Occlusion of the HIV poly(A) site

Caroline Weichs an der Glon, Joan Monks, and Nick J. Proudfoot

Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE UK

To investigate the selective use of poly(A) sites in the 3' long terminal repeat (LTR) but not the 5' LTR of retroviruses, we have studied the poly(A) site of the human immunodeficiency virus (HIV-1). Using hybrid HIV/a-globin gene constructs, we demonstrate that the HIV poly(A) site is inactive or occluded when adjacent to an active promoter, either the homologous HIV promoter or the a-globin gene promoter. Furthermore, this occlusion of the HIV poly(A) site occurs over a considerable distance of up to at least 500 bp. In contrast, two nonretroviral poly(A) sites [a-globin and a synthetic poly(A) site] are active when close to a promoter. We also show that a short fragment of ~60 nucleotides containing the HIV poly(A) site is fully active when placed at the 3' end of the human a-globin gene or within the rabbit β-globin gene. This result rules out the requirement of more distant upstream elements for the activity of the HIV poly(A) site, as has been suggested for other viral poly(A) sites. Finally, we show that the GT-rich downstream region of the HIV poly(A) site confers poly(A) site occlusion properties on a synthetic poly(A) site. This result focuses attention on this more variable part of a poly(A) site in retroviruses as a possible general signal for poly(A) site occlusion.

[Key Words: Poly(A) sites; HIV-1, LTRs; retroviral transcription]

Received August 9, 1990, revised version accepted November 26, 1990.

The duplication of transcriptional control sequences in the long terminal repeats (LTRs) of retroviruses presents unusual regulatory features. The promoter in the 5' LTR initiates transcription that reads through the adjacent poly(A) site into the viral genes beyond. Transcription then proceeds into the 3' LTR forming a polyadenylated 3' terminus at the identical poly(A) site in the 3' LTR. As shown in Figure 1A, this pattern of transcription is achieved by somehow inactivating both the poly(A) site in the 5' LTR and the promoter in the 3' LTR even though identical copies of these two transcription signals are highly active in the opposite LTR (Varmus 1988).

A molecular explanation for the inactivity of the 3' LTR promoter may be that of transcriptional interference. Cullen et al. (1984) demonstrated that in the avian leukosis retrovirus (ALV), transcription from the 5' LTR directly inactivates the 3' LTR promoter. This inactivation could be alleviated by either deleting the 5' LTR promoter or by placing SV40 termination of transcription signals between the two LTRs. Presumably, transcription reading into the 3' LTR blocks the formation of a transcription initiation complex, as has been shown in RNA polymerase I genes (Bateman and Paule 1988; Henderson et al. 1989).

The molecular basis of retroviral poly(A) site occlusion is uncharacterized. In some cases, such as the Rous sarcoma virus (Ju and Cullen 1985), the transcription initiation site lies between the two parts of the poly(A) signal: the A3UA3 and GU-rich sequence elements (Humphrey and Proudfoot 1988). Consequently, the 5' LTR poly(A) site is inactive because the complete RNA processing signal is absent from the transcript. However, more commonly, the poly(A) signals of retroviruses are 50–100 bp 3' to the transcription start site so that the whole RNA processing signal should be within the RNA transcript (Varmus 1988).

Such is the situation in the human immunodeficiency virus HIV-1, where the transcription start site is 72 bp 5' to the poly(A) signal (Varmus 1988). Both the high levels of transcription achieved by HIV, as well as the intense interest in characterizing the molecular mechanisms that regulate it, prompted us to investigate HIV poly(A) site occlusion in more detail. To do this, we have set up a model system by making hybrid HIV LTR/human α2-globin gene constructs with the LTR in place of the α-globin promoter, as well as positioned at the 3' end of the α-globin gene. Using these constructs, we demonstrate both poly(A) site occlusion in the 5' LTR and transcriptional interference in the 3' LTR. We have gone on to characterize the poly(A) site occlusion more thoroughly. First, the HIV poly(A) site functions efficiently when separated from its adjacent promoter and placed either at the 3' end of the α-globin gene or within the rabbit β-globin gene. However, the HIV poly(A) site does not function significantly when positioned immediately 3' to an active promoter, either its homologous HIV promoter or the α-globin promoter. Second, in contrast with the HIV poly(A) site, other poly(A) sites are active when adjacent to either the HIV or α-globin promoters. Our results suggest that the HIV poly(A) site and, in particular, its downstream GT-rich region possess specific sequence features that render it sensitive to transcription from an immediately adjacent promoter.
Occlusion of HIV poly(A) site

Results

Hybrid HIV LTR/α-globin gene constructs reproduce viral transcriptional regulation

To set up a convenient experimental system to investigate poly(A) site occlusion in the HIV LTR sequence, we constructed chimeric DNAs in which portions of the HIV LTR were fused to equivalent parts of the human α2-globin gene. Two such constructs were initially investigated, one containing the HIV LTR promoter and leader sequence in place of the α-globin promoter region (La) and the other containing a second copy of the HIV LTR sequence in place of the 3' end of the α-globin gene, including its poly(A) site (LaL). The HIV LTR sequence present in La and LaL starts from an Aval site (−160) in the 5'-flanking region and includes all of the HIV enhancer and promoter elements [Jones 1989]. The HinfI site (+128) at the 3' end of the LTR fragment immediately follows the GT-rich sequences that are part of the HIV poly(A) signal, as directly demonstrated by Bohnlein et al. (1989). The HIV LTR sequence is shown in Figure 1B while La and LaL are depicted in Figures 2A and 3A. La and LaL are within the transient expression vector pSved [Proudfoot et al. 1984], which contains both the SV40 origin of replication and enhancer sequences. We have recently discovered that the HIV promoter is activated by DNA replication [N.J. Proudfoot and J. Monks, in prep.]. We therefore transfected HeLa cells with either La or LaL pSved, together with a second plasmid pBSVpBR328. This plasmid expresses both T antigen, thereby allowing replication of the pSved plasmids, and rabbit β-globin mRNA, which provides a useful control for efficiency of transfection [Grosfeld et al. 1982].

Figure 2B shows RNA mapping experiments performed on cytoplasmic RNA purified from HeLa cells transfected with either Lo or LaL pSved. Panel I shows the RNase-protected bands obtained using a riboprobe as indicated in Figure 2A. La gave one band that corresponds to a transcript that initiates at the HIV cap site and then reads through the HIV poly(A) site into the α-globin gene beyond. The HIV poly(A) site therefore appears to be inactive in La (see below).

LaL, as well as giving a 5' LTR readthrough band, gave a second larger product of ≈250 nucleotides, which corresponds in size to transcripts reading into the 3' LTR sequence and then ending at the 3' LTR poly(A) site. No transcripts were detected that initiate at the 3' LTR promoter. This 3' LTR readthrough band would be 10 nucleotides larger than the 5' LTR readthrough band because of polylinker sequences present both in the riboprobe and 3' LTR sequence but absent from the 5' LTR sequence. To confirm the transcription pattern observed for the 3' LTR in LaL, a second RNA mapping experiment was carried out using an end-labeled DNA probe specific for the 3' LTR (labeled at an XbaI site in the polylinker sequence 5' to the 3' LTR; see Fig. 2A). By S1 analysis as shown in Figure 2B, panel II, LaL but not La gave a specific band of 260 nucleotides that corresponds to transcripts reading into the 3' LTR sequence and ending at the 3' LTR poly(A) site.

The results presented in Figure 2 indicate that the HIV
polylinker sequence not present in La. 

3' SI probe, a band of 260 nucleotides is detected in LaL, confirming the usage of the 3' LTR poly(A) site by transcripts initiating from the 5' LTR promoter. There are no bands detected in the construct La because the probe was labeled at a site in the polylinker sequence not present in La.

LTR/α-globin constructs reproduce both poly[A] site occlusion in the 5' LTR and transcriptional interference in the 3' LTR. In effect, we have a convenient system in which to investigate these two retroviral regulatory mechanisms. In the following sections we describe experiments that further define the molecular basis of poly[A] site occlusion in HIV.

**The HIV poly(A) site is an efficient 3' processing signal but inactive when part of the 5' LTR sequence**

The observed inactivity of the HIV poly(A) site when part of the 5' LTR [Lo] could be accounted for by the fact that it requires specific sequences 5' to the cap site. These sequences would be present in the active 3' LTR poly[A] site but not the 5' LTR poly[A] site. Indeed, such a sequence requirement has been demonstrated recently for the poly(A) site of the pararetrovirus hepatitis B (Rusnak and Ganem 1990) as discussed below. Alternatively, the HIV poly(A) site could be an intrinsically weak RNA-processing signal. These conclusions are based on the observation that if two identical poly(A) sites are placed in tandem, the 5' site is predominantly utilized (Gil and Proudfoot 1987), while if a stronger poly(A) site is positioned 3' to a weaker site, the stronger 3' poly(A) site predominates (Levitt et al. 1989). As predicted, aML gave no αA band but only the LpA band, consistent with the inactivity of the mutant α-globin poly(A) site.

To characterize the HIV poly(A) site further, we purified an AvHindIII fragment from the HIV LTR [see Fig. 1B] that contains only 7 nucleotides 5' to the AATAAA sequence with its 3' end directly following the GT-rich region. This fragment was inserted into the PvuII site in the 3'-flanking region of the α2-globin gene as in αWL to give the subclone αWL*, as well as in a quite different location in the third exon of the rabbit β-globin gene, 80 nucleotides 5' to the β-globin poly(A) site [BL]. As shown in Figure 3C, the shorter HIV poly(A) site fragment with an inactive poly[A] signal, isolated from a rare form of α-thalassemia (aML) [Higgs et al. 1983]. Figure 3B shows the RNA analysis of αWL and aML pSVed compared with α-pSVed, when transfected into HeLa cells. Using a 3' S1 probe from αWL (Fig. 3A), α-pSVed gave an S1 product of 230 nucleotides as expected for the 3' end of α-globin mRNA. In contrast, both aML and αWL gave a larger band of 360 nucleotides corresponding to utilization of the HIV poly(A) site. The fact that only 10% of mRNA utilized the α-globin poly(A) site, as compared with 90% at the HIV poly(A) site in αWL demonstrates that the HIV poly(A) site is a much stronger RNA-processing signal. These conclusions are based on the observation that if two identical poly(A) sites are placed in tandem, the 5' site is predominantly utilized [Gil and Proudfoot 1987], while if a stronger poly(A) site is positioned 3' to a weaker site, the stronger 3' poly(A) site predominates (Levitt et al. 1989).
works efficiently in competition with the α-globin poly(A) site. Thus, while La gives a strong αpA site S1 band, αWL* largely reads through the α-globin poly(A) site to yield an HIV poly(A) site band at ~90% level. In the clone βL it is evident that almost 50% of transcripts utilize the HIV poly(A) site with the remaining 50% reading through to the strong rabbit β-globin poly(A) site. Lane βL1 shows S1 analysis using a homologous EcoRI end-labeled probe. The βMM band detects mRNA from the cotransfected rabbit β-globin plasmid, which diverges from βL at the BglIII site shown in the diagram (Fig. 3C). As a control experiment, an EcoRI probe was used from the cotransfected rabbit β-globin plasmid, which gives two S1 bands (lane βL2): one corresponding to the β-globin poly(A) site and the other to transcripts from βL, which again diverge in sequence at the BglIII site. On the basis of these results we conclude that the HIV poly(A) site is more efficient than the human α-globin poly(A) but less efficient than the rabbit β-globin poly(A) sites. However, the fact that the minimal HIV poly(A) site functions in two quite different locations indicates that it does not possess an intrinsic
requirement for RNA sequences upstream of the HIV cap site.

Because it is evident that the HIV poly(A) site is efficient in isolation, we wished to confirm that it is wholly inactive or occluded in La. We therefore compared the levels of La mRNA 3' ends [utilizing the α-globin poly(A) site] with a similar construct, with La in which the HIV poly(A) site is deleted (ΔLa, Fig. 3A). Each transfection was controlled by the rabbit ß-globin mRNA cotransfection signal. As indicated by duplicate experiments in Figure 3B, Lo and ΔLa gave the same levels of mRNA 3' ends utilizing the ß-globin poly(A) site. That is, the presence or absence of the HIV poly(A) site had no effect on the level of transcripts reading through to the ß-globin poly(A) site. Figure 3D directly analyzes HIV poly(A) site utilization in Lo. Using either an S1 probe (panel I) or riboprobe (panel II) La gave no signal for utilization of the HIV poly(A) site in Lo, while aML gave a strong signal and therefore provided a positive control for the S1 and riboprobes.

From these experiments, we conclude that the HIV poly(A) site is an efficient RNA-processing signal, requiring only sequences between the AflIII and HindIII sites (i.e., within the 5' LTR transcript). However, this poly(A) site is completely inactive when it is part of the actively transcribing 5' LTR in La. These results therefore rule out an involvement of RNA signals 5' to the HIV cap site in HIV poly(A) site activation. Furthermore, the HIV poly(A) site is an efficient RNA-processing signal, consistent with the need for high levels of expression of the HIV genome.

The HIV poly(A) site is occluded by closely positioned promoters while other poly(A) sites remain active

We wished to investigate the specificity of HIV poly(A) site occlusion. Are other poly(A) sites occluded by closely positioned promoters and can the HIV poly(A) site be occluded by a different promoter? To test these possibilities we made a number of hybrid constructs as indicated in Figure 4A. La was modified to include a synthetic poly(A) site (Levitt et al. 1989) between the HIV promoter and poly(A) site are 70 nucleotides apart, whereas in aSa they are only 20 nucleotides apart. The AT deletion removes the Tar sequence, which is known to signal Tat transactivation of the HIV promoter (Cullen and Greene 1989). This sequence is thought to form a stable hairpin loop structure in the RNA and may interact directly with the Tat protein. The proximity of Tar to the HIV poly(A) site raises the possibility that it also has a role in poly(A) site occlusion. The ΔA deletion removes the AATAAA sequence of the HIV poly(A) site as described above. The α-globin promoter was also modified as indicated in Figure 5A so that either the HIV poly(A) site [αLo] or synthetic poly(A) site [αSa] was positioned directly adjacent to the α promoter. In αLo the promoter and poly(A) site are 70 nucleotides apart, whereas in αSa they are only 20 nucleotides apart.

Figure 4B shows the RNA analysis of these various hybrid constructs following their transfection into HeLa cells. In this experiment, RNA that reads through the promoter-adjacent poly(A) site is detected using an S1 probe for exon 3 of the α gene. As demonstrated before,
sequences [solid box], and important restriction sites are indicated. (B) 5' nuclelease analysis of RNA from HeLa cells transfected with the following plasmid constructs: α, Lo, LαL, LSo, LαTo, LΔαα, αLa, and ααα and cotransfected with RSVpBR328. The 5' DNA probes used in this assay were a BstEII digestion of an α-globin gene plasmid (α1 pSVed, Higgs et al. 1983). The α2 mRNA signal is at 110 nucleotides [plus some overdigestion S1 products], the position at which the α1 gene probe and α2 mRNA 3' ends diverge in sequence [Michelson and Orkin 1980]. The same EcoRI 5' S1 probe was used giving signals as indicated, the β larger band being the authentic β mRNA 3' end with the two bands below being overdigestion products (due to AT-rich sequence). The amount of α2-globin readthrough signal (a3') indirectly determines the level of utilization of the various inserted or deleted poly(A) sites in Lα or α, taking into account the β cotransfection control signal.

Lα has significant levels of α 3' signal because the poly(A) site is occluded. Similarly, LΔαα gives full levels of α readthrough although the transfection in this case was less efficient as judged by the fainter p cotransfection signal. Interestingly, LΔαTo also gave full levels of α 3' signal indicating that the Tar sequence is not involved in poly(A) site occlusion. LαL gave no α 3' signal. This is expected since the 3' end of the α gene is replaced by a second copy of the HIV LTR sequence in LαL. However, LSo also gave only low levels of α 3' signal, indicating that the SPA is active when positioned next to the HIV promoter.

Both αLa and ααα gave complementary results to the Lα and LSo gene constructs in that αLa had full levels of α 3' signal and ααα had greatly reduced levels.

From these results we can make the following conclusions. First, the HIV poly(A) site is occluded when adjacent to either the HIV or α-globin promoters and does not require the HIV Tar sequence for this effect. Second, the SPA is active when adjacent to either the HIV or α promoters as judged by significant reduction in the α 3' signal in LSo and αS. This latter conclusion (that the SPA is functional when positioned adjacent to the HIV or α-globin promoter) is based on lack of readthrough transcription and is therefore indirect.

We have attempted to identify the tiny transcripts, which we predict must exist, that initiate on the α or HIV promoters and end at the closely following poly(A) site [SPA]. We have detected these ~50 nucleotide transcripts at very low levels [data not shown] and presume that their small size, proximity of cap, and poly(A) tail, as well as lack of introns, makes them highly unstable.

We therefore made additional constructs to test whether the lack of readthrough transcription observed when a non-HIV poly(A) site is positioned close to a promoter is due to the presence of a functional poly(A) signal [see Fig. 5A]. First, we have placed the SPA in LSo both in the forward [LS + α] and reverse [LS - α] orientations. In the reverse orientation the SPA still possesses an A2TA3 sequence, but lacks GT-rich downstream signals required for poly(A) site activity [Gil and Proudfoot 1987; Levitt et al. 1989]. We have also cloned an α2-globin poly(A) site fragment in the same place as SPA in LSo but have used both the wild-type α-globin poly(A) site [LαWα] as well as a point mutation [A2TA3 → A2TA3G] derived from α-thalassemia, which is a completely inactive poly(A) site [LαMα] [Higgs et al. 1983]. Figure 5B shows the RNA analysis of these constructs. First, it is clear that the SPA in LSo is orientation specific as judged by a fivefold reduction in readthrough α transcripts for LS + α but full levels for LS - α. With the Lα constructs, both LαM + α and LαM - α [with the α-thalassemia poly(A) site fragment in reverse orientation] gave full levels of α readthrough transcripts. In contrast, LαWα drastically reduced the level of readthrough transcript. These results demonstrate that both a functional A2TA3 sequence, as well as downstream GT-rich sequences, are required to prevent readthrough transcription. That is, the same elements that are required for a functional poly(A) site are also required to block readthrough transcripts. We conclude that both the SPA and α-globin poly(A) sites are functional when adjacent to the HIV promoter while the HIV poly(A) site is fully occluded.

Figure 4. (A) Schematic representation of Lα derivative constructs introducing (1) the insertion of a functional synthetic poly(A) site [SPA] (Levitt et al. 1989) into the LTR, adjacent to the HIV LTR promoter region; (2) two deletions within the HIV LTR, by excision through restriction enzyme digestion of, first, the HIV Tar sequence [ΔT] (Cullen and Greene 1989) and second the poly(A) site itself [ΔA]. Also illustrated are derivative α2-globin constructs in which the SPA as well as the BglII–Hinfl LTR fragment, containing the HIV LTR poly(A) site, have been inserted into the α2-globin derivative vector, 30 bp adjacent to the α2-globin promoter. Exons [hatched boxes], introns [open boxes], HIV fragments [dashed boxes], flanking sequences [thin boxes], promoters [arrows], poly(A) sites [labeled]; polylinker

GENES & DEVELOPMENT 249
not itself confer poly(A) site occlusion, whereas the GT-rich region was a likely candidate. We therefore made synthetic oligonucleotides containing a 5′-linker sequence followed by the HIV GT-rich sequence as shown in Figure 6A. This duplex oligonucleotide was inserted into the plasmid LS + α immediately following the AATAAA of the SPA in place of its downstream GT-rich sequence [called LSHα]. A second construct was then made in which the hybrid SH poly(A) site was excised from LSHα and inserted downstream of the intact α-globin genes so that the α and SH poly(A) sites are positioned in tandem [called αSH]. The details of how LSHα and LSH were constructed are described in Materials and methods.

Figure 6B shows RNA analysis of LSHα and αSH, compared with La and LS + α, transfected into HeLa cells. Using an end-labeled DNA probe for the 3′ end of the α-globin gene [as in Fig. 3A], full levels of α-globin poly(A) site utilization are evident in the La lane, indicating occlusion of the HIV poly(A) site, while at least a fivefold reduction in the α-globin poly(A) site signal [as judged by the β cotransfection control signal] is observed in LS + α, indicating that the SPA is active adjacent to the HIV promoter as shown previously [Fig. 5B]. Analysis of LSHα reveals full levels of α-globin poly(A) site utilization [as in La], indicating that the SH poly(A) site is inactive when next to the HIV promoter unlike the SPA [in LS + α]. However, the SH poly(A) site is functional when placed at the 3′ end of α as in αSH. As shown in Figure 6B, the SH poly(A) site band is at least twice as

Figure 5. [A] Nomenclature and schematic illustrations of La-derivative constructions, into which were inserted the SPA in both orientations [described previously in one orientation as αSa in Fig. 5A], the wild-type α2-globin poly(A) site, and the nonfunctional point-mutated α2-globin poly(A) site in both orientations. All poly(A) site inserts are positioned closely adjacent to the HIV LTR promoter. [B] S1 analysis of RNA from HeLa cells transfected with La, LS + α, LS - α, LoM + α, and LoMα. The cells were cotransfected with RSVpBR328. The S1 DNA probe used in this assay was obtained from a BsrII digestion of α and labeled with Klenow enzyme. The utilization of the inserted poly(A) sites is indirectly measured through α-globin readthrough level. It is important to take into account the quality of transfections based on the β cotransfection control.

The GT-rich downstream sequence of the HIV poly(A) site confers poly(A) site occlusion

We wished to determine which part of the HIV poly(A) site causes it to occlude when close to an active promoter. It has been demonstrated previously that only two specific sequences are required to form an efficient poly(A) site: AATAAA and a downstream GT-rich region [Levitt et al. 1989]. The AATAAA sequence is known to be a general feature of mammalian poly(A) sites although in several cases, variants such as ATATAA and AGTAAAA do exist [Wickens and Stephenson 1984]. In contrast, the downstream GT-rich sequences are relatively variable between poly(A) sites, although more extensive GT-rich regions are often associated with more efficient poly(A) sites. On the basis of these considerations we reasoned that the AATAAA of the HIV poly(A) site could
strong as α. These results demonstrate that the downstream GT-rich region of the HIV poly(A) site confers poly(A) site occlusion properties on a poly(A) site.

Only large-scale spacing between the HIV promoter and poly(A) site prevents poly(A) site occlusion

We have demonstrated that the HIV poly(A) site is inactive when very close to an active promoter (as in La or αLa) but highly active when positioned >1 kb away from a promoter (as in LoL, αL, and βL). This behavior of the HIV poly(A) site is in contrast to the α and synthetic poly(A) sites, which are both active adjacent to promoters. We wished to define more precisely how large a space is required between the promoter and HIV poly(A) site to overcome the occlusion effect. We, therefore, constructed variants of La in which increasingly large spacer fragments (derived from phage λ DNA) were inserted into the 5′ flanking site between the HIV promoter and the poly(A) site (Fig. 7A). Figure 7B shows the S1 analysis of La RNA probing for α-globin poly(A) site utilization. As indicated with a spacer of 75 or 240 nucleotides, full levels of readthrough transcription as compared with La were observed. Interestingly, with spacers of 310 or 420 nucleotides, a significant drop in the level of αpA signal was demonstrated, suggesting that the HIV poly(A) site was now active. However, when the 5′ end of the La constructs was analyzed no detectable HIV poly(A) site band was observed (data not shown), suggesting that these larger spacer constructs were producing unstable mRNAs rather than activating the HIV poly(A) site. To confirm this view, a deleted La construct was tested in which the HIV poly(A) was removed ([LΔAα, Fig. 4A]. The spacers of 240, 310, and 420 nucleotides were added to LΔAα, and α-globin poly(A) site utilization was measured as before. Figure 7B shows that the spacers of 310 and 420 nucleotides in LΔAα still gave significantly reduced levels of αpA signal, confirming that these larger inserts either inhibited the HIV promoter or had reduced mRNA stability, rather than activating the HIV poly(A) site. Indeed, the ratio of αpA signal between LΔAα and its two spacer clones 310 and 420 is close to the αpA ratio for La and its two spacer clones 310 and 420. This suggests that these larger inserts still have no effect on increasing the usage of the HIV poly(A) site. We conclude that the HIV poly(A) site is still occluded by an upstream promoter 500 bp away, but at some point beyond this distance it becomes fully active.

Discussion

The results presented in these studies argue that the HIV poly(A) site possesses specific sequence features that cause it to be occluded when close to a functional promoter. Thus, we demonstrate that the HIV poly(A) site is fully occluded when adjacent to either its homologous HIV promoter or the α2-globin gene promoter. In contrast, two other poly(A) sites are functional when adjacent to these two promoters. These observations are in contrast to two recently described pararetroviral poly(A) sites, one in the ground squirrel hepatitis B virus (Russnak and Ganem 1990) and the other in the plant cauliflower mosaic virus (CaMV; Sanfaçon and Hohn 1990). Pararetroviruses are transcribed as circular minichromosomes and possess a single LTR sequence. Transcription initiates on the LTR promoter proceeding through the adjacent poly(A) site and around the whole viral genome, and finally utilizes the LTR poly(A) site the second time around. Pararetroviruses are therefore functionally analogous to retroviruses and require a mechanism to occlude the LTR poly(A) site. In the case of the hepatitis B virus, sequences upstream of the cap site are required for the poly(A) site to function. Furthermore, LTR sequences from other retroviruses including HIV can apparently function instead of the hepatitis upstream sequences to activate the hepatitis poly(A) site [Russnak and Ganem 1990]. In contrast, the CaMV appears to only partially occlude its promoter-adjacent poly(A) site and does so by
a simple spacing effect. This poly(A) site works at 66% efficiency on transcripts initiating on the adjacent promoter (190 bp upstream) but works with 100% efficiency when artificially separated from the promoter by ~300 nucleotides (Sançafon and Hohn 1990). A second plant poly(A) site was tested in this system (for the nopaline synthetase gene, Nos) and was similarly more active when moved away from the CaMV promoter. This suggests that the sensitivity of poly(A) sites to nearby promoters may be a more general feature of plants rather than specific to pararetroviruses.

The occlusion mechanism for the HIV poly(A) site is surprisingly different from both of the above pararetroviruses. Upstream sequences to the cap site are not required for it to work efficiently. Indeed, a minimal HIV poly(A) site fragment of 60 nucleotides functions in two different locations (both in the 3' flanking region of the human α2-globin gene and in the third exon of the rabbit β-globin gene). The possibility that both locations inadvertently possess upstream sequence elements that activate the HIV poly(A) site seems improbable. Similarly, a simple spacing effect does not account for HIV poly(A) site occlusion. The fact that poly(A) site occlusion appears to be specific to the HIV poly(A) site and not to two other mammalian poly(A) sites immediately distinguishes the HIV poly(A) site from the situation in CaMV. Also, because the HIV poly(A) site is completely occluded adjacent to a promoter and is only active when positioned > 500 bp away, a much more efficient poly(A) site occlusion mechanism is indicated. We are therefore forced to conclude that in each of these three cases of poly(A) site occlusion, different mechanisms must operate. Possibly, the very different origins of these three viruses have led to the convergent evolution of these different approaches to poly(A) site occlusion.

These studies demonstrate that the HIV poly(A) site occlusion sequences must be directly associated with the poly(A) signal itself. First, we show that the sequences between the AII and HindIII sites which border the A2TA3 and GT-rich elements are required for occlusion. Second, we demonstrate that 26 nucleotides of the HIV poly(A) site downstream region including two GT-rich regions confer poly(A) site occlusion properties on a synthetic poly(A) site. We are currently attempting to define these downstream sequences more precisely, as well as to test whether or not other retroviral poly(A) sites possess similar occlusion sequences. Such information may allow us to address the molecular mechanism of this intriguing phenomenon.

Materials and methods

Plasmid Constructions

α, La, aLa, LaL. The human α2-globin gene, containing plasmid called α throughout these studies, was originally called α2W3′PS pSved, as described by Whitelaw and Proudfoot [1986], and contains the α gene in three exons with 1.1 kb of 5′-flanking sequence and 2.4 kb of 3′-flanking sequence. The pSved vector is based on pBR322 but contains the SV40 origin/enhancer region to allow expression of the α-globin gene in mammalian cells [Proudfoot et al. 1984; Whitelaw et al. 1989]. Digestion of α with Smal excises ~860 bp of a 5′-flanking region through to the middle of intron 1 of the α-globin gene and thereby deletes the entire α-globin promoter. In place of the α-globin promoter region, a 300-bp Aval–HindIII fragment containing the HIV-1 promoter and leader sequence [Fig. 1B] was inserted to generate Lo. Lo contains a shorter HIV fragment [Aval–BglIII] and therefore lacks most of the HIV leader sequence including the poly(A) signals. LoL was constructed by excising a 700-bp BstEII fragment, including the third exon of the α-globin gene and the immediate 3′-flanking region from Lo, and replacing this with the HIV-1 Aval–HindIII promoter leader sequence. In this case, the HIV-1 insert derives from pSP65 HA as an EcoRI–SalI fragment and therefore contains polylinker sequence on either side of the HIV insert.

pSP65–HIV LTR HIV-1 LTR fragments Aval–HindIII and BglII–HindIII were inserted into the Smal site of pSP65 [Melton et al. 1984] to generate pSP65 HIV HA and HB.

αWL, BgL, aWL*, BL α and a poly(A) site point mutation of α (called α2M3′PSpSved, Whitelaw and Proudfoot 1986) were linearized with PvuII, which cuts 100 bp 3′ to the poly(A) site, and either the HIV-1 poly(A) site fragment of 110 bp, BglII–HindIII, was inserted to yield αWL or αML, or a shorter 60-bp AII–HindIII poly(A) site fragment was inserted to yield αWL*. The same AII–HindIII HIV poly(A) site was inserted into the BglII site in the middle of the third exon of the rabbit β-globin gene in the plasmid BS/pBR328 (Grosveld et al. 1982; Levitt et al. 1989), to generate BL.

LaL, LAAa, LATa, LSA–α, LSWa, LAa, LAM–α. Sau3AI fragments from phage λ of 75, 240, 310, and 420 bp were inserted into the BglII site between the HIV promoter and poly(A) site in Lo and LAAa to generate LaL, LAAa, LTa and LLAa. LTa and LLAa were constructed by deleting HIV LTR sequences between BglIII and AII (ΔT) and AII and HindIII (ΔA) in Lo. A synthetic poly(A) site (SPA) of 54 bp [Levitt et al. 1989] was inserted into the HIV BglII site in both orientations in La to generate LS + α and LS − α. Finally, an EcoO109 fragment of 130 bp, containing either the poly(A) site of the α2-globin gene or the same mutated poly(A) site [both orientations], was inserted into the same BglII site in La to generate LaWa, LaWa + α, and LaM − α.

αLa, aSaα. A functional promoter fragment from the Smal site at −90 to the HindIII site was cloned into the α2-globin gene also containing some polylinker sequence including the BamHI site in pUC9 immediately following the α-globin + 20 HindIII site) was inserted into Smal-cut α. This yielded an α-globin gene construct lacking most of exon 1 and half of intron 1 but containing a polylinker site into which either the SPA or HIV LTR BglII–HindIII fragments were inserted to generate aSa and αLa, respectively.

LSHα, αSH. Duplexed synthetic oligonucleotides [for detailed sequence, see Fig. 6A] were blunt-end-ligated into LS + α cut at a BglII site immediately 3′ to the SPA AATAAA sequence and a HindIII site in exon 2 of the α-globin gene to form LSHα. αSH was made by purifying an SH poly(A) site fragment excised from LSHα by cutting at the PvuII site [just 5′ to the HIV cap site] and a HindII site immediately 3′ to the HindIII site in exon 2 of α-globin. This fragment was inserted into the 3′-flanking region of α at the PvuII site.

In all the plasmid constructions described above, the presence and orientation of inserts were determined by restriction enzyme analysis. The isolation of vector and insert DNA fragments was obtained by either agarose or polyacrylamide gel electrophoresis purification, depending on the size of the DNA fragment in question. If vector or insert DNA fragments were rendered blunt-ended, this was achieved by filling in ends with Klenow DNA polymerase enzyme and dNTP. Ligation was
performed at room temperature, and transformations were carried out in the NM554 strain of Escherichia coli.

**Cell transfection and RNA analysis**

The methods used for growth and transient transfection of HeLa cells, isolation of cytoplasmic RNA, and S1 nuclease analysis have been described previously (Johnson et al. 1986; Levitt et al. 1989). All HeLa cell transfections with the various HIV-α-plasmids were cotransfected with a second plasmid RpSVpBR328. This plasmid has the dual function of providing SV40 T antigen to allow replication of all the HIV-α-plasmids (which contain SV40 origins of replication) as well as rabbit β-globin mRNA, which is used as internal control in most of the RNA mapping experiments described in these studies.

RNAse protection analysis of cytoplasmic RNA was carried out as described by Melton et al. (1984), with the following modifications of conditions: Hybridizations were performed at 51°C and RNase A and T1 digestions were done at 16°C. All S1 nuclease or RNase digestion products were electrophoresed through 4–8% polyacrylamide/7 M urea sequencing gels with a variation in the percent of polyacrylamide, depending on the size range of the products expected. Size markers were generated by digestion of sequenced plasmids with various restriction enzymes.

**Acknowledgments**

C.W. is funded by an AIDS MRC studentship and by a grant from The Joseph Goerres Gesellschaft. These studies were funded by a Wellcome project grant (16354) to N.J.P. We gratefully acknowledge William James and Tim Humphrey for advice and encouragement throughout these studies.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

**References**

Bateman, E. and M. Paule. 1988. Promoter occlusion during rRNA transcription. *Cell* 54: 985–992.

Bohnlein, S., J. Hauber, and B.R. Cullen. 1989. Identification of a US-specific sequence required for efficient polyadenylation within the human immunodeficiency virus long terminal repeat. *J. Virol.* 63: 421–424.

Cullen, B.R. and W.C. Greene. 1989. Regulatory pathways governing HIV1 replication. *Cell* 58: 423–426.

Cullen, B.R., P.T. Lomedico, and G. Ju. 1984. Transcriptional interference in avian retroviruses—Implications for promoter insertion model of leukaemogenesis. *Nature* 307: 241–245.

Gil, A. and N.J. Proudfoot. 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA 3' end formation. *Cell* 49: 399–406.

Grosveld, G.C., E. deBoer, C.K. Shewmaker, and R.A. Flavell. 1982. DNA sequences necessary for transcription of the rabbit β-globin gene in vivo. *Nature* 295: 120–126.

Henderson, S.L., K. Ryan, and B. Sollner-Webb. 1989. The promoter-proximal rRNA termination augments initiation by preventing disruption of the stable transcription complex caused by polymerase read-in. *Genes & Dev.* 3: 212–223.

Higgs, D.R., S.E.Y. Goodbourn, J. Lamb, J.B. Clegg, D.J. Weatherall, and N.J. Proudfoot. 1983. α-Thalassaemia caused by a polyadenylation signal mutation. *Nature* 306: 398–400.

Humphrey, T. and N.J. Proudfoot. 1988. A beginning to the biochemistry of polyadenylation. *Trends Genet.* 4: 243–245.

Johnson, M., C. Norman, M.A. Reeve, J. Scully, and N.J. Proudfoot. 1986. Tripartite sequences within and 3' to the sea urchin H2A histone gene display properties associated with a transcriptional termination process. *Mol. Cell. Biol.* 6: 4008–4018.

Jones, J.A. 1989. HIV trans-activation and transcriptional control mechanisms. *The New Biologist* 1: 127–135.

Ju, G. and B.R. Cullen. 1985. The role of avian retroviral LTRs in the regulation of gene expression and viral replication. *Adv. Virus Res.* 30: 179–223.

Levitt, N., D. Briggs, A. Gil, and N.J. Proudfoot. 1989. Definition of an efficient poly(A) site. *Genes & Dev.* 3: 1019–1025.

Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12: 7035–7056.

Michelson, A.M. and S.H. Orkin. 1980. The 3' untranslated regions of the duplicated human α-like globin genes are unexpectedly divergent. *Cell* 22: 371–377.

Proudfoot, N.J., T.R. Rutherford, and G.A. Partington. 1984. Transcriptional analysis of human zeta globin genes. *EMBO J.* 3: 1533–1540.

Ratner, L.A., L.L. Fisher, M. Jagodzinski, H. Mitsuya, R.-S. Liou, R.C. Gallo, and F. Wong-Staal. 1987. Complete nucleotide sequences of functional clones of the AIDS virus. *AIDS Res.* 3: 57–69.

Russnak, R. and D. Ganem. 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes & Dev.* 4: 764–776.

Sanfaçon, H. and T. Hohn. 1990. Proximity to the promoter inhibits recognition of cauliflower mosaic virus polyadenylation signal. *Nature* 346: 81–84.

Varmus, H. 1988. Retroviruses. *Science* 240: 1427–1435.

Whitelaw, E. and N.J. Proudfoot. 1986. α-Thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human γ-globin gene. *EMBO J.* 5: 2915–2922.

Whitelaw E., P. Hogben, O. Hanscombe, and N.J. Proudfoot. 1983. α-Thalassaemia caused by a US-specific sequence required for efficient polyadenylation in the human 4-bit 3 globin gene. *Nature* 305: 81–84.

Henderson, S.L., K. Ryan, and B. Sollner-Webb. 1989. The promoter-proximal rRNA termination augments initiation by preventing disruption of the stable transcription complex caused by polymerase read-in. *Genes & Dev.* 3: 212–223.

Higgs, D.R., S.E.Y. Goodbourn, J. Lamb, J.B. Clegg, D.J. Weatherall, and N.J. Proudfoot. 1983. α-Thalassaemia caused by a polyadenylation signal mutation. *Nature* 306: 398–400.
Occlusion of the HIV poly(A) site.

C Weichs an der Glon, J Monks and N J Proudfoot

*Genes Dev.* 1991, 5:
Access the most recent version at doi:10.1101/gad.5.2.244

References

This article cites 23 articles, 8 of which can be accessed free at:
http://genesdev.cshlp.org/content/5/2/244.full.html#ref-list-1

License

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.