The EMBO Journal vol.12 no.5 pp.2119-2128, 1993

Tat-dependent occlusion of the HIV poly(A) site

Caroline Weichs an der Glon, Mark Ashe, Jan Eggermont and Nick J.Proudfoot[1]

Sir William Dunn School of Pathology, South Parks Road, Oxford OX1 3RE, UK
[1]Corresponding author

Communicated by N.J.Proudfoot

Retroviruses must ensure that poly(A) signals in the 3′ LTR are highly active, while identical signals in the 5′ LTR are inactive (occluded). In the case of HIV-1, both promoter proximity in the 5′ LTR and U3 sequences in the 3′ LTR may contribute to this regulation. We have discovered a novel regulatory mechanism for poly(A) site occlusion in HIV-1. When transcription initiation from the HIV promoter is activated by Tat, the HIV poly(A) site is specifically occluded, while other poly(A) sites are unaffected by Tat. Nucleotide signals associated with this Tat effect are immediately adjacent to the AAUAAA sequence of the HIV-1 poly(A) signal. These data suggest that elongating RNA polymerase II, activated by Tat specifically occludes the HIV poly(A) site.

Key words: HIV-1/poly(A) site occlusion/RNA polymerase II/Tat

Introduction

The polyadenylation of transcripts synthesized by RNA polymerase II (Pol II) is a key step in the maturation of translation competent mRNA (Proudfoot and Whitelaw, 1989; Jackson and Standart, 1990). In chromosomal genes this 3′ end processing mechanism occurs at varying distances beyond the translation stop codon and defines the 3′ end of the last exon. The signals that specify this polyadenylation process in mammals are relatively simple comprising an AAUAAA sequence 20–30 nucleotides upstream of the cleavage and polyadenylation site, which is immediately followed by a GU- or U-rich sequence that effectively determines their efficiency (Proudfoot, 1991). These polyadenylation signals have a second function in that termination of Pol II transcription in the 3′ flanking region of the gene is linked to 3′ end processing (Logan *et al.*, 1987; Proudfoot and Whitelaw, 1988; Connelly and Manley, 1988; Enriquez-Harris *et al.*, 1991).

In general these polyadenylation and termination mechanisms do not appear to be regulated events in chromosomal genes. In retroviruses, however, the situation is very different. The long terminal repeat (LTR) sequences at either end of the proviral genome are exactly duplicated and contain all of the promoter and poly(A) signals of the virus, necessitating the tight regulation of poly(A) site usage. As shown in Figure 1 the poly(A) signals in the 5′ LTR are inactive (or occluded), while identical signals in the 3′ LTR are fully active. In some retroviruses this differential usage of poly(A) sites is achieved by placing the AAUAAA of the poly(A) signal before the start site of viral transcription (in the U3 sequence), but the cleavage site with adjacent GU-rich sequence after the start site. Thus the only complete poly(A) signals in the transcript will be in the 3′ LTR transcript (Ahmed *et al.*, 1991; for a review see Proudfoot, 1991). However, a number of retroviruses, including the human immunodeficiency virus HIV-1 possess complete poly(A) signals downstream of the transcriptional start site, so that other mechanisms must exist to restrict usage of the poly(A) site to the 3′ LTR region of the viral transcript (Weichs an der Glon *et al.*, 1991). Two mechanisms have been suggested to contribute to this regulatory process. First, both the HIV and spleen necrosis virus (SNV) poly(A) sites appear to be sensitive to inhibition from closely positioned promoters (as found in the 5′ LTR). When the poly(A) signals are moved away from their respective promoters (or heterologous promoters) to a distance of ~500 nucleotides or more, they become activated (Iwasaki and Temin, 1990, 1992; Weichs an der Glon *et al.*, 1991). However, in addition to this promoter proximity effect, several studies have demonstrated that U3 sequences also significantly activate the HIV poly(A) site. In some studies this U3 effect appears to be the major factor in activating the HIV poly(A) site (Brown *et al.*, 1991; DeZazzo *et al.*, 1991; Valsamakis *et al.*, 1991), while in a recent study comparing U3 and promoter proximity effects, promoter proximity appeared to be predominant (Cherrington and Ganem, 1992).

An additional feature of the HIV LTR sequence between the cap and poly(A) sites of the viral transcript (R, see Figure 1) is the TAR (*t*rans-*a*ctivating *r*esponse element) RNA stem–loop sequence to which the transactivator Tat binds (Berkhout *et al.*, 1989; Dingwall *et al.*, 1989). The close proximity of TAR to the immediately adjacent poly(A) signal may reflect interactions between these two processes. The Tat–TAR interaction has been shown to activate transcription at the HIV promoter [see Sharp and Marciniak (1989) and Cullen (1990) for reviews]. In the absence of Tat the HIV promoter initiates transcripts that fail to elongate efficiently. Tat binds to these attenuated transcripts and activates HIV transcription (Kao *et al.*, 1987; Laspia *et al.*, 1989, 1990). While some experiments argue that this stimulation of HIV transcription occurs at initiation (Hauber *et al.*, 1987; Southgate and Green, 1991), other results point to a direct effect on transcriptional elongation (Laspia *et al.*, 1990; Feinberg *et al.*, 1991; Marciniak and Sharp, 1991; for a review see Frankel, 1992; Kato *et al.*, 1992).

The results presented in this paper describe further analysis of the mechanism of poly(A) site regulation in HIV-1. In our previous studies (Weichs an der Glon *et al.*, 1991) the occlusion of the HIV poly(A) site when in its *in vivo* position 80 nucleotides downstream of the transcription start site was investigated. We found that the HIV poly(A) site, in contrast to either the human α globin or a synthetic poly(A) site (SPA), was efficiently read through resulting in full levels

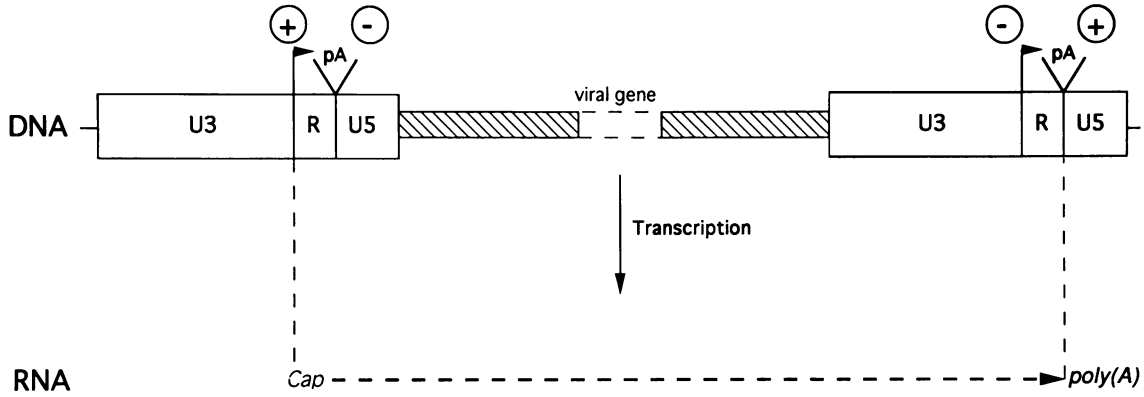

Fig. 1. Diagram showing the structure of the LTR sequences of retroviruses, such as HIV-1. The positions of the U3, R and U5 sequences are indicated and the activity or inactivity of the duplicated promoters and poly(A) signals are indicated by + and −.

of transcripts that utilized a second poly(A) site further downstream in the gene construct. Since we were unable to directly measure utilization of the HIV, α or SPA poly(A) sites close to the HIV promoter (presumably because such truncated mRNAs are unstable), we relied on the level of read-through RNA as indirect measure of poly(A) site usage upstream. Here we present complementary studies showing the direct analysis of the HIV poly(A) site positioned some 450 nucleotides downstream of the HIV promoter. At this position the HIV poly(A) site is partly occluded, while other non-viral poly(A) sites are fully active. More significantly we show that when the HIV promoter is transactivated by Tat, near complete occlusion of the HIV poly(A) site occurs while other poly(A) sites are unaffected. This Tat effect requires the TAR sequence adjacent to the HIV promoter and is associated with sequences immediately adjacent to the AAUAAA sequence. We further show that U3 sequences fully activate the HIV poly(A) site when it is otherwise close enough to the promoter to be affected by promoter proximity and that this U3 effect completely blocks the Tat occlusion effect. These results demonstrate that the HIV poly(A) site can be occluded both by promoter proximity and Tat effects and that these mechanisms are antagonized by U3 sequences. We suggest that all three of these mechanisms may operate in the intact HIV-1 provirus to specify the usage of the viral poly(A) site to the 3′ but not the 5′ LTR sequence.

Results

Tat activation of the HIV promoter causes specific read-through of the HIV poly(A) site

Our previous studies on the HIV poly(A) site involved the analysis of gene constructs (transiently transfected into HeLa cells) in which transcription was initiated on the HIV promoter and then read through the closely spaced HIV poly(A) site into a downstream reporter α globin gene, producing mRNA that utilized the α globin poly(A) site. We concluded that the HIV poly(A) site was inactive when adjacent to the HIV promoter (in its natural 5′LTR context) due to promoter proximity effects, but very active when moved >500 bp away from an active promoter (Weichs an der Glon *et al.*, 1991). In these experiments the HIV promoter was activated by DNA replication, since we have shown that significant levels of HIV transcription can be induced under these conditions in the absence of Tat transactivation (Proudfoot *et al.*, 1992). We have further investigated the occlusion effect of the HIV poly(A) site in two ways. First, we have moved the HIV poly(A) site 450 bp downstream of the HIV promoter in the HIV−α gene construct (see Figure 2A). This allows us to directly probe for HIV poly(A) site use, since mRNA of this size can be readily detected. Secondly, we have co-transfected the HIV gene constructs with a plasmid that expresses Tat (Adams *et al.*, 1988) to determine whether Tat transactivation influences HIV poly(A) site usage.

Figure 2A shows a diagram of the HIV−α globin gene construct used in these experiments. The full HIV promoter sequence (from −135 to +125 bp) including the NFκB enhancer region, the three SP1 sites and the TATA box are present, but not the more distant negative elements (for review of the HIV promoter see Jones, 1989). Downstream of the transcription start site the TAR sequence is present (except in ΔT, see below), but the poly(A) signals have been deleted to generate ΔA (see Materials and methods). All of the HIV−α gene clones used in these studies have this ΔA deletion except for the ΔTTH construct. This HIV LTR sequence is then fused to the human α2 globin gene in place of its promoter. The HIV−α globin fusion occurs in the middle of intron 1 in the α globin gene. The HIV−α globin gene is present in the transient expression vector pSVod, which contains the SV40 origin of replication sequence (Mellon *et al.*, 1981). On transfection into HeLa cells, together with a control plasmid, RβSVpBR328, which expresses both T antigen to allow DNA replication of pSVod plasmids, as well as rabbit β globin mRNA, which provides a transfection efficiency control (Grosveld *et al.*, 1982), we obtain high levels of HIV−α globin transcripts as the HIV promoter is activated by DNA replication (Proudfoot *et al.*, 1992). Also shown in Figure 2A is the site of insertion of various poly(A) signals (including HIV) into the exon 2 *Tth*111I site (T). Using an end-labelled probe from the *Bst*XI site 50 bp upstream (X), it was possible to distinguish mRNA poly(A) site usage at the *Tth*111I site as opposed to read-through transcripts that utilize the α poly(A) site (detected by the 5′ splice site of intron 2, where the RNA will diverge in sequence from the genomic DNA probe).

Figure 2B shows such an S1 analysis of ΔAHIV−α mRNA with three different poly(A) signals inserted into the exon 2 *Tth*111I site. In each case transfections were carried out with and without a third co-transfected plasmid that expresses the HIV Tat protein (Adams *et al.*, 1988). It is clear from these results that Tat expression increases the level

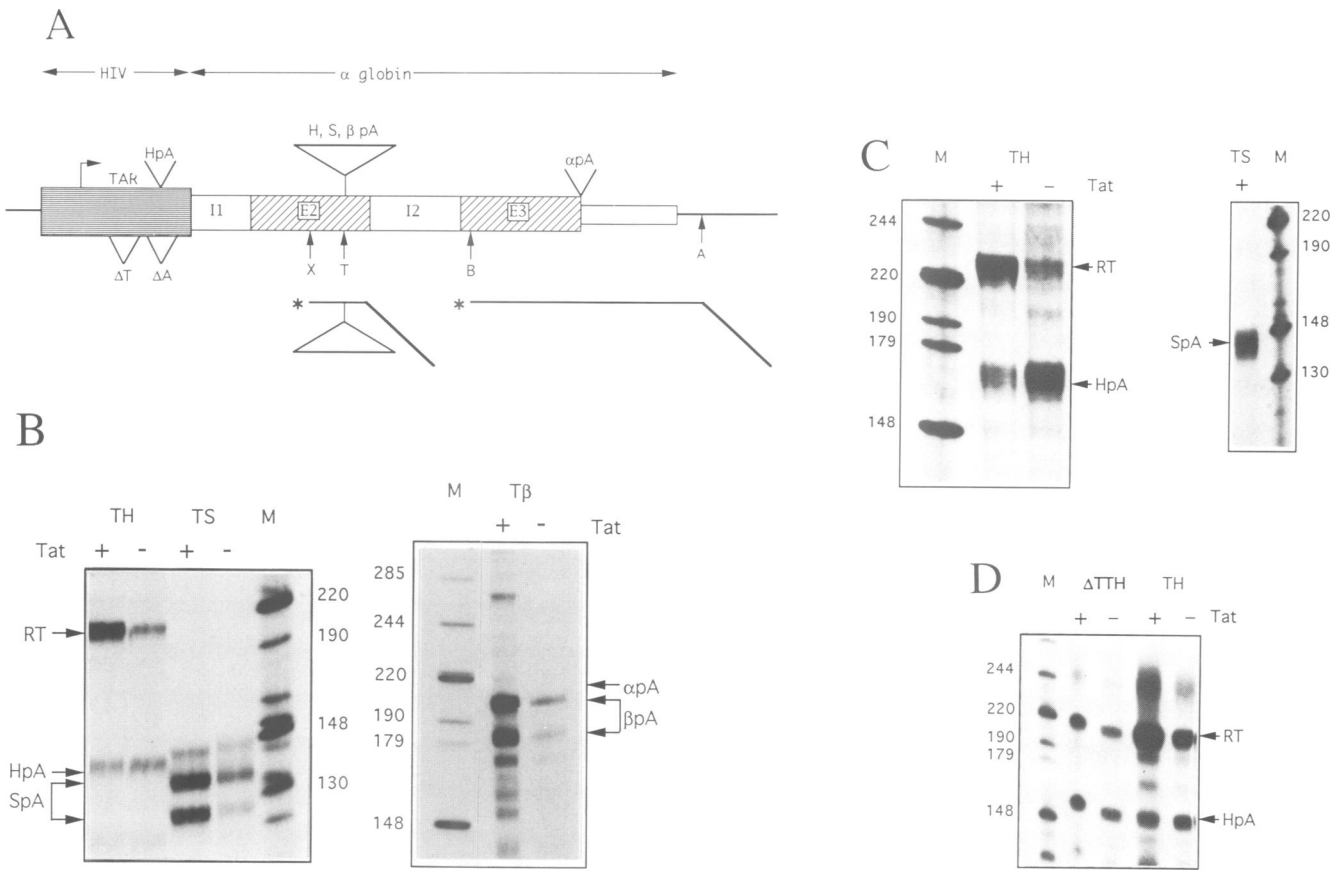

Fig. 2. A. Diagram showing the arrangment of the HIV−α globin gene construct. Horizontally hatched box denotes HIV LTR sequence, open box denotes introns or 3′ untranslated sequence of α globin gene, diagonally hatched box denotes α globin exon and lines denote pSVod vector sequences. ΔT and ΔA denote HIV TAR and poly(A) site deletions, HpA and αpA indicate positions of the HIV and α globin poly(A) sites while X, T, B and A indicate positions of *Bst*XI, *Tth*111I, *Bst*EII and *Ase*I restriction sites in HIV−α. The position of the inserted poly(A) sites at the *Tth*111I site is also indicated. Finally the positions of the two end-labelled DNA probes are shown. * denotes ^{32}P end label while the other end of the probe is angled away from the gene map at the position where it diverges from the mRNA sequence. **B**. S1 analysis of TH, TS and Tβ ΔAHIV−αpSVod co-transfected with (+) or without (−) Tat-expressing plasmid. Homologous *Bst*XI DNA probes were used in each case giving either poly(A) bands (pA) or readthrough bands (RT) to the intron 2 splice site. TH gives a 6-fold Tat response based on densitometric quantification. The faint band visible in the Tβ+ lane is a partial S1 product as it is not the correct size for an authentic readthrough product. **C**. Exonuclease VII analysis of TH and TS ΔAHIV−αpSVod using the TH *Bst*XI probe. TH Tat response in this case is 14-fold, although this higher figure may relate in part to exonuclease VII. **D**. S1 analysis of ΔTTH as compared with TH ΔAHIV−αpSVod using the TH *Bst*XI probe.

of mRNA detected (~5-fold based on the co-transfected control β mRNA levels, data not shown) and that with the rabbit β and SPA signals (called Tβ and TS respectively) nearly 100% use of the inserted poly(A) site occurs, with no qualitative change in the S1 patterns observed with or without Tat. In contrast the HIV poly(A) site (called TH) gives a near equal distribution of poly(A) site and read-through bands without Tat, but switches to producing predominantly read-through transcripts in the presence of Tat. Quantification of these results indicates a 6-fold increase in read-through transcripts when transactivated by Tat. This value for the Tat effect is calculated as the change in ratio between read-through and poly(A) transcript bands with or without Tat.

The patterns of S1 bands produced with these three poly(A) signals in the absence of Tat is entirely consistent with our previous studies showing that the SPA poly(A) site was unaffected by promoter proximity, while HIV was occluded (Weichs an der Glon *et al.*, 1991). As shown in these studies SPA works efficiently in the *Tth*111I site, which is 400 bp away from the transcription start site in HIV−α. The same is true for rabbit β, which is a related poly(A) signal to SPA (Levitt *et al.*, 1989). We previously showed that the HIV poly(A) signal is occluded by promoter proximity up to a distance of ~500 bp. At the 450 bp distance as in these constructs, it is clear that we obtain an intermediate level of occlusion. The observation that with Tat-activated transcription the HIV poly(A) site switches to near complete read-through (occlusion) is a surprising and potentially critical property of HIV transcription. The characterization of this phenomenon provides the principal experimental theme of these studies.

Figure 2C shows the analysis of TH and TS mRNA using the alternative RNA mapping enzyme exonuclease VII. Again for TH a switch from HIV poly(A) site usage to read-through when transactivated by Tat is demonstrated. Higher levels of RNA were used in the TH− lane to allow a clearer comparison with TH+. Quantification of these data indicate a 14-fold switch to read-through transcripts in the presence of Tat. The larger Tat effect seen in this experiment is atypical as a 6-fold effect is more normally obtained. It is possible that using exonuclease VII rather than S1 accentuates the Tat effect. These results also indicate that the lower bands seen with TS and Tβ (Figure 2B) are S1 over digestion

A

SEQUENCES OF HYBRID POLY(A) SITES

					TAT-SENSITIVITY	POLY(A) SITE STRENGTH
H	TTAAGCCTC	AATAAA	GCTTGCCTTGAGTGCTTCAAGTA	GTGTGTGCCCGTCTGTTGTGTG	6.3	30%
S	GATCC	AATAAA	AGATCTTTATTTTCATTAGATCT	GTGTGTTGGTTTTTTGTGTG	1.0	80%
SH	GATCC	AATAAA	AGATCTCTCGAGAGATCTCAAGTA	GTGTGTGCCCGTCTGTTGTGTG	1.0	5%
HS	TTAAGCCTC	AATAAA	GCTAGATCTCTCGAGAGATCTCAAGTA	GTGTGTTGGTTTTTTGTGTG	5.9	5%
HΔS	TTAAGCCTC	AATAAA	GCTTTTATTTTCATTAGATCT	GTGTGTGTGTG	> 6	70%
SΔS	GATCC	AATAAA	AGATCTTTATTTTCATTAGATCT	GTGTGTGTGTG	1.0	65%

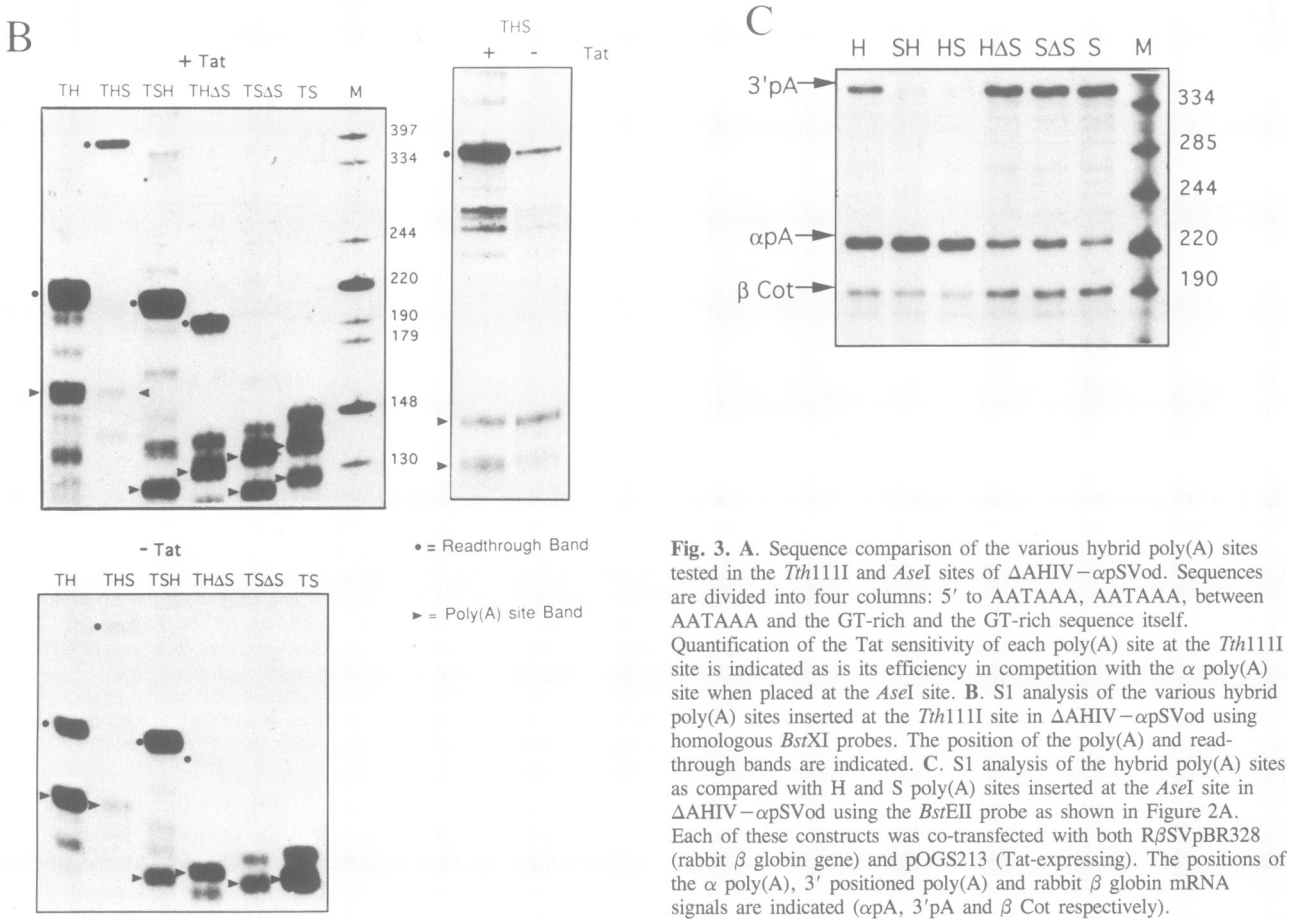

Fig. 3. A. Sequence comparison of the various hybrid poly(A) sites tested in the *Tth*111I and *Ase*I sites of ΔAHIV−αpSVod. Sequences are divided into four columns: 5′ to AATAAA, AATAAA, between AATAAA and the GT-rich and the GT-rich sequence itself. Quantification of the Tat sensitivity of each poly(A) site at the *Tth*111I site is indicated as is its efficiency in competition with the α poly(A) site when placed at the *Ase*I site. **B**. S1 analysis of the various hybrid poly(A) sites inserted at the *Tth*111I site in ΔAHIV−αpSVod using homologous *Bst*XI probes. The position of the poly(A) and read-through bands are indicated. **C**. S1 analysis of the hybrid poly(A) sites as compared with H and S poly(A) sites inserted at the *Ase*I site in ΔAHIV−αpSVod using the *Bst*EII probe as shown in Figure 2A. Each of these constructs was co-transfected with both RβSVpBR328 (rabbit β globin gene) and pOGS213 (Tat-expressing). The positions of the α poly(A), 3′ positioned poly(A) and rabbit β globin mRNA signals are indicated (αpA, 3′pA and β Cot respectively).

products as the sequence around these two poly(A) signals is especially AU-rich. Exonuclease VII does not possess any single strand endonuclease activity (as does S1) and so does not give these extra bands.

A possible explanation for the effect of Tat on the HIV poly(A) signal could be that TAR, the RNA binding site for Tat, is directly adjacent to the poly(A) site so that Tat binding sterically inhibits polyadenylation in HIV. However, in TH the DNA fragment containing the HIV poly(A) site (cloned into the *Tth*111I site) does not contain an intact TAR stem−loop sequence (only the 3′ side of the loop) and so cannot bind to Tat. As shown in Figure 2D, the Tat occlusion effect on the HIV poly(A) site in TH does require TAR in the the HIV promoter. Thus ΔTTH (ΔT refers to deletion of the TAR sequence in the HIV LTR sequence, see Figure 2A and Materials and methods) with or without Tat co-transfection gives an equal ratio of HIV poly(A) site and read-through bands in contrast to TH, which again responds to Tat by increasing occlusion by at least 5-fold. These results lead to the unexpected conclusion that an RNA processing signal 450 bp away from the promoter directly responds to the way in which that promoter is activated. In a number of recent studies Tat activation has been shown to act at the level of transcriptional elongation (Laspia *et al.*, 1990; Feinberg *et al.*, 1991; Marciniak and Sharp, 1991; Kato *et al.*, 1992). Such observations may well relate to this poly(A) site occlusion phenomenon (see Discussion).

Definition of Tat-specific poly(A) signals

We wished to determine which specific sequence associated with the HIV poly(A) site conferred the Tat-dependent, read-through effect described in the previous section. We reasoned

that the construction of hybrid poly(A) signals combining different parts of the HIV and SPA signals should identify Tat-specific signals, since HIV responds while SPA does not respond to Tat. Figure 3A shows the sequences of the various hybrid poly(A) sites investigated compared with the intact HIV and SPA signals, while Figure 3B presents the S1 analysis of HIV−α mRNA produced with these different poly(A) signals inserted at the *Tth*111I site (see in Figure 2A). As shown previously, the HIV poly(A) site responds to Tat while SPA does not. With THS in the absence of Tat, the poly(A) site and read-through bands are of nearly equal intensity, while with Tat the read-through band predominates. The lower level of S1 signals obtained with THS reflects a lower specific activity probe used in this particular experiment. For this reason the experiment with THS (+ or − Tat) is repeated, as shown in a separate panel. Again the same switch from poly(A) transcripts to predominantly read-through is observed. Similarly THΔS gives just a poly(A) site band without Tat, but produces equal amounts of read-through and poly(A) bands with Tat. In contrast the pattern of S1 bands for TSH and TSΔS remains unchanged with or without Tat. It should be noted that the poly(A) bands in THΔS, TSΔS and TS break up into several close together S1 products due to S1 over digestion at AU-rich sequences. In fact, the Tat S1 products are somewhat more digested giving predominantly the lower poly(A) bands. Quantification of the data presented in Figure 2B is indicated against each of the hybrid poly(A) site sequences shown in Figure 2A. As shown, the Tat-responsive poly(A) sites H, HS and HΔS each give a 6-fold or greater switch in ratio from poly(A) site use to read-through.

The experiments described in these studies, showing that Tat has a specific effect on the HIV but not other poly(A) sites are unexpected. Since Tat has been shown to increase the elongation efficiency of Pol II, it might be predicted that its effect on poly(A) site usage merely relates to the relative strength of the poly(A) signal. We have therefore carried out experiments to determine whether the relative strengths of the different poly(A) signals used in Figure 3 related to their Tat responsiveness. Figure 3C shows the analysis of the four different hybrid poly(A) sites as well as HIV and SPA when placed 160 bp downstream of the α poly(A) site in HIV−α [at the *Ase*I site (A); see Figure 2A]. The relative use of these downstream positioned poly(A) sites to α gives a measure of their efficiency as RNA processing signals. Using an S1 3′ end-labelled probe from the *Bst*EII site in HIV−α (see Figure 2A), the proportion of transcripts that use the α or downstream positioned poly(A) sites can be directly measured. Quantification of these results is presented in Figure 3A. It is clear that SPA is the strongest poly(A) site in competition with α as 80% of the mRNA uses SPA rather than α. Similarly HΔS and SΔS are both equally strong, though somewhat less so than SPA. The HIV poly(A) site gives a band of ~30% the level of α, while HS and SH both produce bands at ~5% the level of α. With the two weakest poly(A) sites, HS responds to Tat and SH does not, while with the much stronger poly(A) sites, HΔS and SΔS, only HΔS responds to Tat. These results indicate that the Tat response effect described in these studies is independent of intrinsic poly(A) site strength.

The only sequences common to HIV and the two hybrid poly(A) signals that respond to Tat (HS and HΔS) are positioned immediately before and after the AAUAAA signal (see Figure 3A). To investigate these Tat-responsive sequences more precisely, we have generated specific mutations in the HIV poly(A) site as shown in Figure 4A. These mutated poly(A) sites were constructed by oligonucleotide replacement using rare cutting restriction sites in the HIV poly(A) site region, as described in the Materials and methods. They were again tested for Tat responsiveness by insertion into the *Tth*111I of HIV−α followed by transfection into HeLa cells with or without Tat. As shown in Figure 4B panel I, TH and TH* (which has a single base change creating an *Xho*I) give a 6-fold Tat response. Quantification of these experiments repeated four times reveals a consistent 6-fold Tat effect (Figure 4A). Mutation of the sequence 5′ to AAUAAA from TTAAGCCTC>$CTAG_5$ (clone 5′ΔTH) only slightly reduces the Tat effect to 4-fold (Figures 4A and B, panel II). In contrast clone 3′ΔTH1 which mutates GCTTGC>GGAAAT gives a greater reduction in the Tat response to 2.5-fold (Figures 4A and B, panel III). This particular mutant sequence was selected as it is based on the rabbit β globin poly(A) site, which is not Tat-responsive (Figure 2B). Since both the point mutation in TH* and the five base change in 3′ΔTH resulted in a significantly weaker poly(A) site than in TH (data not shown), it proved necessary to increase the spacing between the inserted poly(A) site and the HIV promoter (3′ΔTH1, see next section) to enhance the usage of the 3′Δ poly(A) site and so obtain a more accurate measure of the Tat response effect. As a control for 3′ΔTH1, TH1 Tat responsiveness was also measured in duplicate and gave a 7-fold effect, similar to TH.

Taken together these data confirm the analysis of the hybrid HIV−SPA poly(A) sites, which implicates the sequences immediately before and after the AAUAAA sequence in the Tat response. However, it is apparent that the sequences immediately 3′ to the AAUAAA of the HIV poly(A) site are quantitatively more significant.

The effect of promoter proximity and U3 sequences on Tat responsiveness of the HIV poly(A) site

The two previously described mechanisms for HIV poly(A) site occlusion (promoter proximity and a requirement for U3 sequences) might seem more than sufficient to account for the inactivation of the HIV poly(A) site in the 5′LTR. It was therefore unexpected that there is yet another, third mechanism to occlude the 5′ LTR poly(A) signals in HIV. We therefore tested both the effect of promoter proximity and U3 sequence on the HIV poly(A) site using Tat-activated HIV transcription.

As shown in Figure 5A, the HIV poly(A) site was moved with increasing distance away from the HIV promoter in HIV−α by inserting it into either the *Bst*EII site to make clone BH (630 bp away from the promoter) or the *Ase*I site to make clone AH (1000 bp away). Also a monomer and dimer fragment of 'neutral DNA' (170 bp in length, derived from pUC119) was inserted immediately upstream of the HIV poly(A) site when in the *Tth*111I site of HIV−α to make clones TH1 and TH2 (increasing the spacing from 450−620 and 790 bp respectively). Each of these spacer clones was transfected into HeLa cells and co-transfected with both the Tat expression plasmid and the other control plasmid that expresses T antigen and rabbit β globin mRNA. Figure 5B (panel I) shows the S1 analysis of these spacer clones using the 3′ end-labelled *Bst*EII probe for HIV−α as an indirect measure of HIV poly(A) site occlusion. Based on quantification of the α poly(A) signal relative to the co-

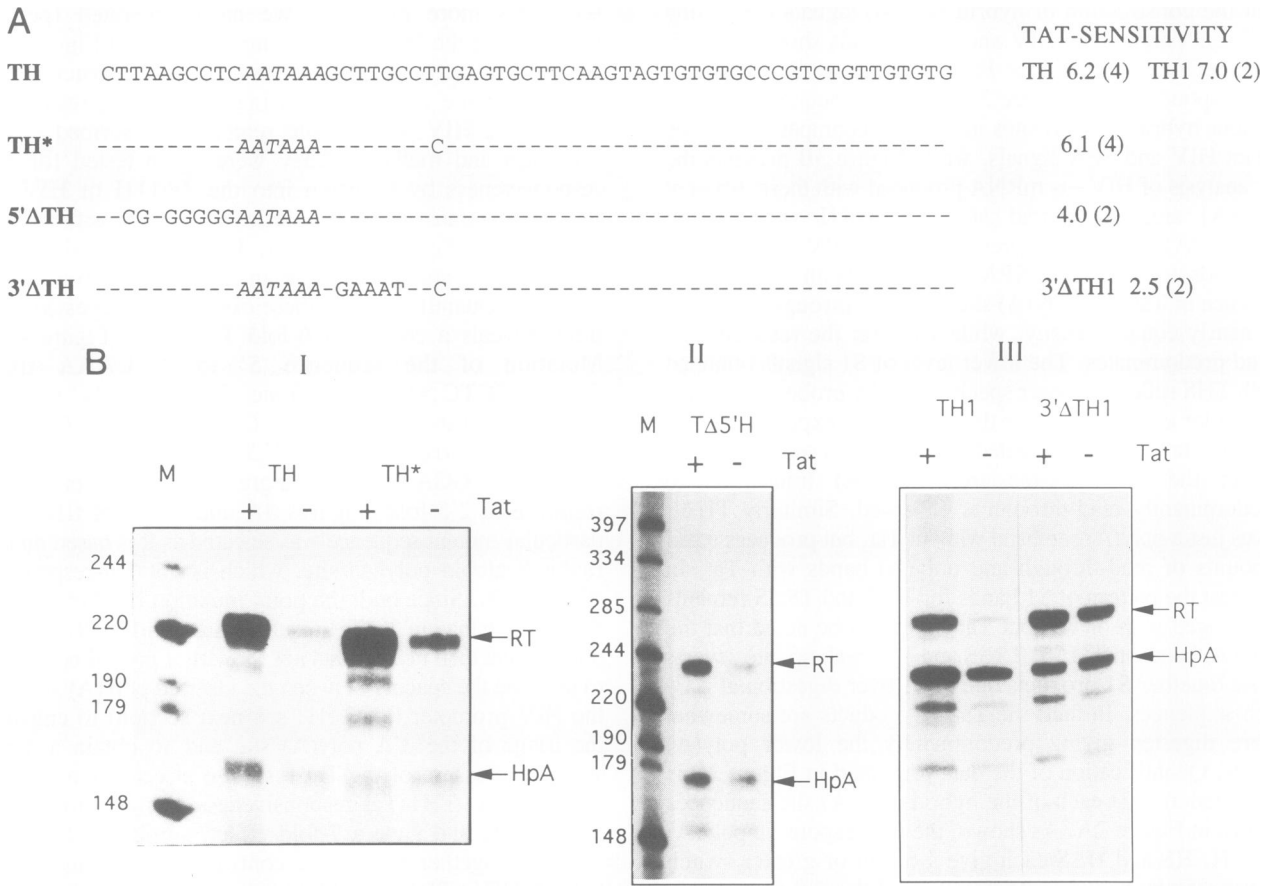

Fig. 4. A. Sequence comparison of the mutated HIV poly(A) sites. The AAUAAA sequence is written in italics and the altered nucleotides are aligned under the HIV poly(A) site sequence. Dashed lines denote identity with the HIV sequence. The quantification of the Tat effect data for each mutated poly(A) site is also indicated. The figure in brackets indicates the number of experiments from which the average value is derived. **B.** S1 analysis data of representative experiments for the effect of Tat on TH and TH* (panel I), 5′ΔTH (panel II) and 3′ΔTH1 (panel III). In each case homologous *Bst*XI probes were used. The HIV poly(A) site and read-through bands are again indicated.

transfection β mRNA control signal, TH gives 75% levels of read-through transcripts that utilize the α poly(A) site as compared with HIV$-\alpha$ [with no inserted HIV poly(A) site], confirming that the HIV poly(A) site is largely occluded in the presence of Tat at 450 bp spacing from the HIV promoter. This result is entirely consistent with the direct analysis of TH as shown in Figure 2B where >80% of transcripts read through the HIV poly(A) site. BH gives at least a 2-fold drop in read-through transcripts to a 30% level of transcripts that utilize the α poly(A) site (as compared with HIV$-\alpha$), while AH shows that the HIV poly(A) site is now quite active and competes significantly with the upstream positioned α poly(A) site (40% in this analysis). In contrast, SPA inserted at the *Tth*111I site (TS) totally blocks read-through transcripts to the α poly(A) site and largely out-competes the α poly(A) site (to a >90% level) when placed downstream at the *Ase*I site (AS). Again direct analysis of the same TS RNA using the *Bst*XI probe confirms that all transcripts utilize the inserted SPA (Figure 2B). The other spacer constructs, TH1 and TH2, confirm these results (panel II) as quantification of TH1 RNA gives a 65% use of the HIV poly(A) site, while TH2 gives a >95% use of the HIV poly(A) site. In these experiments the HIV poly(A) site is directly probed for using the *Bst*XI S1 probe as described in Figure 2A. Taken together, these spacing experiments suggest that the Tat-mediated occlusion effect of the HIV poly(A) site is no longer effective when the spacing exceeds ~700 bp. At this position the HIV poly(A) is fully active in the presence of Tat. Between 500 and 700 bp spacing we observe intermediate Tat occlusion of the HIV poly(A) site, as indicated by BH and TH1.

To investigate the effect of U3 sequences on Tat occlusion of the HIV poly(A) site, a 160 bp fragment containing 135 bp upstream of the HIV transcription start site and 25 bp downstream was inserted 5′ of the HIV poly(A) site in TH and AH to form THU3 and AHU3 (as shown in Figure 5A). This U3 fragment contains all of the sequence elements implicated by Valsamakis *et al.* (1991) and DeZazzo *et al.* (1991) as required for U3 activation of the HIV poly(A). Figure 5C (panel I) shows S1 analysis of RNA from THU3 and TH transfections in the presence of Tat using the HIV$-\alpha$ 3′ *Bst*EII S1 probe as an indirect measure of HIV poly(A) site usage. With THU3, only low levels of read-through transcripts to the α poly(A) site are detected as compared with full levels of read-through transcripts obtained with TH. This effect cannot be attributed solely to spacing as in the case of BH, where the spacing of the HIV promoter to poly(A) site is about the same as for THU3, significant levels of read-through transcripts that utilize the α poly(A) site are detected (Figure 5B). In addition, using the direct *Bst*XI S1 probe for THU3, it can be seen in Figure 5C (panel II) that with or without Tat the HIV poly(A) site is fully active in contrast to TH1, whereas Tat promotes 30% read-through of the HIV poly(A) site. These results confirm in our

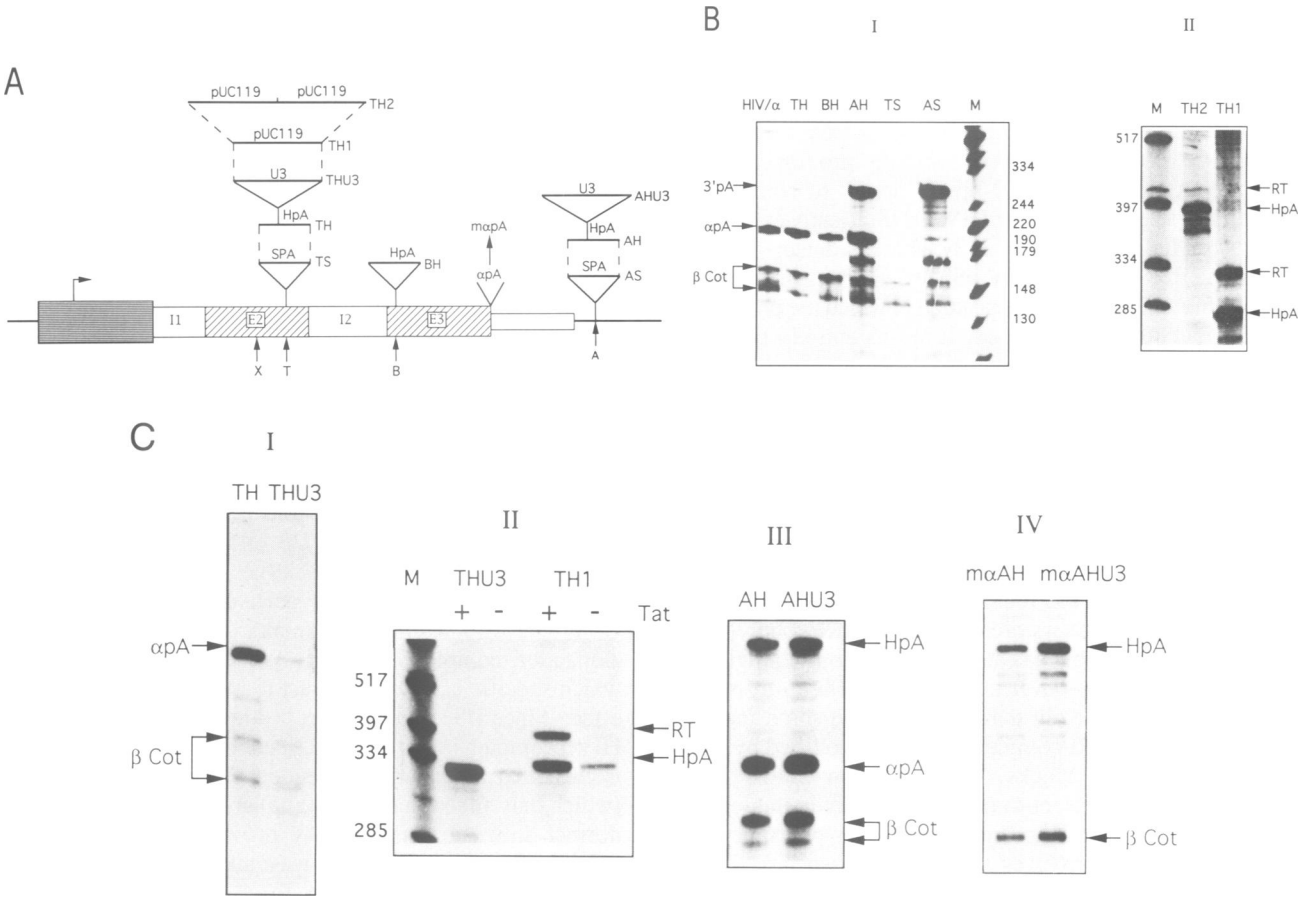

Fig. 5. A. Diagram of the various spacer and U3 constructs based on ΔAHIV−αpSVod (as in Figure 2A). The HIV poly(A) site (HpA) and SPA are inserted into the *Tth*111I, *Bst*EII and *Ase*I sites and either pUC119 or U3 sequence are in turn inserted into TH and AH. The α poly(A) site is mutated to an inactive form in mαpA. **B**. Panel I: S1 analysis of TH, BH, AH, TS and AS ΔAHIV−αpSVod as compared with ΔAHIV−αpSVod (HIV−α), using the *Bst*EII probe as in Figure 3C. Panel II: S1 analysis of TH2 and TH1 ΔAHIV−αpSVod using homologous *Bst*XI probes as in Figure 2A. **C**. Panels I, III and IV: S1 analysis of TH, THU3, AH, AHU3, mαAH and mαAHU3 ΔAHIV−αpSVod all using the *Bst*EII probe as in Figure 3C. Panel II: S1 analysis of THU3 compared with TH1 ΔAHIV−αpSVod using homologous *Bst*XI probes.

experimental system that U3 sequences have an activating effect on the HIV poly(A) site and more importantly, this U3 effect appears to block Tat-mediated occlusion. Finally, with the construct AHU3 (panel III), we show that the usage of the HIV poly(A) site is unaffected by U3 sequence when positioned at a 1 kb distance from the HIV promoter, either in competition with the α poly(A) site or when the α poly(A) site is inactivated by a point mutation (panel IV). In effect we can only detect a U3 effect on the HIV poly(A) site when it is in a position where it would otherwise be inhibited by Tat-activated transcription. The effects of both spacing between promoter and poly(A) site with or without Tat, and also U3 sequence on the usage of the HIV poly(A) site, are summarized in Table I.

Table I. The effects of both spacing between promoter and poly(A) site with or without Tat and U3 sequence on the usage of the HIV poly(A) site.

	400n (*Tth*111I)		1000n (*Ase*I)	
	−Tat	+Tat	−Tat	+Tat
−U3	+	− (5′ LTR)	++	++
+U3	++	++	++	++ (3′ LTR)

*Tth*111I and *Ase*I indicate the positions used in HIV−αpSVod. 5′ and 3′ LTR indicate conditions described in these studies that are closest to the two poly(A) sites in the HIV-1 provirus. −, inactive; +, partially active; ++, fully active HIV poly(A) site.

Discussion

We have uncovered a novel activity associated with Tat *trans*-activation of the HIV promoter that mediates the sequence-specific read-through of the HIV poly(A) site. Previously the proximity of the HIV poly(A) site to its promoter in the 5′LTR and the presence of U3 RNA sequence elements upstream of the HIV poly(A) site in the 3′ LTR have both been argued to account for poly(A) site selection in HIV. However, this third mechanism of HIV poly(A) site regulation has all of the hallmarks of a physiologically significant mechanism, as both the key HIV transactivator, Tat and a specific HIV RNA sequence are required. All of the experiments to date on HIV poly(A) site regulation have used heterologous gene constructions expressed on episomal plasmids, transfected into tissue culture cell lines. Clearly the chromosomally integrated HIV provirus may display different parameters to those observed in the currently employed experimental systems. Establishing the relative importance of these three mechanisms for regulation of HIV poly(A) site usage must await analysis of appropriately mutated proviruses stably transfected into

human T cells. However, it should be mentioned that all experimental systems employed to study HIV polyadenylation other than our own HIV−α globin reporter system have used heterologous promoters that may therefore obscure the true mechanism of HIV poly(A) site regulation. Even so we have shown that the promoter proximity of the HIV poly(A) site of only 80 bp is enough to completely occlude the poly(A) site in our HIV−α transiently expressed constructs. Indeed we were only able to detect the Tat occlusion effect when the HIV poly(A) site was moved 450 bp away from the promoter, a position at which the proximity effect is much less pronounced. It seems entirely possible that in the actively transcribing provirus, proximity effects are insufficient to completely block the 5′LTR poly(A) signals, so that this further Tat-specific mechanism needs to operate.

The molecular basis of the Tat-specific occlusion effect is as yet uncharacterized, but must account for a number of features established in these studies. First, the HIV poly(A) site possesses specific sequences immediately 5′ and 3′ to AAUAAA that are required for the Tat effect and to which *trans*-acting factors presumably bind. Secondly, this effect is dependent on a Tat-modified Pol II elongation complex, but only acts over a distance of 700 bp from the promoter. It seems likely that there may be some feature of the elongating Pol II complex that when modified by Tat, activation somehow blocks the recognition of the HIV poly(A) site in the nascent transcript, but only in the region <700 bp away from the promoter. It seems plausible that this blocking effect involves recognition of specific RNA sequences on either side of AAUAAA so that the access of polyadenylation factors to AAUAAA would be impeded. That Tat mediates its effect on HIV transcription at the elongation level is now a well described phenomenon shown both *in vivo* (Laspia *et al.*, 1990; Feinberg *et al.*, 1991) and *in vitro* (Marciniak and Sharp, 1991; Kato *et al.*, 1992). As described in these experiments Tat mediates its effect on the HIV poly(A) site up to 700 bp away from the promoter. Therefore the direct involvement of Tat in HIV poly(A) usage is entirely consistent with a transcriptional elongation role for Tat.

Our previous studies on HIV poly(A) occlusion (Weichs an der Glon *et al.*, 1991) not only demonstrated that the poly(A) site was directly inhibited by promoter proximity but also that this effect was specific to the GU-rich downstream sequence of the HIV poly(A) site. The α globin poly(A) site and SPA were not affected by promoter proximity. These intrinsic occlusion properties of the HIV poly(A) site can now be seen as only a part of the mechanism that blocks the HIV poly(A) site in the 5′ LTR. Since HIV proviral transcription will in general occur in the presence of Tat, the Tat-specific occlusion mechanism may prove to be the quantitatively more important. However, it is possible that when the provirus is initially transcribed, before the levels of Tat are high enough to activate the promoter, the intrinsic ability of the HIV poly(A) site to be occluded by promoter proximity in the 5′ LTR is enough to allow transcription to read through the viral genome. Once the level of Tat transactivation increases to the stage where the HIV promoter is extremely active, a further mechanism may be needed to prevent usage of the poly(A) site in the 5′ LTR. What better way than to involve the Tat transactivating molecule itself in this mechanism?

The role of U3 sequences in activating the HIV poly(A) site remains a clear phenomenon described in a number of heterologous transient expression systems. U3 sequences have also been shown to increase the activity of the HIV poly(A) site *in vitro* using nuclear extracts, although the level of activation is lower than that found *in vivo* (Gilmartin *et al.*, 1992; Valsamakis *et al.*, 1992). As described in these studies, we have also observed activation of the HIV poly(A) site by U3 sequences in our own HIV−α gene system. However, we can only detect a significant U3 activation effect when the HIV poly(A) site is occluded by promoter proximity. At a distance of >700 bp we find that the HIV poly(A) site is active without U3 sequences and is not further activated by their presence. These results contrast with those of Valsamakis *et al.* (1991) and DeZazzo *et al.* (1991) who both reported a significant activation of the HIV poly(A) site by U3 sequences when placed over 1 kb from the promoter. It is possible that the heterologous promoters used in these studies have influenced the U3 effect. The fact that U3 sequences can completely block the Tat-specific occlusion effect as described here is provocative and may suggest that U3 can have some direct role in modifying the Tat-activated elongation complex of Pol II. We are currently defining which specific U3 sequences act to block the Tat occlusion effect. Since U3 sequences are >9 kb away from the active HIV promoter in the 5′ LTR sequence of the provirus, it is clear that the position-dependent U3 effects we observe (which only function up to 700 bp) are inconsistent with the normal dimensions of the HIV provirus.

In conclusion our results suggest that the HIV poly(A) site uses a tripartite mechanism to regulate its utilization: promoter proximity and suppression by Tat-activated transcription, which combined with the absence of U3 sequences all contribute to the occlusion of the HIV poly(A) site in the 5′ LTR. However, as is true for all of these HIV poly(A) site occlusion and activation mechanisms, their true physiological importance must await analysis of the whole provirus.

Materials and methods

DNA construction

HIV−αpSVod. This is described by Proudfoot *et al.* (1992). Essentially an HIV-1 LTR fragment from a *Sca*I site in U3 (−135) to a *Hin*fI site in R (+125) was fused to the human α2 globin gene at a *Sma*I site in the middle of intron 1 to form an HIV−α gene hybrid. This gene construct was then inserted between the *Sca*I and *Pst*I sites of pSVod in the Amp^r gene using the *Pst*I site at the 3′ end of the α globin gene. pSVod is a transient expression vector containing the SV40 origin (but not adjacent enhancer) and lacks the pBR322 replication poison sequence, as described by Lusky and Botchan (1981).

ΔT and ΔAHIV−αpSVod. This was obtained by deletion of the HIV TAR sequence from *Bgl*II−*Afl*II (+25 to +65) and HIV poly(A) signals from *Afl*II−*Hin*dIII (+65 to +80) respectively. With ΔA the deleted sequence was replaced with a linker oligonucleotide containing *Xho*I and *Hin*dIII sites.

TH, BH, AH, TS and AS ΔAHIV−αpSVod. These constructions have the HIV and SPA poly(A) signals inserted into the *Tth*111I site (exon 2 of the α globin gene, T), the *Bst*EII (exon 3 of the α globin gene, B) or the *Ase*I site (3′ vector sequence of pSVod, A). H is the HIV poly(A) signal excised from pSP65HIVLTR, which has the AvaI−*Hin*fI HIV-1 LTR sequence inserted into the polylinker at the *Sma*I of pSP65 (see Weichs an der Glon *et al.*, 1991). H is either the 40 bp *Afl*II(HIV)−*Xba*I(polylinker) fragment inserted into the T site as in Figure 2B, or the 80 bp *Bgl*II(HIV)−*Xba*I (polylinker) fragment inserted into T, B and A sites of HIV−αpSVod, as in subsequent Figures. S is the synthetic poly(A) site (Levitt *et al.*, 1989) excised from a subclone (called pySPA; Enriques-Harris *et al.*, 1991), in

which the SPA is inserted into the polylinker sequence of pUC119 (at the *Sma*I site). The whole polylinker with inserted SPA was excised by cutting with *Eco*RI and *Hin*dIII and then inserted into the T and A sites of HIV−α pSVod.

Tβ *ΔAHIV−αpSVod*. This construct has a 140 bp *Bgl*II−*Nde*I fragment containing the rabbit β globin poly(A) site inserted between unique *Xba*I and *Asp*718 sites in the polylinker sequence on either side of the SPA in TS.

T and A HS, SH, HΔS and SΔS ΔAHIV−αpSVod. These constructions are various hybrid poly(A) sites inserted between unique *Xba*I and *Asp*718 sites in the polylinker sequence on either side of the SPA in TS and AS. HS poly(A) site was constructed by first replacing the 3′ side of the HIV AATAAA sequence in pSP65HIVLTR from *Hin*dIII−*Hin*dIII (in polylinker) with an oligonucleotide as shown in Figure 3A that contains a linker sequence followed by the SPA GT-rich sequence. HS was excised from this intermediate construct by digesting with *Afl*II (in HIV R) and *Pvu*II (in pSP65) generating a 255 bp fragment. SH poly(A) site was constructed as an intermediate clone by replacing the sequence 3′ to AATAAA in pySPA, with an oligonucleotide as shown in Figure 3A containing the same linker sequence as in HS followed by the HIV GT-rich sequence. SH was excised as an *Eco*RI−*Hin*dIII fragment and inserted directly into the *Tth*111I and *Ase*I sites of ΔAHIV−αpSVod. HΔS and SΔS poly(A) sites were both constructed by oligonucleotide replacements of H and S poly(A) sites. With HΔS, the HIV poly(A) site in pSP65HIVLTR was deleted from the *Hin*dIII site (immediately 3′ to the AATAAA sequence) to *Bam*HI (in polylinker) and replaced with an oligonucleotide of S sequence, but with a truncated GT-rich sequence (see Figure 3A). HΔS was excised from this intermediate clone as an *Afl*II−*Xba*I fragment and inserted into the TS and AS *Xba*I- and *Asp*718-cut vectors. SΔS was made by directly replacing sequence from *Bgl*II−*Asp*718 in TS and AS with the same oligonucleotide sequence as in HΔS.

5′Δ and 3′Δ TH ΔAHIV−αpSVod. These were constructed as follows. Δ5′TH poly(A) site was made as an intermediate clone by first mutating the sequence 5′ to the AATAAA sequence of the HIV poly(A) site in pSP65HIVLTR from TTAAGCCTC > TTAACTCGAGGGGG by replacing the *Afl*II−*Bam*HI (in the pSP65 polylinker sequence) fragment with a mutant oligonucleotide. This mutated HIV poly(A) site was then excised as a *Bgl*II−*Xba*I fragment and inserted into TS and AS as before.

3′ΔTH. This was constructed by first generating an *Xho*I site 3′ to the HIV AAUAAA sequence by oligonucleotide replacement to form TH*. This construct was further modified by replacing the *Afl*II−*Xho*I fragment with another mutant oligonucleotide sequence to specifically mutate the sequence 3′ to the HIV AAUAAA-GCTTCG to that of the rabbit β globin poly(A) site-GGAAAT (which is not responsive to Tat; Figure 2B). 3′ΔTH was further modified by insertion of a spacer fragment as for TH1.

TH1, TH2, THU3 and AHU3(mα) ΔAHIV−αpSVod. These have inserts into the HIV poly(A) signal *Afl*II site (5′ to AATAAA) in TH and AH. 1 and 2 refer to a monomer or dimer pUC119 spacer fragment of 170 bp while U3 refers to the 160 bp HIV U3-R region from *Sca*1 (−135 bp) to *Bgl*II (+25 bp). mα indicates that the α poly(A) site in HIV−α has been converted from AATAAA > AATAAG as found in an α thalassaemia gene (Higgs *et al.*, 1983).

Transfection of HeLa cells

Subconfluent HeLa Cells were transiently transfected with 15 μg HIV−αpSVod plasmid (with various alterations), 5 μg RβSVpBR328 (expresses rabbit β globin mRNA and T antigen) and with or without 5 μg pOGS213 (expresses Tat; Adams *et al.*, 1988) using calcium phosphate co-precipitation (see Weichs an der Glon *et al.*, 1991). After 2 days cells were harvested and cytoplasmic RNA was purified according to Maniatis *et al.* (1982).

RNA mapping techniques

DNA probes. *Bst*XI was used to linearize the various HIV−α plasmids. This DNA was then 'tailed' with [α-^{32}P]dTTP using terminal deoxynucleotide transferase (Boehringer). *Bst*EII-cut HIV−αpSVod and *Eco*RI-cut RbSVpBR318 was filled in with [α-^{32}P]dNTPs using Klenow DNA polymerase. ^{32}P DNA was purified by G100 Sephadex chromatography. The same HIV−αpSVod *Bst*EII probe was used for all the different HIV−α gene constructs, so that the S1 band obtained for poly(A) sites positioned downstream of the α poly(A) site is a mismatch band (due to sequence divergence between the probe and mRNA at the *Ase*I site in HIV−αpSVod). The intensity of this S1 band directly correlates with the strength of the downstream poly(A) site.

SI nuclease and exonuclease VII RNA mapping. These procedures have been described in detail by Johnson *et al.* (1986). In outline 20−100 c.p.s. of DNA probe was annealed to 50 μg of transfected HeLa cell cytoplasmic RNA in R-loop buffer at 52°C and then diluted into S1 or *Exo*VII buffer plus enzyme. The digested DNA was then fractionated on denaturing polyacrylamide gels. The RNA mapping data obtained was quantitated by densitometric analysis of the radioautographs. The numerical data is presented in the text and/or in the figures or their legends. The Tat sensitivity values are calculated as the change in ratio between read-through and poly(A) transcript bands with or without Tat transactivation.

Acknowledgements

We gratefully acknowledge the contribution made by Joan Monks to the production of several HIV−α constructs used in these experiments. We also thank members of the NJP lab for help and encouragement throughout these experiments as well as in production of the paper. C.W. and M.A. are both supported by Aids Directed Programme MRC studentships. C.W. also received generous additional support from the Joseph Gorres Foundation and Edward P.Abraham Research Fund. J.E. is supported by an EMBO long-term fellowship and is a senior research assistant of the National Fund for Scientific Research of Belgium. Finally this work was supported by a Programme Grant from the Wellcome Trust (no. 032773) and a project grant from the MRC AIDS-directed programme (no. G9017434).

References

Adams,S.E., Johnson,I.D., Braddock, M., Kingsman,A.J., Kingsman,S.M. and Edwards,R.M. (1988) *Nucleic Acids Res.*, **16**, 4287−4297.
Ahmed,Y.F., Gilmartin,G.M., Hanly,S.M., Nevins,J.R. and Greene,W.C. (1991) *Cell*, **64**, 727−737.
Berkhout,B., Silverman,R.H. and Jeang,K.-T. (1989) *Cell*, **59**, 273−282.
Brown,P.H., Tiley,L.S. and Cullen,B.R. (1991) *J. Virol.*, **65**, 3340−3343.
Cherrington,J. and Ganem,D. (1992) *EMBO J.*, **11**, 1513−1524.
Connelly,S. and Manley,J.L. (1988) *Genes Dev.*, **2**, 440−452.
Cullen,B.R. (1990) *Cell*, **63**, 655−657.
DeZazzo,J.D., Kilpatrick,J.E. and Imperiale,M.J. (1991) *Mol. Cell. Biol.*, **11**, 1624−1630.
Dingwall,C. *et al.* (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6925−6929.
Enriquez-Harris,P., Levitt,N., Briggs,D. and Proudfoot,N.J. (1991) *EMBO J.*, **10**, 1833−1842.
Feinberg,M.B., Baltimore,D. and Frankel,A.D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4045−4049.
Frankel,A.D. (1992) *Curr. Opin. Genet. Dev.*, **2**, 293−298.
Gilmartin,G.M., Fleming,E.S. and Oetjan,J. (1992) *EMBO J.*, **11**, 4419−4428.
Grosveld,G.C., de Boer,E., Shewmaker,C.K. and Flavell,R.A. (1982) *Nature*, **295**, 120−126.
Hauber,J., Perkins,A., Heimer,E.P., Cullen,B.R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6367−6368.
Higgs,D.R., Goodbourn,S.E, Lamb,J., Clegg,J.B., Weatherall,D.J. and Proudfoot,N.J. (1983) *Nature*, **306**, 398−400.
Iwasaki,K. and Temin,H.M. (1990) *Genes Dev.*, **4**, 2299−2307.
Iwasaki,K. and Temin,H.M. (1992) *Gene Expr.*, **2**, 7−18.
Jackson,R.J. and Standart,N. (1990) *Cell*, **62**, 15−24.
Johnson,M.R., Norman,C., Reeve,M.A., Scully,J. and Proudfoot,N.J. (1986) *Mol. Cell. Biol.*, **6**, 4008−4018.
Jones,K.A. (1989) *New Biol.*, **1**, 127−135.
Kao,S.Y., Calman,A.F., Luciw,P.A. and Peterlin,B.M. (1987) *Nature*, **330**, 489−493.
Kato,H., Sumimoto,H., Pognonec,P., Chen,C.H., Rosen,C.A. and Roeder,R.G. (1992) *Genes Dev.*, **6**, 655−666.
Laspia,M.F., Rice,A.P. and Mathews,M.B. (1989) *Cell*, **59**, 283−292.
Laspia,M.F., Rice,A.P. and Mathews,M.B. (1990) *Genes Dev.*, **4**, 2397−2408.
Levitt,N., Briggs,D., Gil,A. and Proudfoot,N.J. (1989) *Genes Dev.*, **3**, 1019−1025.
Logan,J., Falck,P.E., Darnell,J.J. and Shenk,T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8306−8310.
Lusky,M. and Botchan,M. (1981) *Nature*, **293**, 79−81.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning, A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Marciniak,R.A. and Sharp,P.A. (1991) *EMBO J.*, **10**, 4189−4196.
Mellon,P., Parker,V., Gluzman,Y. and Maniatis,T. (1981) *Cell*, **27**, 279−288.

Proudfoot,N.J. (1991) *Cell*, **64**, 671–674.
Proudfoot,N.J. and Whitelaw,E. (1988) In Hames,B.D. and Glover,D.M. (eds), *Transcription and Splicing*. IRL Press, Oxford, UK.
Proudfoot,N.J., Lee,B. and Monks,J. (1992) *New Biol.*, **4**, 368–381.
Sharp,P.A. and Marciniak,R.A. (1989) *Cell*, **59**, 229–230.
Southgate,C.D. and Green,M.R. (1991) *Genes Dev.*, **5**, 2496–2507.
Valsamakis,A., Zeichner,S., Carswell,S. and Alwine,J.C. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2108–2112.
Valsamakis,A., Schek,N. and Alwine,J.C. (1992) *Mol. Cell. Biol.*, **12**, 3699–3705.
Weichs an der Glon,C., Monks,J. and Proudfoot,N.J. (1991) *Genes Dev.*, **5**, 244–253.
Whitelaw,E. and Proudfoot,N. (1986) *EMBO J.*, **5**, 2915–2922.

Received on December 8, 1992; revised on February 4, 1993