and Wickens, 1988). Together, the hexanucleotide and downstream G/U clusters (where present) constitute the core or basal signals involved in 3′ end processing.

While for many RNAs polyadenylation is unregulated, for other transcripts, particularly those containing multiple poly(A) signals, the differential use of these signals is an important locus of regulation. One important class of transcripts that must regulate poly(A) site use is the genomic RNA of the so-called retroid elements, genetic elements that replicate via reverse transcription of RNA; these include both viruses (retroviruses, hepatitis B viruses, caulimoviruses) and transposons (e.g. *Drosophila copia* and yeast *Ty* elements). Their RNA templates for reverse transcription are terminally redundant and signals for polyadenylation are contained within each redundancy. Thus, synthesis of a full length transcript requires bypass of a polyadenylation signal in the upstream portion of the RNA (i.e. in the 5′ redundancy) but use of this site in the 3′ redundancy (see Figure 1). The experiments reported here investigate the mechanisms governing this differential poly(A) site use in the human immunodeficiency virus, HIV-1.

Three classes of models can be envisaged to explain how poly(A) sites are used during the processing of HIV genomic RNA. 1. The leaky signal model: the poly(A) signal may merely be inefficient, allowing a certain fixed percentage of RNAs to read through on each pass. In this model, no specific regulation of pA site use is required. This mechanism seems to be at play during processing of transcripts of the circular genome of polyoma virus (Acheson, 1984). 2. The promoter proximity model: there may be a minimum distance that the transcribing complex must travel in order to recognize the poly(A) signal and assemble a cleavage/polyadenylation complex. In this model, the core poly(A) signals are sufficient for efficient polyadenylation but the close proximity of the 5′ poly(A) signal to the mRNA start site results in suppression of their recognition or use. This mechanism has been implicated in regulating polyadenylation of spleen necrosis virus (Iwasaki and Temin, 1990), cauliflower mosaic virus (Sanfacon and Hohn, 1990) and hepatitis B viruses (J.M.Cherrington, R.Russnak and D.Ganem, manuscript in preparation). In both of these models, no additional sequence elements beyond those needed for normal polyadenylation are required. 3. The upstream activator model: in this model, the core viral poly(A) signals are postulated to be inefficient; additional sequences 5′ to the poly(A) signal and genomic RNA cap

site must be transcribed before the signal can be efficiently used. Therefore, the 5′ poly(A) signal is ignored because these sequences are not yet present on the transcript, while at the 3′ LTR the sequences are present and the poly(A) signal can then be used efficiently. Upstream sequences that increase polyadenylation efficiency have been identified in several other viruses, including SV40 (Carswell and Alwine, 1989), adenovirus (DeZazzo and Imperiale, 1989), hepatitis B viruses (Russnak and Ganem, 1990) and caulimoviruses (Sanfacon *et al.*, 1991). The possibility that HIV polyadenylation might be regulated as in model 3 has been raised by the recent demonstration by several groups that the U3 region of HIV can upregulate the use of homologous or heterologous downstream poly(A) signals (Russnak and Ganem, 1990; Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991).

Controversy exists regarding the mechanisms of differential poly(A) site use in HIV, with some groups (Weichs an der Glon *et al.*, 1991) favoring distance-related suppression of 5′ poly(A) site use (model 2) while others (Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991) emphasize the role of upstream sequences (model 3). However, to date experiments interpreted as supporting model 3 have not directly examined processing events at the 5′ signals. Rather, they rely exclusively upon estimates of the efficiency of the core HIV poly(A) signals when found in a 3′ context, in the presence or absence of U3 elements. The case for model 3 hinges upon the assertion that the core HIV poly(A) signals are surprisingly inefficient (DeZazzo and Imperiale, 1991; Valsamakis *et al.*, 1991); from this it is inferred that 5′ signals are bypassed due to the lack of U3 sequences.

In this paper, we test directly each of the described models, both in transient transfection assays and in the more authentic context of integrated proviruses established by retroviral infection. Our results indicate that the principal determinant of 5′ poly(A) site bypass is promoter proximity, which suppresses an otherwise fairly efficient processing signal. However, the absence of U3 sequences makes an additional, more modest, contribution, chiefly by increasing the sensitivity of the signal to the suppressive effects of cap proximity. In our view the major role of U3 polyadenylation elements in the viral life cycle is to enhance the processing efficiency at the 3′ LTR.

Results

Poly(A) site use in HIV is regulated

Figure 1A shows the genetic organization of the HIV DNA provirus. Transcription from the provirus results in a 9 kb, terminally redundant genomic RNA which serves as the template for reverse transcription. Additionally, this full length transcript can be extensively spliced to generate RNAs coding for the envelope and regulatory gene products (reviewed in Greene, 1990). As previously discussed, the synthesis of genomic RNA requires that the 5′ poly(A) signal be ignored, while the 3′ signal is used. However, this requirement alone does not indicate that use of the poly(A) site is regulated: regulation implies that the processing efficiency at the 5′ poly(A) site differs from that at the 3′ poly(A) site. Therefore, we directly determined the amount of processing which occurs at each poly(A) signal in proviral DNA.

RNA was harvested from a mass culture of $CD4^+$ HeLa cells bearing multiple stably integrated proviruses derived by infection with the HIV derivative, HIV-gpt (Page *et al.*, 1990). RNase protection analysis was performed using a probe that would distinguish between transcripts polyadenylated at the 5′ LTR, those polyadenylated at the 3′ LTR, and those reading through both signals (Figure 1A). RNAs polyadenylated at the 5′ site would yield a 90 nt fragment, while RNAs reading through the 5′ signal would generate a 180 nt fragment, and RNAs polyadenylated in the 3′ LTR would yield a band of 240 nt. Transcripts reading through both signals would yield a 320 nt fragment. The RNase protection analysis of cytoplasmic (cyto), $poly(A)^+$ (pA) and whole cell RNA (total) from these cells is shown in Figure 1B. The strong intensity of the 180 and 240 nt fragments indicates that the vast majority of transcripts traversing the body of the provirus are polyadenylated efficiently at the 3′ LTR. Additionally, the minor band visible at 320 nt in the whole cell RNA fraction (total RNA lane) further attests that polyadenylation at the 3′ LTR is very efficient (>95%). By contrast, the 90 nt fragment representing the product of processing in the 5′ LTR was not detectable. The failure to detect the small transcript could in principle be due to transcript instability. As will be shown below, however, this is not likely to be the case: mutations that inactivate the 5′ hexanucleotide AAUAAA do not result in increased levels of RNA processed at a downstream signal, as would be predicted by the instability model (see below and Figure 5C). We conclude that polyadenylation at the 5′ LTR is indeed very inefficient, with <10% of transcripts being processed at this site.

Efficiency of the HIV poly(A) signal

Inefficient processing at the 5′ LTR could be due to (i) an efficient poly(A) signal that is suppressed by the proximity of the cap site, or (ii) a poor poly(A) signal that requires upstream sequences to function efficiently. To distinguish between these possibilities, it was important to determine the efficiency of the core HIV poly(A) signals in a context independent of promoter proximity. To do this we employed a tandem poly(A) site assay [sometimes referred to as a poly(A) site competition assay (Weichs an der Glon *et al.*, 1991); see below]. The core HIV poly(A) signals (AAUAAA and downstream GU-rich region; Bohnlein *et al.*, 1989) along with varying amounts of 5′ flanking sequence were cloned downstream of a 1.6 kb c-*src* cDNA; 3′ to the HIV poly(A) site is an SV40 early poly(A) signal (Figure 2A; in our numbering system +1 corresponds to the first A of the HIV AAUAAA). The constructs are driven by the SV40 early promoter; transcripts that are not polyadenylated at the HIV signal can be processed by the downstream SV40 poly(A) signal. These constructs were transfected into COS cells and 48 h later, $poly(A)^+$ RNA was harvested and subjected to Northern analysis using an *src* cDNA probe (Figure 2B). The ratio of transcripts processed at the HIV poly(A) signal and those reading through to the SV40 poly(A) signal was quantified using a phosphorimager (Molecular Dynamics). This ratio provides an estimate of the relative processing efficiency of the HIV pA signal only if both transcripts are equally stable. Therefore, actinomycin D chase experiments were conducted on clones src.HIV−9/SV and src.HIV−94/SV and it was determined that in both cases the processed and readthrough transcripts displayed identical half-lives (data not shown).

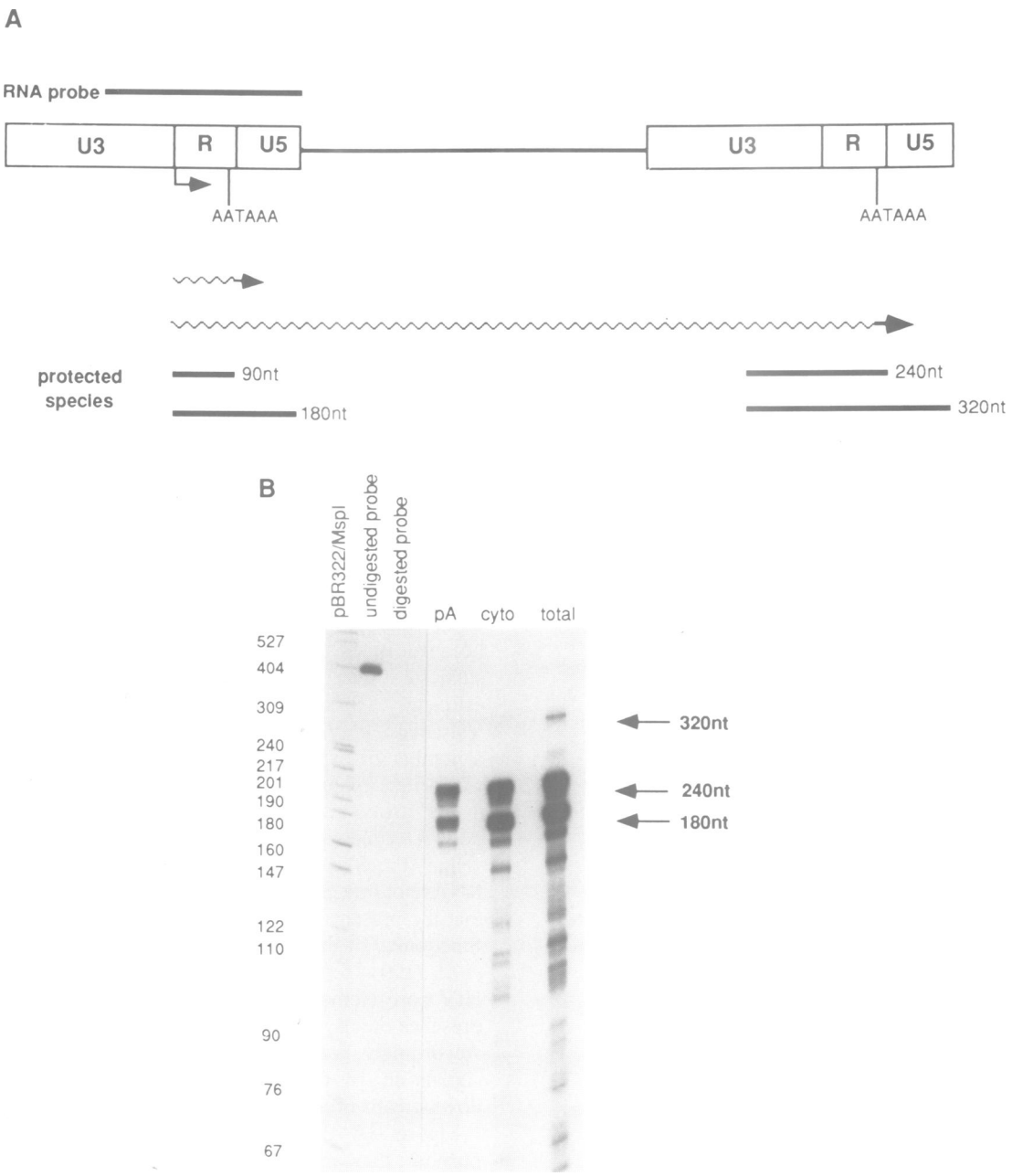

Fig. 1. (**A**) General organization of the HIV provirus and diagram of the RNase protection strategy. The wavy lines depict transcription beginning in the 5′ LTR and polyadenylation occurring at either the 5′ or 3′ pA signal. The RNA probe used for the protection analysis is shown above the provirus and contains HIV sequences from −219 to +100, relative to the first A of the AAUAAA hexanucleotide as +1. 5′ LTR processing, 5′ readthrough, 3′ LTR processing and 3′ readthrough can be distinguished by hybridization to this uniformly labeled RNA probe. The predicted protected fragments of the analysis are shown by bold lines below the provirus. (**B**) RNase protection analysis of pA^+ (pA), cytoplasmic (cyto) or whole cell (total) RNA isolated from $CD4^+$ HeLa cells stably expressing the integrated provirus, HIV-gpt. Positions of protected species predicted in (A) are indicated by arrows at right.

A construct that contained the minimal HIV pA signal (−9 to +100 relative to the HIV hexanucleotide), src.HIV−9/SV, directed efficient processing; 65% of the transcripts were processed at the HIV pA signal while the remaining 35% read through that signal (Figure 2B, lane 1). Clone src.HIV−94/SV contains all the sequences that are normally found between the cap site and the pA signal (the HIV cap site is at −74 relative to its pA signal) and therefore contains nearly the same processing information that would be present during transcription of genomic RNA from the 5′ LTR. This clone also directed efficient processing (Figure 2B, lane 3). Similar processing is observed in src.HIV−55/SV which contains sequences from just downstream of the HIV cap site through the poly(A) signal (Figure 2B, lane 2). These data show that the basal HIV poly(A) signals direct fairly efficient processing (~65%).

In order to determine the effect of U3 sequences on processing efficiency (i.e. reflecting the situation at the 3′ LTR), src.HIV−219/SV was constructed (Figure 2A). As shown in Figure 2B (lane 4), the inclusion of sequences from −94 to −219 increased the processing efficiency of the HIV pA signal to >95%; no readthrough transcript

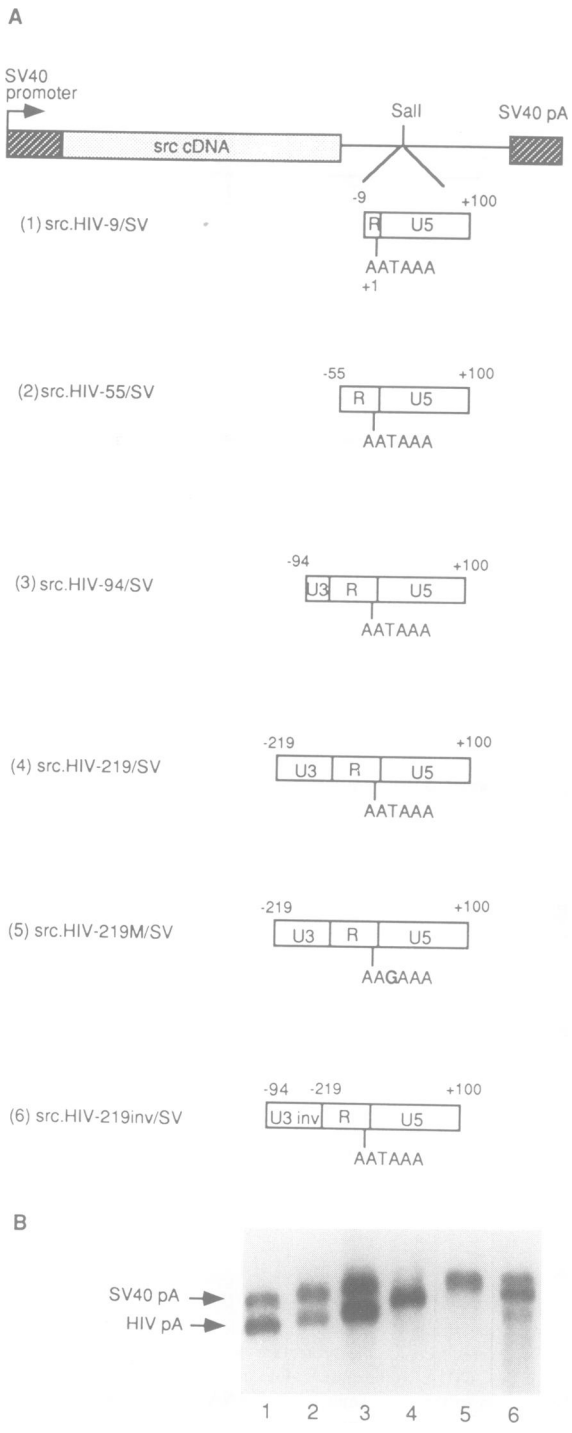

Fig. 2. A system for studying 3′ end processing using an SV40 based expression vector. (**A**) The HIV pA signal containing various amounts of upstream flanking sequence (constructs 1–6) was cloned downstream of a c-*src* cDNA in a vector containing the SV40 origin of replication, allowing for copy number amplification upon transfection into COS cells. c-*src* cDNA expression is driven by the SV40 early promoter; readthrough transcripts are processed at a downstream SV40 pA signal. The 5′ and 3′ limits of the HIV sequences cloned into the *Sal*I site are shown, and are numbered relative to the HIV hexanucleotide as +1. (**B**) Determination of the efficiency of the HIV pA signal by Northern analysis of cytoplasmic pA^+ RNA harvested from COS7 cells transfected with constructs 1–6. The lane numbers correspond to the construct numbers in (A). Arrows to the left show transcripts processed at the HIV pA signal or at the SV40 pA signal. A uniformly labeled RNA probe, containing the 3′ portion of the c-*src* cDNA sequences from nt 884 to 1712, was used in Northern analysis.

was detectable by Northern analysis. To ensure that our failure to detect the readthrough transcript was not due to its instability, a mutant hexanucleotide (AAGAAA) was introduced into the HIV LTR in src.HIV−219/SV producing src.HIV−219M/SV. The hexanucleotide mutation abolished HIV poly(A) site use, and the longer 'readthrough' transcript was readily detected (Figure 2B, lane 5); instability thus cannot explain its absence from lane 4.

To examine the orientation dependence of the U3 sequences in the up-regulation of RNA processing efficiency, the sequences between −219 and −94 were inverted in the construct src.HIV−219inv/SV. Figure 2B, lane 6, shows that these sequences do not function in the opposite orientation. This is consistent with the idea that they function as an RNA element, but further studies will be required to establish this firmly.

Examination of assay variables

The distinction between models 2 (promoter proximity) and 3 (upstream elements) hinges on the assessment of the efficiency of the basal HIV processing signals. We find that the core HIV pA signals are quite efficient, processing 65% of the transcripts which traverse them (Figure 2B). However, because other groups using different constructs have suggested that these signals function substantially less efficiently (Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991), we thought it important to examine a number of experimental variables that might affect our assessment of processing efficiency.

Our assay involves the use of constructs bearing two poly(A) signals, the tester HIV element and a downstream signal (the SV40 early pA signal) used to polyadenylate RNAs not processed at the HIV site. It is formally possible that the presence of a downstream element could alter the functioning of the upstream signal. For instance, DeZazzo *et al.* (1991) reported that the processing efficiency of the HIV core elements increased from 8% to ~50% in the presence of a downstream poly(A) signal from adenovirus. Accordingly, we analyzed the function of the HIV poly(A) signals in *src*-based constructs lacking a heterologous downstream processing signal.

Figure 3A depicts the constructs designed for this purpose; they are analogous in all respects to their cognate 'two poly(A) site' clones of Figure 2. These constructs were cotransfected into COS cells with the β-galactosidase expression vector, pON249 (Geballe *et al.*, 1986); in parallel, the corresponding 'two poly(A) site' src.HIV/SV constructs were similarly examined. Transfection efficiency was determined for each sample by assay of β-galactosidase activity. Poly(A)$^+$ RNA prepared from each transfection was then analyzed by Northern blot hybridization with a labeled *src* probe (Figure 3A); the amount of RNA loaded in each lane was adjusted to normalize for transfection efficiency. As previously observed, in the 'two poly(A) site' constructs the core HIV signals directed processing at ~65% efficiency (lanes 1–3), and inclusion of U3 sequences (lane 4) increased the efficiency from 65 to >95%. Lanes 5–8 demonstrate that when present singly the core HIV pA signals also direct efficient processing and that all of the src.HIV constructs direct processing to a similar degree. If the basal HIV elements were very inefficient, a substantial difference would have been anticipated between constructs lacking (lanes 5–7) and those bearing (lane 8) U3 sequences.

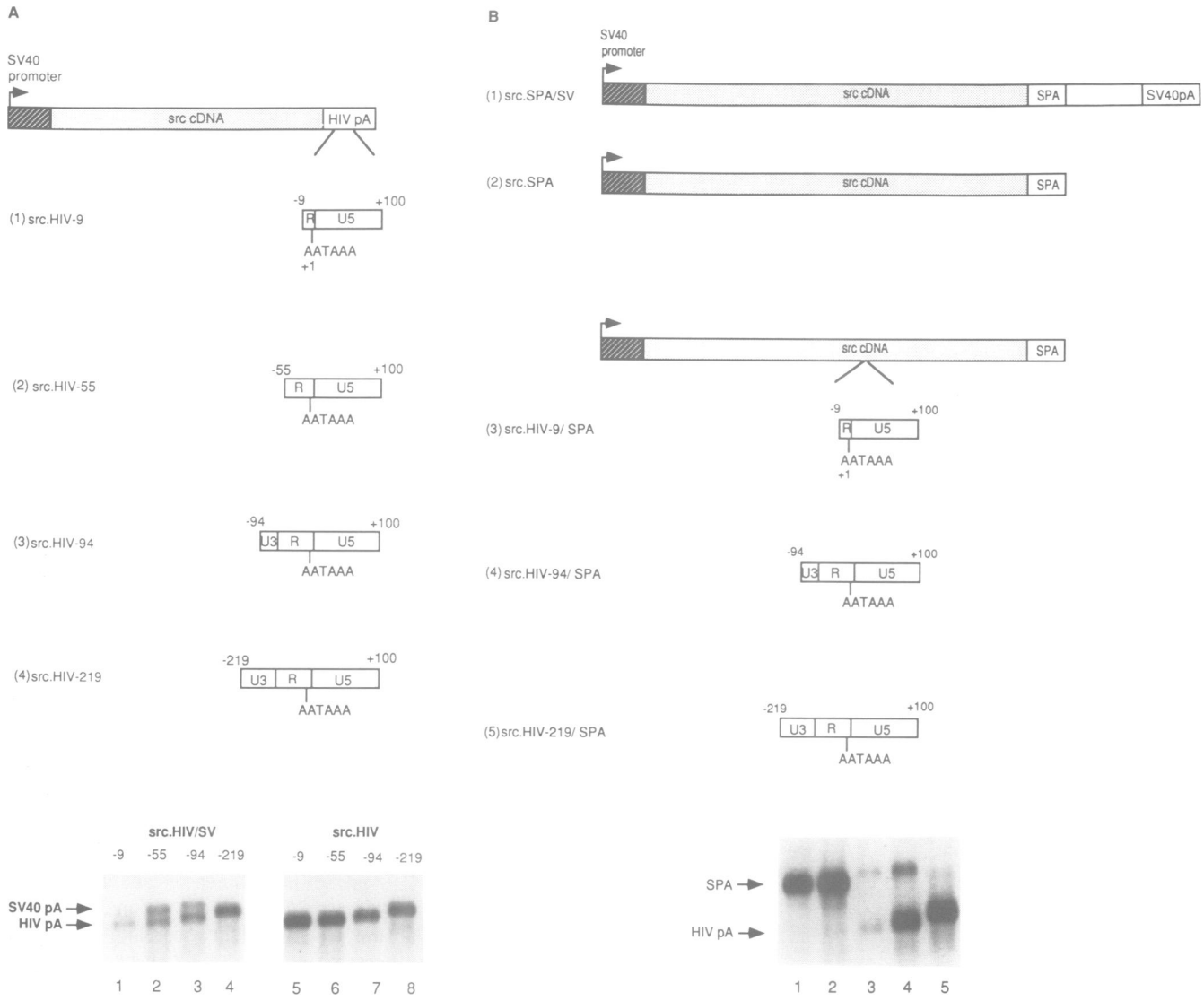

Fig. 3. (**A**) Measurement of the efficiency of the HIV pA signal in the absence of a downstream pA signal. Lanes 1–4 show the results of Northern analysis conducted on pA^+ RNA harvested from COS cells cotransfected with src.HIV/SV constructs 1–4 (Figure 2A and B) and a β-gal expression vector, pON249. Lanes 5–8 show the results of Northern analysis conducted on pA^+ RNA harvested from COS cells cotransfected with src.HIV constructs 1–4 and pON249. The amount of RNA loaded per lane was corrected for transfection efficiency. The numbers above the lanes indicate the 5′ endpoint of the HIV sequences present in the construct being analyzed in that lane. src.HIV/SV denotes constructs with two pA signals; src.HIV denotes constructs with only the HIV pA signal. Arrows to the left show transcripts processed at the HIV pA signal or at the SV40 signal. A uniformly labeled RNA probe, containing the 5′ portion of the c-*src* cDNA sequences from nt 91 to 884, was used in Northern analysis. (**B**) Measurement of the efficiency of the HIV pA signal in the presence of downstream SPA signal. Construct 1 contains the SPA signal positioned 3′ of *src* sequences on a construct containing a downstream SV40 pA signal. Construct 2 contains the SPA signal on a construct containing no rescue signal. Constructs 3, 4 and 5 contain the HIV pA signal and 5′ flanking sequence (numbered relative to the hexanucleotide as +1) cloned into the *src* cDNA. Northern analysis was conducted on cytoplasmic, pA^+ RNA harvested from COS cells transfected with constructs 1–5. The lane numbers correspond to the construct numbers. Arrows to the left show transcripts processed at the HIV pA signal or at the V40 signal. A uniformly labeled RNA probe, containing the 5′ portion of the c-*src* cDNA sequences from nt 91 to 884, was used in Northern analysis. Construct 3 reproducibly generated a lower amount of RNA than constructs 4 and 5; we attribute this to RNA instability or low transcription rate, but have not yet directly distinguished between these possibilities.

(The small stimulatory effect of U3 sequences is not readily observed in this assay, because the magnitude of the U3 effect is within the range of the error of our normalization procedure for transfection efficiency.)

Phosphorimager quantitation of this experiment also shows that the tandem poly(A) sites do not function completely independently of one another. Comparison of the signal intensities of lanes 2–4 with those of their cognates in lanes 6–8 reveals that the intensity of the single bands of the 'one poly(A) site' clones equals the sum of the intensities of the two bands present in the 'two-site' clones. That is, the total amount of $poly(A)^+$ RNA which is generated from a particular construct is the same whether one or two poly(A) signals are present on that construct. This suggests that there is *cis* competition between the signals for components of the processing reaction, as previously proposed by others (Weichs an der Glon *et al.*, 1991).

To examine this issue in another way, we assayed selected 'one poly(A) site clones' by nuclease protection. From cells transfected in parallel with src.HIV−55 or src.HIV−219 (cf. Figure 3A) total RNA was prepared and examined by RNase mapping using the same probe as described in

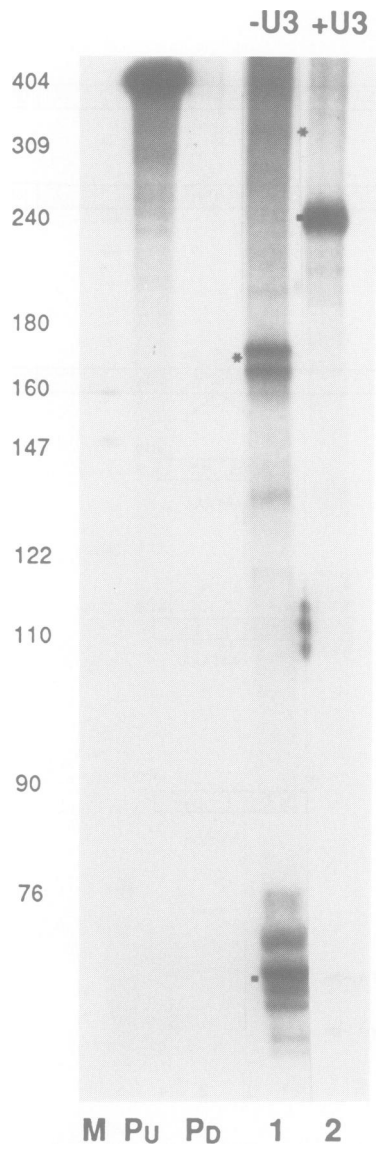

Fig. 4. RNase protection analysis of total RNA derived from two constructs containing one poly(A) site. The constructs transfected for this analysis are shown in Figure 3A as construct 2, src.HIV−55 (−U3) and construct 4, src.HIV−219 (+U3). The uniformly labeled RNA probe described in Figure 1A was used in this analysis. RNA derived from the transfection of the construct without U3 sequences (lane 1) would yield predicted products of 75 nt representing transcripts processed at the HIV pA signal and 155 nt representing readthrough transcripts. RNA derived from the transfection of the construct containing U3 sequences (lane 2) would yield predicted products of 240 nt representing transcripts processed at the HIV pA signal, and 320 nt representing readthrough transcripts. The cluster of protected bands corresponding to transcripts processed at the HIV pA signal are marked by a square, readthrough transcripts are marked by an asterisk. Mol. wt markers shown to the left are pBR322 DNA digested with *Msp*I. PU denotes undigested probe. PD denotes digested probe.

Figure 1. As shown in Figure 4 (lane 1) the src.HIV−55 RNA lacking U3 sequences was still efficiently processed at the HIV poly(A) site, with transcripts polyadenylated there representing ~70% of the total protected material (multiple repeats of this experiment have consistently given values of 70−90% processing). When U3 sequences are present (src.HIV−219; lane 2), virtually all of the protected RNA corresponds to molecules processed at the HIV poly(A) site. This analysis once again indicates that the core signals function rather efficiently, though they can be further bolstered to some degree by U3 elements.

In other experiments we examined whether the distance between the poly(A) sites in 'two poly(A) site' clones would affect the processing efficiency estimates. These studies revealed that the core HIV pA signal of src.HIV−9/SV functions equally efficiently at a distance of 300 bp or 1.0 kb from the downstream SV40 pA signal (data not shown). Next we asked whether the character of the downstream poly(A) signal influenced the use of the HIV signal. To do this we constructed the src.HIV/SPA plasmids shown in Figure 3B. In these constructs the HIV poly(A) signals were cloned into the body of a src cDNA, 900 bp upstream of a synthetic poly(A) signal derived from β-globin (designated SPA; Levitt *et al.*, 1989). These constructs were transfected into COS cells, and poly(A)$^+$ RNA harvested at 48 h post-transfection was analyzed by Northern blotting as before (Figure 3B, lanes 3−5). Lanes 1 and 2 in Figure 3B are controls to show that, as expected, the SPA signal functions efficiently in this assay. When present upstream of an SV40 poly(A) site in plasmid src.SPA/SV, >95% of transcripts are processed at the SPA signal and very little polyadenylation occurs at the SV40 pA signal (lane 1). Construct src.SPA demonstrates that the SPA signal is able to function very efficiently even in the absence of a downstream pA signal (lane 2). Despite their differing levels of expression, the src.HIV/SPA series again revealed processing at the minimal HIV poly(A) signal to be ~60−65% efficient (Figure 3B, lanes 3 and 4). Furthermore, addition of U3 sequences to these constructs also increased the processing efficiency to >95% (lane 5).

These results reaffirm our initial measurements of the basal and stimulated HIV processing efficiencies and indicate that our estimates are not strongly dependent upon a particular configuration of the tester and downstream processing signals. Since processing at the 5′ LTR in proviral DNA is quite inefficient (Figure 1), a mechanism must exist which suppresses the use of the 5′ poly(A) signal.

Promoter proximity reduces the efficiency of the HIV polyadenylation signal

To test the hypothesis that the short distance (74 nt) between the HIV cap site and poly(A) signal is the determinant that suppresses polyadenylation in the 5′ LTR of HIV, we constructed the deletion series diagrammed in Figure 5A. In these constructs, *src* sequences separating the SV40 promoter and HIV pA site were varied from 1800 to 180 nt. The 5′ endpoint of the HIV sequences (−94 relative to the hexanucleotide) present in these constructs was chosen to mimic the situation as it exists at the 5′ LTR, i.e. containing sequences from the cap site to the poly(A) signal. These constructs were transfected into COS cells, RNA was harvested 48 h later and cytoplasmic RNA was analyzed by RNase protection using a probe that distinguishes between processed and readthrough transcripts (diagrammed in Figure 1A). With this probe, a protected fragment of 110 nt represents processing at the HIV pA signal while a 195 nt species represents transcripts which read through the HIV pA signal.

The results of the RNase protection analysis are shown in Figure 5B. Lanes 1−3 demonstrate that at a distance of >600 bp from the promoter the HIV poly(A) signal directs

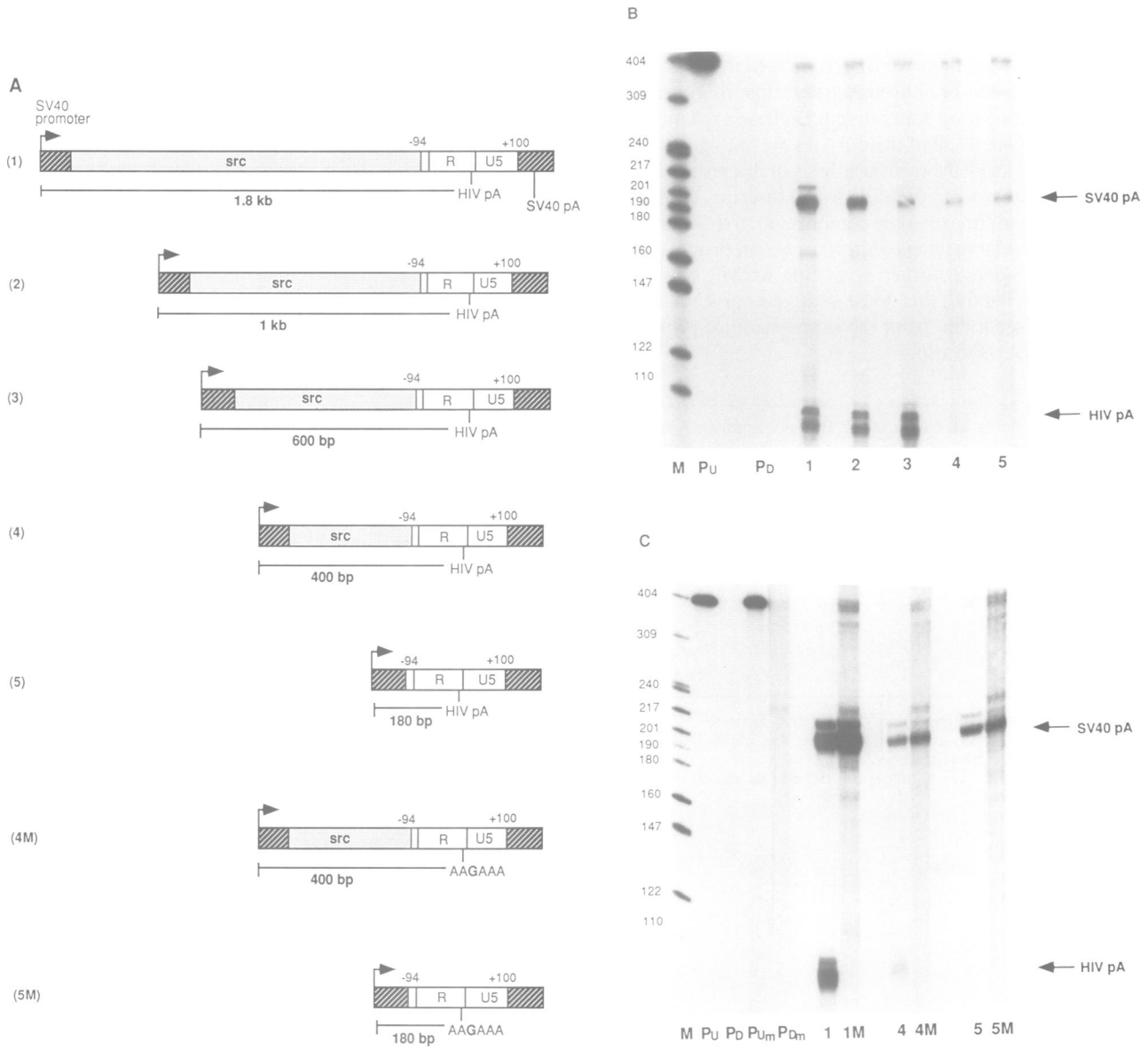

Fig. 5. Cap site proximity affects processing efficiency at the 5′ HIV pA signal. (**A**) In order to decrease the distance between the SV40 cap site and the HIV pA signal, deletions were made within the *src* sequences of src. −94HIV/SV (construct 1), to produce derivatives 2, 3, 4 and 5. The approximate distance which remains between the cap site and the pA signal in each construct is shown. Constructs 1M, 4M and 5M contain a hexanucleotide mutation in constructs 1, 4 and 5, respectively. (**B**) RNase protection analysis of cytoplasmic RNA harvested from COS cells transfected with constructs 1−5. A uniformly labeled RNA probe described in Figure 1A was used in this analysis. (**C**) RNase protection analysis of cytoplasmic RNA harvested from COS cells transfected with constructs 1, 1M, 4, 4M, 5 and 5M. The probe described in Figure 1A was used for the analysis of RNA from constructs 1, 4 and 5. For RNA from constructs 1M, 4M and 5M, an identical probe was used, save for the presence of a mutation in the hexanucleotide from AATAAA to AAGAAA. Predicted products of 110 nt representing transcripts processed at the HIV pA signal, and 195 nt representing readthrough transcripts are shown by arrows to the right. Lane numbers correspond to the numbered constructs shown in (A). Mol. wt markers shown to the left are pBR322 DNA digested with *Msp*I. PU denotes wild-type undigested probe, PD denotes wild-type digested probe, PU_m denotes undigested probe containing a hexanucleotide mutation and PD_m denotes digested probe containing a hexanucleotide mutation.

efficient processing (65%) in agreement with our Northern analysis (Figure 2, lane 4). However, at a distance of 400 bp (Figure 5B, lane 4) or 180 bp (lane 5) from the promoter, the HIV poly(A) signal is no longer utilized efficiently: processing is reduced to 40% at 400 bp while at 180 bp < 10% of the transcripts are processed at the HIV pA signal (cf. Table I). Identical results were also obtained when the HIV sequences included in these clones extended to −55 rather than −94 (data not shown).

Inspection of Figure 5B reveals another striking feature: in addition to the shift from processing to readthrough at the HIV poly(A) site, there is a dramatic reduction in the total level of cytoplasmic *src* RNA recovered from cells transfected with constructs in which the cap site is within 400−500 nt from the pA site. Evidence of this phenomenon can be found in other studies of this type (Brown *et al.*, 1991; Weichs an der Glon *et al.*, 1991), but has previously been assumed to be due to RNA instability. However, this suggestion has not been experimentally tested. We think it highly unlikely that RNA instability alone accounts for this, as we have observed the identical phenomenon for a large number of poly(A) signals (HIV, duck hepatitis B virus,

ground squirrel hepatitis virus and the β-globin signal SPA) when closely approximated to the SV40 early promoter. In support of this contention we directly assayed the stability of the processed and readthrough transcripts of constructs 2 and 5 of Figure 5A by actinomycin D chase experiments. As shown in Figure 6, all of these RNAs are similarly stable.

We do not yet know the molecular basis of this phenomenon; it could be due to suppression of transcript initiation, elongation or transport, or some combination of these effects. Studies to define its mechanism further are in progress and will not be considered further here. But, as will be shown below (Figures 7 and 8) this decrease in overall RNA levels is distinct and separable from the suppression of promoter proximal poly(A) site use.

Table I. Polyadenylation efficiency of the HIV pA signal carrying or not carrying U3 sequences, as a function of the distance from the cap site

Cap site−pA distance (bp)	HIV processing efficiency	
	−U3 (%)	+U3 (%)
1800−2000	60	>95
800−1200	65	>95
600	70	>95
400	40	>95
180	<10	N/A

These results represent the average of at least five separate experiments. As the deviation was never >5%, the results have been rounded to the nearest 5th percentile.
N/A, not applicable—the U3 region is too large to be placed within 180 nt of the pA signal.

To rule out the possibility that the short transcripts expected from processing at the HIV pA signal in constructs 4 and 5 of Figure 5A might be unstable relative to the readthrough transcript, we proceeded as follows. If use of the HIV poly(A) site was efficient in these clones but the resulting short RNA was unstable, then mutational inactivation of the HIV pA signal should result in a significant increase in the quantity of readthrough RNA compared with that generated by the wild-type clone. Accordingly, we introduced a hexanucleotide mutation (AATAAA to AAGAAA) into constructs 1, 4 and 5 (Figure 5A) to generate constructs 1M, 4M and 5M (Figure 5A). These mutant plasmids and their wild-type parents were transfected into COS cells; 48 h later cytoplasmic RNA was examined by RNase protection analysis using a probe containing wild-type HIV sequences from −219 to +100 for constructs 1, 4 and 5 and the same probe but containing the hexanucleotide mutation for the analysis of constructs 1M, 4M and 5M (Figure 5C). As before, lanes 1, 4 and 5 show that construct 1 directs efficient processing whereas constructs 4 and 5 (at a distance of <400 nt from the cap site) do not. Lanes 1M, 4M and 5M show the results of the protection analysis using RNA generated from constructs 1M, 4M and 5M. First, if both transcripts are equally stable, then the intensity of the band representing the readthrough transcript in lane 1M should equal the sum of the intensities of the two transcripts present in lane 1; phosphorimager quantitation reveals this to be the case. Secondly, the hexanucleotide mutations in clones 4M and 5M did not yield an increase in levels of readthrough RNA (compare lanes 4 and 4M or 5 and 5M). Therefore, the observed shift to readthrough at a distance

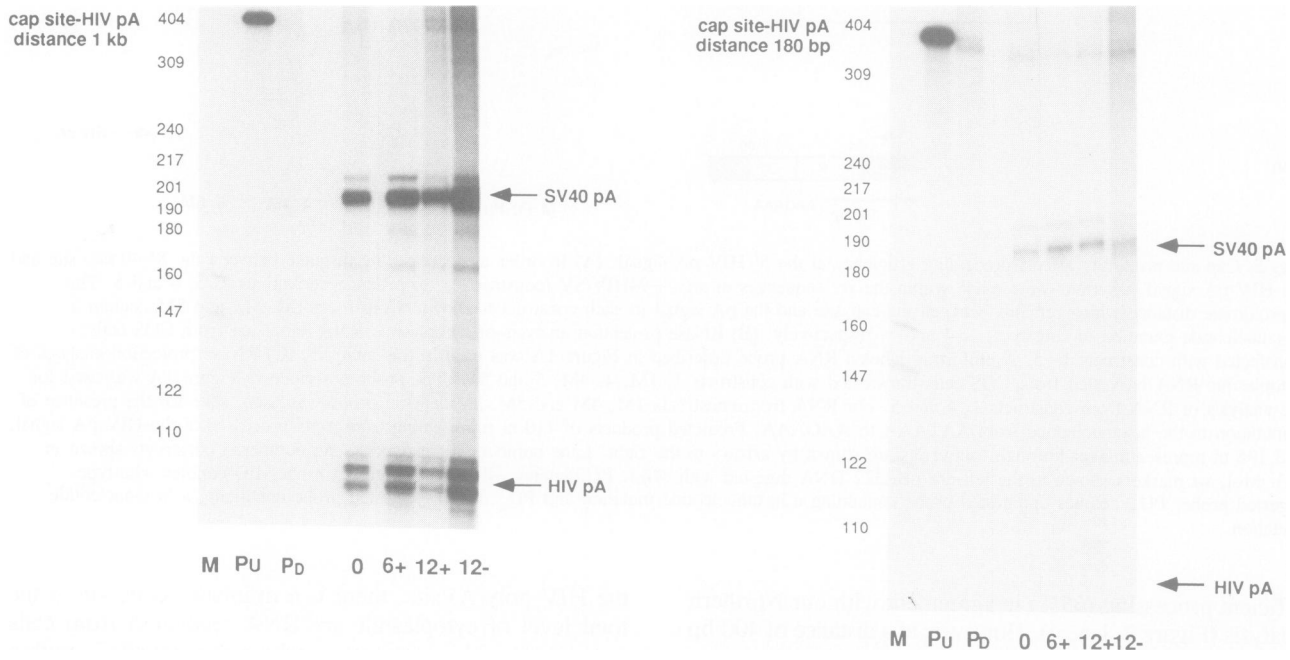

Fig. 6. Stability analysis of RNA transcripts polyadenylated at different distances from the cap site. COS cells were transfected with a construct containing 1 kb of *src* sequence between the cap site and pA signal (left panel; construct 2 of Figure 5A) and a construct containing 180 bp of *src* sequence between the cap site and pA signal (right panel; construct 5 of Figure 5A). At 48 h post-transfection, cells were either treated or not treated with actinomycin D. RNA harvested from cells at 48 h or 60 h post-transfection which did not receive actinomycin D treatment are depicted (0) and (12−), respectively. RNA harvested from cells at 6 or 12 h post-actinomycin D treatment are depicted (6+) and (12+), respectively. The probe described in Figure 1A was used for the analysis of RNA. Predicted products of 110 nt representing transcripts processed at the HIV pA signal, and 195 nt representing readthrough transcripts are shown by arrows to the right. Mol. wt markers shown to the left are pBR322 DNA digested with *Msp*I. PU denotes undigested probe, PD denotes digested probe.

of <400 bp from the cap site is not due to the instability of the short transcript but reflects a true suppression of processing at the HIV pA signal.

Not all poly(A) signals are equally sensitive to suppression by promoter proximity

We next asked whether any poly(A) signal would be suppressed when positioned near the 5′ end of a transcript, or whether this property is a special feature of the HIV signal To do so we constructed a series of plasmids that contain an unrelated poly(A) signal (the β-globin signal SPA) placed at varying distances from the promoter (Figure 7A). These clones were transfected into COS cells, cytoplasmic RNA was harvested 48 h post-transfection and subjected to RNase protection analysis. An RNA species of 120 nt (Figure 7; lanes 1 and 2) or 95 nt (lanes 3, 4 and 5) represents transcripts processed at the SPA site, while a protected fragment of 300 nt (lanes 1 and 2) or 275 nt (lanes 3, 4 and 5) represents readthrough transcripts. As shown in Figure 7B, >95% of the transcripts are processed at the SPA signal even when the cap site is placed only 300 bp from the promoter. The SPA signal continues to show efficient processing (~70%) when it is only 140 nt from the cap site, a distance at which the HIV signal is completely suppressed (lane 5). Thus, the susceptibility of poly(A) signals to suppression by cap site proximity varies, and depends upon some feature of the processing signal itself. A similar conclusion has also been reached by Weichs an der Glon *et al.* (1991), who have suggested that variation in the GU-rich regions may be responsible for such differences.

When the promoter is 300 nt or less from the SPA signal, a drastic decrease in the accumulation of all stable transcripts from the constructs is once again observed (lanes 4 and 5). This indicates that the attenuation of stable RNA accumulation is separable from the suppression of poly(A) site use: the former seems to be invariable when promoters and poly(A) signals are approximated, while the latter depends upon the nature of the poly(A) signal.

U3 sequences reduce the sensitivity of the HIV poly(A) signal to suppression by cap site proximity

One feature shared by the SPA signal and the U3-augmented HIV poly(A) signal is the extremely high efficiency with which both are used (Figure 2B, lane 4 and Figure 3B, lane 1). Once the relative resistance of the SPA site to suppression by cap proximity became clear, we wondered whether the U3-augmented HIV signal would behave similarly. To explore this, the series of plasmids depicted in Figure 8A was constructed. In these constructs the 5′ endpoint of HIV sequences was −317 relative to the HIV pA signal (this includes approximately two-thirds of the proviral U3 region). These HIV sequences were positioned behind variable lengths of *src* DNA so that the cap site−pA site distances ranged from 400 to 2000 nt. These constructs were transfected into COS cells and cytoplasmic RNA was harvested and subjected to RNase protection analysis as described above. In the experiment shown in Figure 8, a 240 nt protected species represents transcripts polyadenylated at the HIV poly(A) signal, while readthrough transcripts protect a 320 nt fragment. As expected, when positioned 600−2000 nt from the cap site these HIV sequences direct processing with >95% efficiency (Figure 8B, lanes 1−4). However, in contrast to what was observed with either the src.HIV−94/SV deletion series (Figure 5B) or the

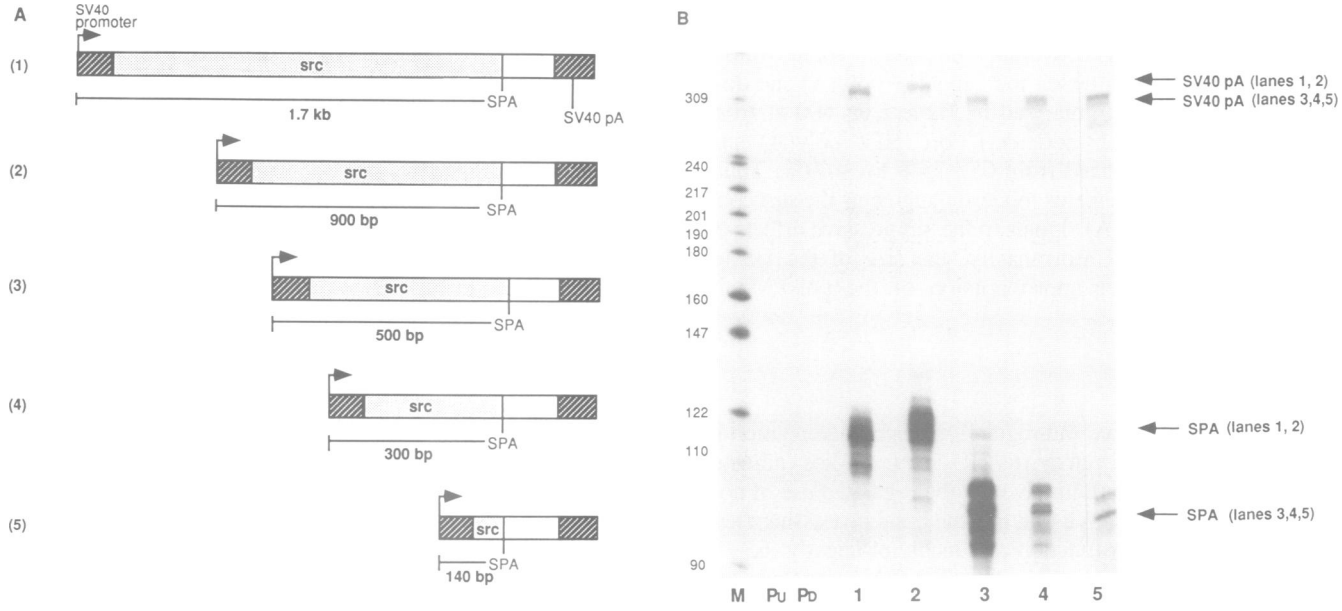

Fig. 7. A synthetic poly(A) site (SPA) is insensitive to suppression by cap site proximity. (**A**) Deletions were made in src.SPA (construct 1), to produce derivatives which decreased the distance between the SV40 promoter and the SPA signal. The approximate distance which remains between the cap site and pA signal is noted for each construct. (**B**) RNase protection analysis of cytoplasmic RNA harvested from COS cells transfected with constructs 1−5. The probe used in this analysis contained sequences at its 3′ end which are not present in the RNA from constructs 3, 4 and 5; therefore, the protected fragments are slightly larger in lanes 1 and 2 as compared with those in lanes 3, 4 and 5. Predicted products of 120 nt (lanes 1 and 2) or 95 nt (lanes 3, 4 and 5) representing transcripts processed at the HIV pA signal, and 285 nt (lanes 1 and 2) or 300 nt (lanes 3, 4 and 5) representing readthrough transcripts are shown by arrows to the right. Lane numbers correspond to the numbered constructs shown in (A). Mol. wt markers shown to the left are pBR322 DNA digested with *Msp*I. PU denotes undigested probe while PD denotes digested probe.

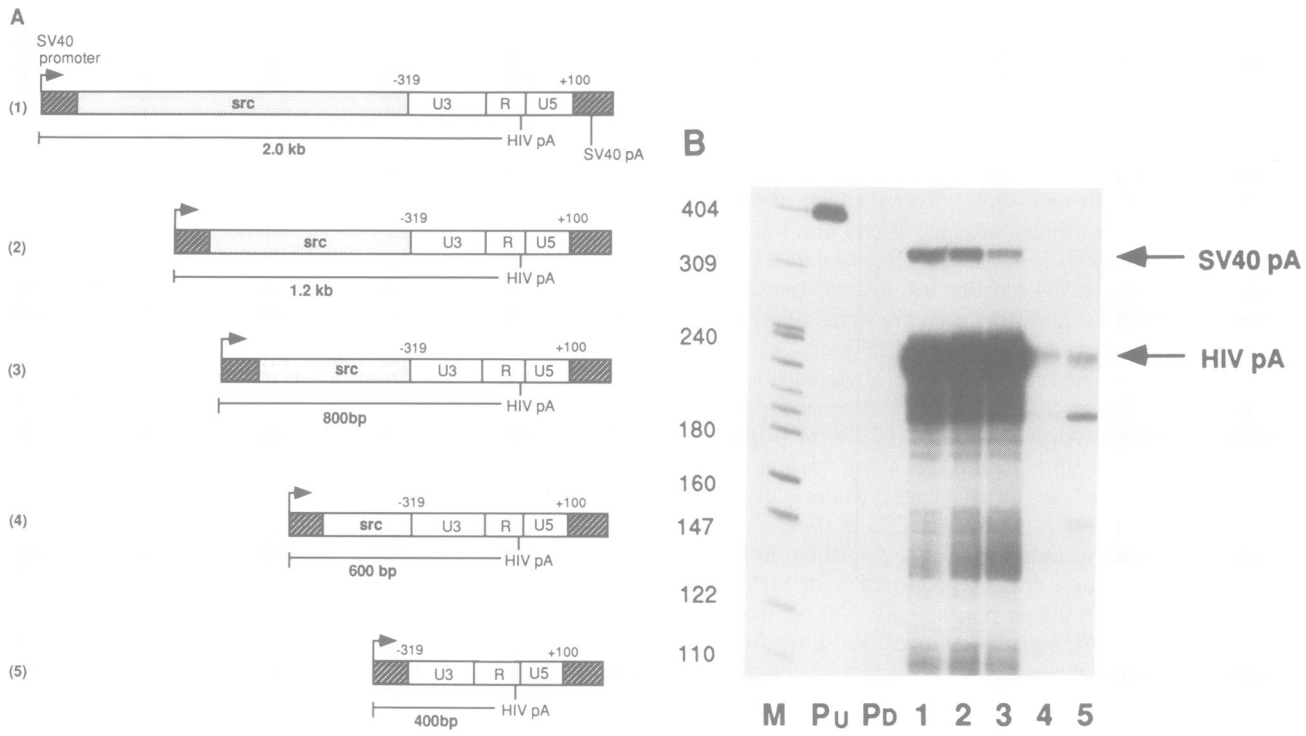

Fig. 8. The U3-augmented HIV pA site displays reduced sensitivity to suppression by cap site proximity. (**A**) Deletions were made in src.−317/SV (construct 1) to produce derivatives (constructs 2, 3, 4 and 5) which decreased the distance between the SV40 promoter and HIV pA signal. The approximate distance which remains between the cap site and pA signal is noted for each construct. (**B**) RNase protection analysis of cytoplasmic RNA harvested from COS cells transfected with constructs 1−5. The probe described in Figure 1A legend was used for this analysis. Predicted products of 240 nt representing transcripts processed at the HIV pA signal, and 320 nt representing readthrough transcripts are shown by arrows to the right. Lane numbers correspond to the numbered constructs shown in (A). Mol. wt markers shown to the left are pBR322 DNA digested with *Msp*I. PU denotes undigested probe while PD denotes digested probe.

src.HIV−55/SV deletion series (data not shown), when the cap site was brought to within 400 nt of the (U3-augmented) HIV poly(A) signal, efficient processing still occurred (lane 5). (Note again the overall reduction in stable transcripts from plasmids in which the cap site−poly(A) site distance is <600.) As summarized in Table I, at 400 nt from the promoter the (U3-deficient) core HIV poly(A) signal's efficiency is reduced from 65−75% to ~40%. Thus, the absence of U3 elements modestly augments the susceptibility of the core poly(A) signals to the suppressive effects of cap site proximity. Unfortunately, the size of the U3 region precluded further approximation of the poly(A) signals to the promoter.

Discussion

We have examined the regulation of polyadenylation in the AIDS-associated retrovirus HIV. Our studies indicate that several features are involved in the regulated use of poly(A) sites in this virus. We find that the basal processing elements of the viral RNA function reasonably efficiently (65−75%) when placed at a distance from the RNA start site (i.e. in a position mimicking that of the 3′ LTR); this efficiency can be enhanced to nearly 100% by the addition of upstream U3 sequences to the transcript. In our hands the intrinsic efficiency of the basal HIV processing signals is too great to account for the high level of readthrough seen at the 5′ LTR in proviral DNA, as would be called for by model 3 (in which upstream sequences increase processing efficiency). At the 5′ LTR, these signals must be actively suppressed. Experiments reported here and elsewhere (Weichs an der Glon *et al.*, 1991) indicate that cap site proximity is largely responsible for this suppression. The absence of U3 sequences at the 5′ poly(A) signal may contribute to this secondarily by modestly enhancing the sensitivity of the core signals to distance-related suppression. But in our view, the major role of U3 polyadenylation elements in the viral life cycle is to enhance the processing efficiency at the 3′ LTR.

Other groups have suggested that model 3 (upstream sequences) may be sufficient to explain the processing of HIV RNA (Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991). However, in these reports events at the 5′ LTR have not been examined directly. In all these cases, estimates of the intrinsic processing efficiency of the basal HIV poly(A) signals have yielded values considerably lower than that reported here (in the range of 10%). At present we do not understand the basis of these differing estimates. As mentioned previously, DeZazzo *et al.* (1991) noted that the magnitude of their estimate varied 6- to 7-fold depending on the presence or absence of a downstream poly(A) signal, in this case from adenovirus. We observe no comparable disparity between the one and two poly(A) site clones we constructed. Perhaps some feature of the adenoviral processing signal accounts for this difference. The estimates we obtain for HIV core element processing efficiency agree well with similar assessments of Proudfoot and colleagues (Weichs an der Glon *et al.*, 1991) and with those of Iwasaki and Temin (1990) for the avian retrovirus SNV.

Despite these differences there is wide agreement that the U3 region does contain sequences that upregulate polyadenylation at both homologous and heterologous poly(A) signals (Russnak and Ganem, 1990; Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991). Upstream elements which increase processing efficiency have also been described in a number of other viral systems (Carswell and Alwine, 1989; DeZazzo and Imperiale, 1989; Russnak and Ganem, 1990; Sanfacon *et al.*, 1991) and are likely to exist in cellular genes as well. Data presented in this paper demonstrate that sequences between −94 and −219 (125 bp of U3) increase the processing efficiency of the HIV pA signal from 65 to >95%. Valsamakis *et al.* (1991) additionally reported that a linker insertion mutation which disrupted sequences from −77 to −94 (in our numbering system) significantly decreased processing efficiency. In our constructs, we did not observe any effect of these sequences in regulating the processing efficiency of the HIV poly(A) signal (see Figure 2, lanes 2 and 3).

How does cap site proximity influence poly(A) site use? At present we can say little about this with certainty. Perhaps a leader RNA of a certain length (>400−500 nt) is required to allow recruitment of the appropriate protein factors involved in cleavage and polyadenylation. If so, however, then certain poly(A) signals must be able to escape this requirement, either by promoting alternative secondary structures more favorable to such recruitment or by directly recruiting additional stimulatory factors. A mechanistic understanding of both the activation of poly(A) signals by upstream elements and their suppression by cap site proximity will doubtless require development of *in vitro* systems in which these phenomena can be faithfully reproduced.

Materials and methods

Plasmid constructions

The HIV proviral clone HXB2 was a gift from Kathleen Page. HIV-gpt has been described previously (Page *et al.*, 1990).

SrcΔs was derived from the plasmid, src-8. The *src* sequences present in src-8 were derived from pGEM7Zsrc as a 1.6 kb *Hin*dIII−*Xho*I fragment. pGEM7Zsrc contains the 1.6 kb *Eco*RI−*Xho*I *src* cDNA fragment from pGC1-src (Hirai and Varmus, 1990). The 1.6 kb *Hin*dIII−*Xho*I fragment was cloned into *Hin*dIII−*Xho*I digested pGSpA.wt (Russnak and Ganem, 1990) which carried a deletion of ground squirrel hepatitis virus (GSHV) sequences from −34 to −9 relative to the GSHV pA signal. SrcΔs was constructed by deleting the 800 bp *Sal*I fragment of src-8 which removed all of the GSHV sequences from the plasmid leaving a construct which contained the SV40 promoter driving c-*src* cDNA, a multicloning site (MCS) downstream of *src* and an SV40 pA signal 150 bp downstream of the MCS.

The src.HIV/SV series (Figure 2) was constructed by cloning the HIV poly(A) signal containing varying amounts of upstream DNA (filled in with Klenow) derived from the provirus, HXB2 (gift from K.Page), into the unique *Sal*I site (filled in with Klenow) present in the MCS downstream of *src*. Src.HIV−9/SV contained the *Afl*II−*Nar* fragment, src.HIV−40/SV contained the *Sac*I−*Nar* fragment, src.HIV−55/SV contained the *Bgl*II−*Nar* fragment, src.HIV−94/SV contained the *Pvu*II−*Nar* fragment, src.-HIV−219/SV contained the *Bsp*EI−*Nar* fragment, and src.HIV−317/SV contained the *Dra*III−*Nar* fragment of HXB2. src.HIV-219.inv was constructed by inserting, in opposite orientation, the filled in 125 bp *Bam*HI−*Pvu*II fragment from src.HIV−219/SV (*Bam*HI is present in MCS upstream of *Sal*I) into src.HIV−94/SV which had been digested with *Bam*HI and filled in with Klenow.

src.SPA/SV was constructed by digesting src-8 with *Bcl*I and *Bam*HI (−20 to +240 relative to the GSHV pA signal), filling in the ends with Klenow polymerase and inserting the synthetic pA signal (SPA signal) derived from β-globin (Levitt *et al.*, 1989) as a double stranded, filled in, 55 bp synthetic oligonucleotide; the sequence of which is 5′-TCGAGAATAAAA-GATCTTTATTTTCATTAGAATCTGTGTGTTGGTTTTTTGTGTGT-3′ and 5′-GATCACACACAAAAAACCAACACACAGATCTAATGAAAA-TAAAGATCTTTTATTC-3′. src.SPA was constructed by deleting the downstream SV40 pA signal of src.SPA/SV as a *Sac*I−*Bsp*EI collapse. The src.HIV/SPA series was constructed by cloning the HIV pA signal fragments (*Afl*III−*Nar*, *Pvu*II−*Nar* and *Bsp*EI−*Nar* from HXB2) into the unique *Mlu*I site of src.SPA. The inserts were derived from the src.HIV/SV series constructs as *Xba*I fragments (*Xba*I sites flank the *Sal*I site in srcΔs). Both the vector and inserts were filled in with Klenow before ligation.

The src.HIV series was constructed by cloning the HIV pA signal fragments (*Afl*II−*Nar*, *Bgl*II−*Nar*, *Pvu*II−*Nar* and *Bsp*EI−*Nar*) derived from the src.HIV/SV series constructs as *Xba*I fragments (*Xba*I sites flank the *Sal*I site in srcΔs) into the unique *Xba*I site (present in MCS downstream of src) of srcΔs/ΔSV40. SrcΔs/ΔSV40 was constructed by deleting the SV40 signal from the plasmid, srcΔs. This was done by deleting the *Sal*I−*Sac*I fragment (resulting in the loss of both sites but leaving *Bam*HI and *Bgl*II site in the MCS downstream of *src*).

src.HIV−94M/SV and src.HIV−219M/SV were constructed using site directed mutagenesis (Kunkel *et al.*, 1987). The HIV *Pvu*II−*Nar* and *Bsp*EI−*Nar* fragments were purified from src.HIV−94/SV and src.HIV−219/SV respectively, as *Xba*I fragments. These fragments were each cloned into the *Xba* site of mp19 and subjected to site directed mutagenesis (Kunkel *et al.*, 1987) using the oligo 5′-GGCAAGCTTTCT-TGAGGCTT-3′ resulting in an AATAAA to AAGAAA change.

Constructs used in the promoter proximity experiments resulted from deletions made within *src* sequences and the 5′ and 3′ flanking polylinker sequences of construct 1 of a particular series. In all cases, digests with the appropriate enzyme(s) were completed, ends were repaired if necessary and the plasmid was recircularized. The src.HIV−94/SV wild-type and hexanucleotide mutation deletion series (Figure 5) and the src.HIV−317/SV deletion series (Figure 8) were constructed in the following way, starting with src.HIV−94/SV or src.HIV−317/SV respectively: construct 2 resulted from an *Mlu*I−*Kpn*I collapse, construct 3 resulted from a *Pst*I collapse, construct 4 resulted from a *Pst*I collapse of construct 2, and construct 5 resulted from a *Bam*HI−*Kpn*I collapse. The src.SPA/SV deletion series (Figure 7) constructs 2, 3 and 4 were constructed by making the same deletions in src.SPA/SVS as were made in src.HIV/SV; construct 5 resulted from an *Oxa*NI−*Kpn*I collapse of src.SPA/SV.

pG.src was constructed by cloning the 1.6 kb *Hin*dIII−*Bam*HI fragment containing the *src* cDNA sequences from src-8 into pGEM3Z digested with *Hin*dIII and *Bam*HI.

pG.5′src was constructed by making an *Mlu*I−*Bam*HI collapse in pG.src thereby deleting the 3′ half of *src* (nt 884−1712).

pG.HIV was constructed by cloning the 319 bp *Xba* fragment from src.HIV−219M/SV (containing HIV sequences from −219 to +100 relative to the hexanucleotide) into pGEM7Zf+ digested with *Xba*I.

pG.HIV_M was constructed by cloning the 320 bp *Xba* fragment from src.HIV−219M/SV (containing HIV sequences from −219 to +100 relative to the mutant hexanucleotide) into pGEM7Zf+ digested with *Xba*.

Cell culture and transfections

$CD4^+$ HeLa cells bearing the integrated provirus HIV-gpt were the kind gift of Drs K.Page and D.Littmann (UCSF). They were derived by infection of HeLa-CD4 cells with helper-free stocks of the HIV vector HIV-gpt; after selection for gpt in mycophenolic acid, 50−100 gpt^+ colonies were pooled and carried as a mass culture (Page *et al.*, 1990). HeLa and COS7 cells were maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 10% fetal calf serum. Transfection was carried out with 2 μg of plasmid DNA per 100 mm dish in the presence of 500 μg/ml DEAE−dextran (Sompayrac and Danna, 1981). Where indicated, actinomycin D (10 μg/ml final concentration in DME supplemented with 10% fetal calf serum) was added to cells 48 h post-transfection for the appropriate period of time.

RNA analysis

Total RNA was isolated from $CD4^+$ HeLa cells bearing the integrated provirus, HIV-gpt, according to the procedure of Chirgwin *et al.* (1979). Cytoplasmic RNA was isolated 48 h post-transfection from tissue culture cells according to the protocol of Kaufman and Sharp (1982). $Poly(A)^+$ RNA was selected using oligo(dT) cellulose as described by Hirsch *et al.* (1988). Northern analysis and RNase protection analysis were carried out as described by Russnak and Ganem (1990).

An RNA probe generated from pGsrc linearized at *Mlu*I and transcribed *in vitro* with T7 RNA polymerase and [^{32}P]UTP (Melton *et al.*, 1984) was used for the Northern analysis of the *src* constructs (Figures 2 and 3A). An RNA probe generated from pG.5′src linearized at *Hin*dIII was transcribed *in vitro* as described above and used for Northern analysis of the src.HIV/SPA constructs (Figure 3B). An RNA probe generated from

pG.HIV linearized at *Eco*RI was transcribed *in vitro* as described above (except that both [^{32}P]UTP and [^{32}P]CTP were used) and used for RNase protection analyses of RNA derived from transfection of constructs containing a wild-type hexanucleotide (Figures 1, 4, 5B, 6 and 8). RNase protection analysis of RNA derived from transfection of constructs containing a mutant hexanucleotide was conducted using an RNA probe generated from pG.HIV$_M$, linearized with *Eco*RI and transcribed *in vitro* as described for pG.HIV (Figure 5C). An RNA probe generated from src.SPA linearized at *Oxa*NI was transcribed *in vitro* (with both [^{32}P]UTP and [^{32}P]CTP) with SP6 polymerase and used for the RNase protection analysis of the src.SPA/SV promoter proximity constructs (Figure 7). In all of the RNase protection analyses, the calculations of processing efficiency included a correction factor which accounted for the U and C content of the probe.

Acknowledgements

We thank Drs E.Blackburn and H.E.Varmus for critical reviews of the manuscript. J.M.C. was supported by an American Cancer Society Postdoctoral Fellowship.

References

Acheson,N. (1984) *Mol. Cell. Biol.*, **4**, 722–729.
Bohnlein,S., Hauber,J. and Cullen,B. (1989) *J. Virol.*, **63**, 421–424.
Brown,P.H., Tiley,L.S. and Cullen,B.R. (1991) *J. Virol.*, **65**, 3340–3343.
Carswell,S. and Alwine,J.C. (1989) *Mol. Cell. Biol.*, **9**, 4248–4258.
Chirgwin,J.M., Przbyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Conway,L. and Wickens,M.P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3949–3953.
Conway,L. and Wickens,M. (1987) *EMBO J.*, **6**, 4177–4184.
DeZazzo,J.D. and Imperiale,M.J. (1989) *Mol. Cell. Biol.*, **9**, 4951–4961.
DeZazzo,J.D., Kilpatrick,J.E. and Imperiale,M.J. (1991) *Mol. Cell. Biol.*, **11**, 1624–1630.
Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251–260.
Geballe,A., Spaete,R.R. and Mocarski,E.S. (1986) *Cell*, **46**, 865–872.
Gil,A. and Proudfoot,N.J. (1984) *Nature*, **312**, 473–474.
Green,T.L. and Hart,R.P. (1988) *Mol. Cell. Biol.*, **8**, 1839–1841.
Greene,W. (1990) *Annu. Rev. Immunol.*, **8**, 453–475.
Hart,R.P., McDevitt,M.A., Ali,H. and Nevins,J. (1985) *Mol. Cell. Biol.*, **5**, 2975–2983.
Higgs,D.R., Goodburn,S., Lamb,J.B., Weatherall,D.J. and Proudfoot,N.J. (1983) *Nature*, **306**, 398–400.
Hirai,H. and Varmus,H.E. (1990) *Mol. Cell. Biol.*, **10**, 1307–1318.
Hirsch,R.C., Colgrove,R. and Ganem,D. (1988) *Virology*, **167**, 136–142.
Iwasaki,K. and Temin,H.M. (1990) *Genes Dev.*, **4**, 2299–2307.
Kaufmann,R.J. and Sharp.P.A. (1982) *Mol. Cell. Biol.*, **2**, 1303–1319.
Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enyzmol.*, **15**, 376.
Levitt,N., Briggs,D., Gil,A. and Proudfoot,N.J. (1989) *Genes Dev.*, **3**, 1019–1025.
McDevitt,M.A., Hart,R.P., Ali,H. and Nevins,J. (1984) *Cell*, **37**, 993–999.
McLaughlan,J., Gaffney,D., Whiton,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1368.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Montell,C., Fischer,E.F., Caruthers,M.H. and Berk,A.J. (1983) *Nature*, **305**, 600–605.
Page,K.A., Laundau,N.R. and Littman,D.R. (1990) *J. Virol.*, **64**, 5270–5276.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
Russnak,R. and Ganem,D. (1990) *Genes Dev.*, **4**, 764–776.
Sadofsky,M. and Alwine,J.C. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3949–3953.
Sanfacon,H. and Hohn,T. (1990) *Nature*, **346**, 81–84.
Sanfacon,H., Brodman,P. and Hohn,T. (1991) *Genes Dev.*, **5**, 141–149.
Sompayrac,L.M. and Danna,K.J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7575–7578.
Valsamakis,A., Zeichner,S., Carswell,S. and Alwine,J.C. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2108–2112.
Weichs an der Glon,C., Monks,J. and Proudfoot,N.J. (1991) *Genes Dev.*, **5**, 244–253.
Wickens,M. and Stephenson,P. (1984) *Science*, **226**, 1045–1051.
Zarkower,D. and Wickens,M. (1988) *J. Biol. Chem.*, **263**, 5780–5788.
Zarkower,D., Stephenson,P., Sheets,M. and Wickens,M. (1986) *Mol. Cell. Biol.*, **6**, 2317–2323.

Received on August 28, 1991; revised on December 12, 1991