MOLECULAR AND CELLULAR BIOLOGY, Sept. 1992, p. 3699–3705
0270-7306/92/093699-07$02.00/0

Vol. 12, No. 9

Elements Upstream of the AAUAAA within the Human Immunodeficiency Virus Polyadenylation Signal Are Required for Efficient Polyadenylation In Vitro

ALEXANDRA VALSAMAKIS, NANCY SCHEK, AND JAMES C. ALWINE*

Molecular Biology Graduate Group, Department of Microbiology, School of Medicine, University of Pennsylvania, 560 Clinical Research Building, 422 Curie Boulevard, Philadelphia, Pennsylvania 19104-6142

Received 17 April 1992/Returned for modification 28 May 1992/Accepted 9 June 1992

Recent in vivo studies have identified specific sequences between 56 and 93 nucleotides upstream of a polyadenylation [poly(A)] consensus sequence, AAUAAA, in human immunodeficiency virus type 1 (HIV-1) that affect the efficiency of 3′-end processing at this site (A. Valsamakis, S. Zeichner, S. Carswell, and J. C. Alwine, Proc. Natl. Acad. Sci. USA 88:2108–2112, 1991). We have used HeLa cell nuclear extracts and precursor RNAs bearing the HIV-1 poly(A) signal to study the role of upstream sequences in vitro. Precursor RNAs containing the HIV-1 AAUAAA and necessary upstream (U3 region) and downstream (U5 region) sequences directed accurate cleavage and polyadenylation in vitro. The in vitro requirement for upstream sequences was demonstrated by using deletion and linker substitution mutations. The data showed that sequences between 56 and 93 nucleotides upstream of AAUAAA, which were required for efficient polyadenylation in vivo, were also required for efficient cleavage and polyadenylation in vitro. This is the first demonstration of the function of upstream sequences in vitro. Previous in vivo studies suggested that efficient polyadenylation at the HIV-1 poly(A) signal requires a spacing of at least 250 nucleotides between the 5′ cap site and the AAUAAA. Our in vitro analyses indicated that a precursor containing the defined upstream and downstream sequences was efficiently cleaved at the polyadenylation site when the distance between the 5′ cap and the AAUAAA was reduced to at least 140 nucleotides, which is less than the distance predicted from in vivo studies. This cleavage was dependent on the presence of the upstream element.

The 3′ ends of most higher eukaryotic mRNAs are produced by a tightly coupled endonucleolytic cleavage and polyadenylation reaction (reviewed in reference 41). This processing occurs 10 to 30 nucleotides downstream of the consensus sequence AAUAAA (32), which has been shown to be required for cleavage in vivo (14, 27, 42) and for cleavage and polyadenylation in vitro (44). However, poly(A) signals consist of sequences in addition to AAUA AA. Additional elements upstream and downstream of AAU AAA have been shown to control the efficiency of utilization of the AAUAAA as a poly(A) signal. Hence, not all AAUA AA elements appearing in a message will be efficiently utilized for 3′-end processing.

Downstream elements (DSEs) have been shown to affect polyadenylation efficiency both in vivo (4, 8, 9, 17, 19, 23–25, 35, 36, 46, 47) and in vitro (20, 34). DSEs do not conform to a consensus sequence but instead are either G+U or U rich. The function of these elements may be subject to spatial constraints, since insertions which increase distance between AAUAAA and DSEs decrease signal efficiency (18, 21, 24). Additionally, human T-cell leukemia virus type I, which has more than 250 nucleotides between the AAUAAA and DSE, has an RNA secondary structure that juxtaposes the two elements (1, 3).

Recently, in vivo studies have identified additional regulatory elements upstream of AAUAAA (upstream elements [USEs]) in a variety of viral poly(A) signals, including simian virus 40 (SV40) late (7), ground squirrel hepatitis virus (33), adenovirus type 2 major late transcription unit (10, 12), cauliflower mosaic virus (37), and human immunodeficiency virus type 1 (HIV-1) (6, 11, 39). Nonhomologous USEs from different poly(A) signals are functionally similar. For example, the USEs from the HIV-1 poly(A) signal can replace the USE of SV40 late (39) and ground squirrel hepatitis virus poly(A) signals (33). Further, the existence of homology between some USEs and DSEs suggests functional similarity (39). In contrast to the above-mentioned finding for the HIV polyadenylation signal, the in vivo studies of Weichs an der Glon et al. (40) have suggested that there are no USEs in the HIV polyadenylation signal. To counter the doubts cast by these studies, we have used in vitro polyadenylation experiments to determine whether USEs are utilized in vitro and to confirm and extend our previous in vivo results.

In vitro polyadenylation systems using HeLa cell nuclear extracts have been shown to faithfully reproduce in vivo events (29, 30), including the requirement for DSEs (19, 34, 46, 47). In this investigation, we have studied the utilization of the HIV-1 poly(A) signal in vitro to establish the requirement for USEs. Using deletion and linker substitution (LS) mutants, we demonstrate that the USEs previously found to be required for maximal 3′-end processing in vivo (39) are also required for cleavage and polyadenylation in vitro. In vivo studies (40) have also suggested that the distance between the 5′ cap and the AAUAAA is critical for the efficient utilization of the HIV-1 poly(A) signal. We show here that precursor RNAs with relatively short distances between the 5′ cap and the AAUAAA are efficiently processed in vitro and that this processing depends on the presence of the USE. Hence, an in vivo effect of proximity to the 5′ cap site must occur by a mechanism other than simply distance.

* Corresponding author.

MATERIALS AND METHODS

Plasmids and transcripts. Substrate RNAs for in vitro polyadenylation reactions were prepared by in vitro transcription from plasmids containing the wild-type HIV-1 poly(A) signal or LS mutants within the upstream region of the polyadenylation signal. These mutants were constructed by polymerase chain reaction (PCR) amplification of HIV-1 nucleotides 9463 to 9727 and subcloning of amplified fragments into pGEM3Zf+. Plasmid pU3R+39 (5) was used as the PCR template for the wild-type HIV-1 polyadenylation signal. Derivatives of pU3R+39 bearing linker substitutions (*Nde*I-*Xho*I-*Sal*I [NXS] 4, 5, 6, 7, 8, 9, and 10) were used as templates for USE mutants. These derivatives were constructed by replacing a 568-bp *Bgl*II-*Hin*dIII region of pU3R+39 with 620-bp *Bam*HI-*Hin*dIII fragments from eight LS mutant chloramphenicol acetyltransferase plasmids (39, 45) in which 18 bp of wild-type HIV-1 sequences between HIV-1 9481 and 9606 were sequentially substituted with an 18-bp NXS polylinker. Primers used for the PCR amplification of wild-type and LS mutant polyadenylation signals were complementary to pU3R+39 nucleotides 2190 to 2206 and HIV-1 nucleotides 9463 to 9480. The latter primer encoded an *Eco*RI site at its 5′ end to facilitate subcloning. The 342-bp PCR products were digested with *Eco*RI and *Xba*I (*Xba*I is located 3′ to HIV-1 9727 in pU3R+39 and the LS derivatives). The resulting 275-bp fragments were subcloned into *Eco*RI- and *Xba*I-digested pGEM3Zf+.

Deletion mutants Δ4-8, Δ4-9, and Δ4-10 were constructed by ligating 150-bp *Xho*I-*Xba*I fragments of mutants NXS 8, 9, and 10 to the 3,350-bp *Xho*I-*Xba*I fragment of NXS 4. Deletion Δ9-10 was constructed by ligating the 150-bp *Xho*I-*Xba*I fragment from NXS 10 to the 3,350-bp *Xho*I-to-*Xba*I fragment of NXS 9. Deletion Δ*Bgl*II was made by insertion of a 115-bp *Bgl*II-*Xba*I (HIV 9612 to 9727) fragment from pU3R+39 into pGEM3Zf+ between *Bam*HI and *Xba*I.

Templates for in vitro transcription of polyadenylation substrate RNAs were linearized with *Xba*I. Reactions were performed under standard conditions (2), using T7 polymerase (Promega Biotec), 50 μCi of [^{32}P]UTP (Amersham), 1 μg of linearized template, and 0.5 mM ^{7}MeGTP (Pharmacia), which allows 5′ capping of synthesized RNAs. ^{32}P-labeled RNAs were extracted with phenol-chloroform-isoamyl alcohol (50:49:1), ethanol precipitated, and purified by electrophoresis through a 5% polyacrylamide–7 M urea gel. RNAs were eluted from gel slices at room temperature for 16 h with 400 mM NaCl–20 mM Tris (pH 7.5)–0.1% sodium dodecyl sulfate. After extractions with phenol-chloroform-isoamyl alcohol and chloroform-isoamyl alcohol (99:1), RNAs were ethanol precipitated. [^{32}P]UTP incorporation was quantitated by liquid scintillation counting.

Nuclear extracts and in vitro polyadenylation-cleavage reactions. HeLa cell nuclear extracts for in vitro polyadenylation were prepared as described elsewhere (28). In vitro polyadenylation reaction mixtures contained 14.5 μl (60% by volume) of HeLa nuclear extract, 1 mM ATP (Pharmacia), 20 mM creatine phosphate (Sigma), 2.6% polyvinyl alcohol, and 3 ng of ^{32}P-labeled substrate RNA (5×10^7 cpm/μg). Cleavage conditions were similar except that the ATP concentration was reduced to 250 μM and 1 mM cordycepin (Boehringer Mannheim) was added. All reactions were performed at 30°C for 30 min. At the end of the incubation, mixes were brought to 300 μl with 50 mM Tris (pH 8.0)–10 mM EDTA–10 mM NaCl, extracted twice with phenol-chloroform-isoamyl alcohol, and precipitated with ethanol. RNAs were analyzed on 40-cm 5% polyacrylamide–7 M urea gels.

Polyadenylated and cleaved products were quantitated with a Molecular Dynamics Phosphorimager. Quantitated values for LS and deletion mutant RNAs were adjusted to account for U content relative to that of wild-type RNA. Percent processed RNA was calculated as processed RNA counts divided by processed plus unprocessed RNA counts.

RNase protection of polyadenylated RNA. ^{32}P-labeled polyadenylated products (337 to 437 nucleotides) of in vitro polyadenylation reactions were gel isolated as described above. The isolated RNA was hybridized to a 200-fold molar excess of unlabeled complementary RNA probe for 14 h at 45°C under standard conditions (2). The 685-nucleotide unlabeled probe was synthesized in vitro as described above and contained HIV-1 nucleotides 9612 to 9727 plus 570 nucleotides of non-HIV-1 sequence at its 5′ end. Hybridization products were digested with RNases T_1 (Bethesda Research Laboratories) and A (Sigma) and then with proteinase K (Sigma), using standard protocols (2). After phenol-chloroform-isoamyl alcohol extraction and ethanol precipitation, protected RNAs were resolved on an 8% polyacrylamide–7 M urea gel.

RESULTS

The HIV-1 polyadenylation signal is functional in vitro. To establish that the HIV-1 polyadenylation signal could be properly processed in vitro, a 280-nucleotide precursor RNA was synthesized by using T7 polymerase and the template shown in Fig. 1. This RNA contains HIV-1 nucleotides 9463 to 9727, including the AAUAAA, the downstream G-U element (5), and all of the upstream U3 sequences previously shown to be required for 3′-end processing in vivo (6, 11, 39). Using standard in vitro polyadenylation conditions with 3 ng of ^{32}P-labeled RNA and 60% nuclear extract by volume, 30% of the 280-nucleotide wild-type HIV-1 RNA substrate was polyadenylated during the 30-min reaction (Fig. 2A). Polyadenylation resulted in the appearance of RNAs migrating between 337 and 437 nucleotides (PA in Fig. 2A).

Precursor RNAs can be polyadenylated in vitro at their 3′ ends without undergoing cleavage (38). In Fig. 2A, such end-polyadenylated precursor RNAs would migrate similarly to properly cleaved and polyadenylated RNA (337 to 437 nucleotides). To confirm that the in vitro-polyadenylated RNAs were in fact cleaved prior to polyadenylation, their 3′ ends were mapped by RNase protection. The 337- to 437-nt polyadenylated species in Fig. 2A were isolated from the gel and hybridized with an unlabeled 685-nucleotide RNA probe. As shown in Fig. 2C, 79- and 121-bp RNase-resistant hybrids would result from the properly cleaved and the uncleaved, end-polyadenylated RNAs, respectively. Figure 2B (lane 1) shows that the isolated in vitro-polyadenylated RNA produced only the 79-bp RNase-resistant hybrid, indicating that proper cleavage occurred prior to polyadenylation. In contrast, precursor RNA formed the expected 121-bp RNase-resistant hybrid (Fig. 2B, lane 2).

U3 sequences are required for in vitro processing. Previous studies (6, 11, 39) indicate that U3 sequences upstream of the AAUAAA are required for efficient 3′-end processing in vivo. To test whether these sequences were important for polyadenylation efficiency in vitro, the level of processing of wild-type precursor RNA was compared with that of a precursor RNA lacking the significant U3 sequences (mutant Δ*Bgl*II; Fig. 1). The only *cis*-regulatory elements encoded by

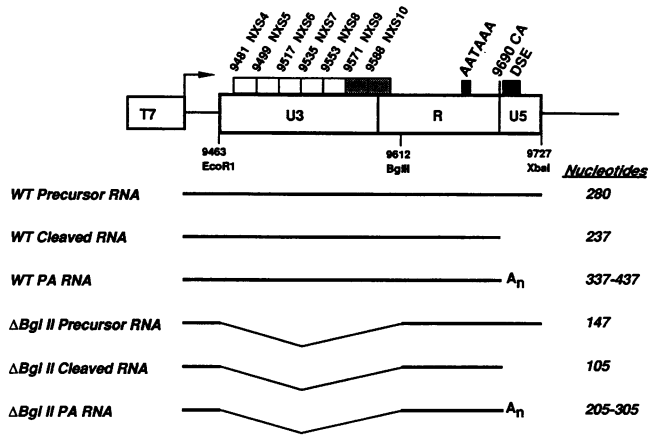

FIG. 1. HIV-1 polyadenylation signal precursor RNAs and expected products of in vitro cleavage and polyadenylation reactions. The wild-type polyadenylation signal containing all sequences required for efficient 3'-end formation in vivo (nucleotides 9463 to 9727 [numbering according to the provirus sequence]) is shown as it appears in a pGEM vector from which precursor RNAs can be transcribed for in vitro analysis. The AAUAAA lies between nucleotides 9665 and 9770; the cleavage site, at the dinucleotide -CA-, is located 25 nucleotides downstream; the DSE is in the U5 region (9691 to 9715 [5]); the region upstream of the AAUAAA contains all of the USEs mapped in vivo (39). The boxes above U3 indicate the positions of LS mutations used to define the USE. The two shaded boxes indicate the LS mutations which affected polyadenylation in vivo (39) and in vitro (these studies). In each LS mutant, 18 nucleotides of wild-type sequence were replaced with an 18-bp NXS polylinker (45). Mutants are numbered in order (NXS 4 to NXS 10). The nucleotide number above each box is the first substituted nucleotide of each LS mutant. Thin lines represent pGEM vector sequences. An *Eco*RI site was inserted upstream of nucleotide 9463 in the wild-type and NXS templates for cloning purposes. In the Δ*Bgl*II mutant, the HIV sequences upstream of the *Bgl*II site (nucleotide 9612) have been deleted. Precursor RNA transcription was driven by the T7 promoter, using templates linearized at the *Xba*I site (nucleotide 9727). Below the map, the expected precursor and processed (polyadenylated [PA]) products are shown. Wild-type (WT) and NXS mutant templates produce a 280-nucleotide precursor which is processed to 237 nucleotides by polyadenylation cleavage and the 337 to 437 nucleotides by cleavage and polyadenylation. The Δ*Bgl*II template produces a 147-nucleotide precursor which is cleaved to 105 nucleotides. Cleavage and polyadenylation of this precursor result in RNAs of 205 to 305 nucleotides.

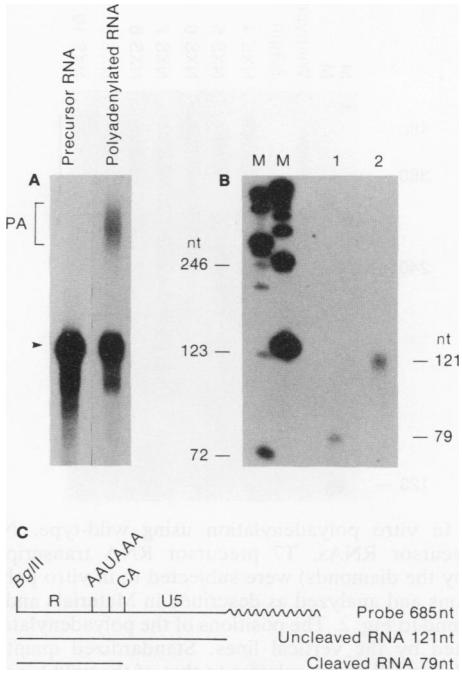

FIG. 2. In vitro polyadenylation using the wild-type HIV polyadenylation signal. (A) Three nanograms of ^{32}P-labeled wild-type precursor RNA (arrowhead) was incubated with HeLa nuclear extract (60% by volume; 25-μl final reaction volume). Reaction products were resolved on 40-cm 5% polyacrylamide–7 M urea gels. Unreacted precursor RNA was electrophoresed as a control. The polyadenylated products (PA) migrate between 330 and 440 nucleotides (see Fig. 1). (B) RNase protection mapping of in vitro polyadenylated RNA. The ^{32}P-labeled polyadenylated RNA shown in panel A was excised and eluted from the gel. Recovered RNA was hybridized to 3 μg of a partially complementary, unlabeled 685-nucleotide (nt) probe (see panel C) and digested with RNases A and T_1. Lanes: M, markers; 1, protected RNA; 2, RNase protection product which results with use of the ^{32}P-labeled wild-type uncleaved, unpolyadenylated precursor. (C) RNase protection probe and expected RNase protection products.

this RNA are AAUAAA and the downstream G+U-rich element.

In vitro polyadenylation of the 147-nucleotide Δ*Bgl*II precursor results in RNAs migrating between 205 and 305 nucleotides (Fig. 1). Polyadenylation of Δ*Bgl*II RNA was diminished relative to that of the wild type (Fig. 3). Quantitation of precursors and polyadenylated products demonstrated that during the 30-min reaction, 31% of wild-type RNA precursor was polyadenylated, whereas only 11% of Δ*Bgl*II precursor was processed (Table 1). This experiment has been repeated numerous times with similar results. Thus, U3 sequences affect polyadenylation efficiency in vitro in a manner similar to their effect in vivo.

Mutational analysis of USEs required for in vitro cleavage and polyadenylation. Previous in vivo studies from this laboratory used LS mutants to define U3 sequences required for efficient 3'-end formation. These experiments showed that LS mutations NXS 9 and 10 (substituting an 18-bp NXS linker for HIV nucleotides 9571 to 9588 and 9589 to 9606, respectively; Fig. 1) affected the efficiency of polyadenylation signal utilization (39). The two mutations decreased polyadenylation to 30 and 50% of the wild-type level, respectively. To determine whether these same sequences were required for in vitro polyadenylation, seven different templates for precursor RNA synthesis, representing mutants NXS 4 to 10, were created (Fig. 1). These templates contain sequential 18-bp; NXS linker substitutions in HIV-1 sequences between nucleotides 9463 and 9606.

The 280-nucleotide precursor RNAs were tested in in vitro polyadenylation reactions. As indicated in Fig. 3 and quantitated in Table 1, precursor RNAs derived from mutants NXS 9 and NXS 10 were polyadenylated less efficiently than was wild-type precursor RNA. These data are in exact agreement with in vivo data. Thus, sequences in the region between HIV nucleotides 9571 and 9606 are required for efficient polyadenylation signal recognition both in vitro and in vivo.

To investigate whether sequences between 9571 and 9606 were required for efficient cleavage in vitro, the NXS mutant precursor RNAs were tested in reactions in which polyadenylation was inhibited by inclusion of 1 mM cordycepin and

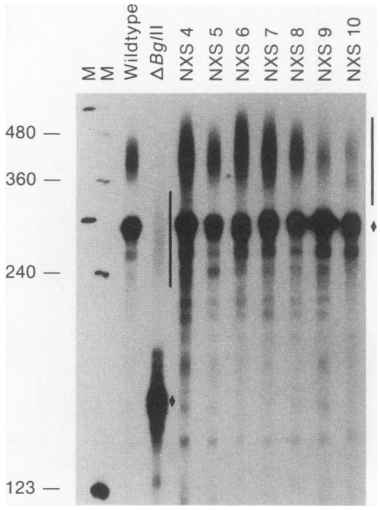

FIG. 3. In vitro polyadenylation using wild-type, NXS, and *ΔBglII* precursor RNAs. T7 precursor RNA transcripts (3 ng; indicated by the diamonds) were subjected to in vitro polyadenylation reactions and analyzed as described in Materials and Methods and the legend to Fig. 2. The positions of the polyadenylated species are indicated by the vertical lines. Standardized quantitation of polyadenylation efficiency relative to that of the wild type is shown in Table 1. Lanes M, markers. Sizes are indicated in nucleotides.

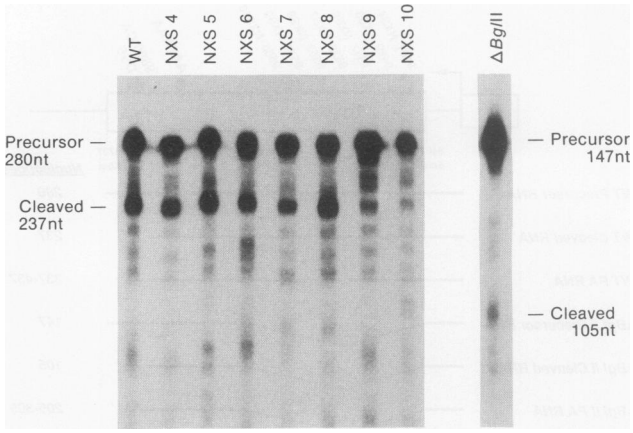

FIG. 4. In vitro cleavage using wild-type (WT), NXS, and *ΔBglII* precursor RNAs. T7 precursor RNA transcripts (3 ng) were subjected to in vitro polyadenylation reactions in the presence of cordycepin and analyzed as described in Materials and Methods and the legend to Fig. 2. Standardized quantitation of polyadenylation cleavage efficiency relative to that of the wild type is shown in Table 1. nt, nucleotides.

reduction of the ATP concentration from 1 mM to 250 μM. Figure 4 and Table 1 indicate that mutants NXS 9 and 10 were cleaved less efficiently than was the wild type. The intensity of the 237-nucleotide cleaved product was reduced by these two mutations. Bands migrating between 240 and 280 nucleotides represent RNA degradation products, since they were visible in lanes with unreacted substrate (not shown). Table 1 summarizes the quantitative results of three separate cleavage experiments performed with two different nuclear extracts.

Precursor RNAs with short distances between the 5′ cap and AAUAAA can be cleaved efficiently in vitro. Previous studies of several retroviral polyadenylation signals suggest that the distance between the cap site and the AAUAAA may influence the efficiency of polyadenylation site utilization (22, 40). For HIV-1, distances greater than 240 nucleotides from cap to AAUAAA were suggested to be necessary for efficient utilization in vivo (40). To determine whether similar distance effects were operative in vitro, a panel of deletion mutants in which this distance was varied in the presence or absence of the significant USEs was constructed (Fig. 5A).

TABLE 1. Quantitation of polyadenylation and cleavage[a]

RNA	Nucleotides affected	% Polyadenylation	% Cleavage
WT		31	25
ΔBglII	Δ9463–9612	11	8
NXS 4	LS 9481–9498	38	22
NXS 5	LS 9499–9516	42	26
NXS 6	LS 9517–9534	41	33
NXS 7	LS 9535–9552	39	25
NXS 8	LS 9553–9570	45	28
NXS 9	LS 9571–9588	11	9
NXS 10	LS 9589–9606	4	13

[a] Data for the cleavage reaction are averages of three experiments; the standard deviation was ±4. The polyadenylation data are results of a single experiment using all of the substrates; significant points (wild type [WT], *ΔBglII*, NXS 9, and NXS 10) have been repeated numerous times with similar results.

Template Δ4-8 produced a 203-nucleotide precursor RNA containing the USEs defined above and 140 nucleotides between the 5′ cap and the AAUAAA (Fig. 5A). This distance is 72 nucleotides shorter than that in wild-type HIV-1 polyadenylation substrate RNA (Fig. 5A) and approximately 160 nucleotides shorter than the minimal effective distance suggested by in vivo studies (40).

The processing efficiency of the Δ4-8 substrate RNA was tested in the in vitro cleavage assay. Quantitation of the properly cleaved 166-nucleotide product indicated this mutant precursor was processed with somewhat greater efficiency than was the wild type (Fig. 5B; 38 and 20% efficiency, respectively). Thus, while at least 250 nucleotides between the 5′ cap and the AAUAAA may be required for polyadenylation signal utilization in vivo (40), precursor RNAs with distances as short as 140 nucleotides are processed with wild-type efficiency in vitro.

Deletion of an additional 18 and 36 nucleotides (Δ4-9 and Δ4-10; Fig. 5A) further reduced the distance from cap to AAUAAA in precursor RNA and resulted in decreased in vitro cleavage efficiency (Fig. 5B). However, this was a predictable result since Δ4-9 and Δ4-10 delete the USEs defined above. Moreover, cleavage of these RNAs was reduced to levels comparable to that of the Δ9-10 precursor RNA (Fig. 5B), a precursor which has a greater distance (194 nucleotides) between the 5′ cap and AAUAAA but lacks the 36-nucleotide USE found to be important for 3′-end formation both in vivo and in vitro. Thus, the absence of the USE, not distance between 5′ cap and AAUAAA, can account for inefficient in vitro cleavage of Δ4-9 and Δ4-10 precursors.

DISCUSSION

The data presented show that the HIV-1 poly(A) signal directs accurate cleavage and polyadenylation in vitro. Significantly, sequences upstream of AAUAAA are required for efficient in vitro processing; mutation of a 36-nucleotide

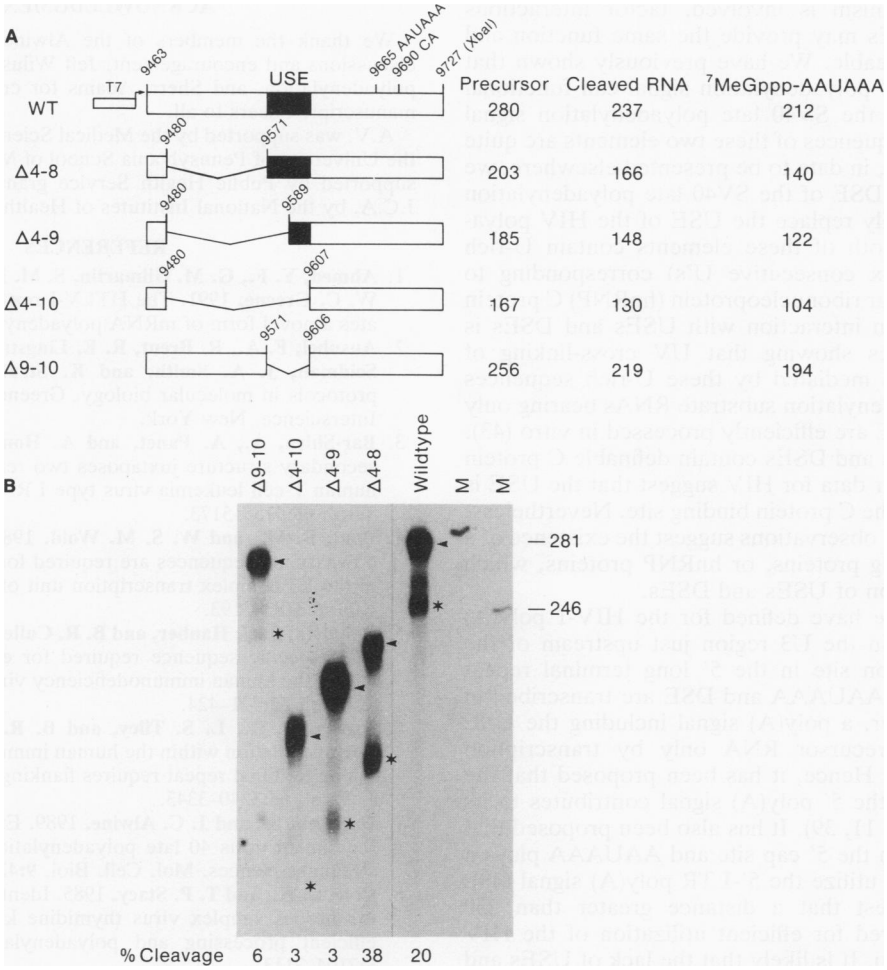

FIG. 5. In vitro cleavage of precursors with short distances between the 5′ cap and AAUAAA. (A) Constructions, precursor and cleaved RNA sizes, and distances between 5′ cap and AAUAAA. All templates encode AAUAAA (9665 to 9670), cleavage site (9690), and DSE (9691 to 9715). The USE is indicated by solid boxes. HIV-1 nucleotide numbers indicate regions deleted in mutant constructions. All precursors were synthesized from an *Xba*I-linearized template as described above. Precursor and cleaved RNA sizes and the distances between 5′ cap and AAUAAA are shown for each deletion mutant at the right. WT, wild type. (B) In vitro cleavage of precursors with short distances between the 5′ cap and AAUAAA. Three nanograms of ^{32}P-labeled precursor RNA was incubated in the in vitro polyadenylation reaction in the presence of cordycepin and analyzed as described in Materials and Methods. Arrowheads and stars indicate precursor and cleaved product RNAs, respectively. Cleavage efficiency, expressed as percent cleavage below each lane, was calculated as described in Materials and Methods. Lanes M, markers. Sizes are indicated in nucleotides.

region lying between 56 and 93 bases upstream of AAUAAA (Fig. 6) reduced the level of cleavage to at least one-third the wild-type level. These data are the first demonstration of the function of USEs in vitro. In addition, the in vitro results are in complete agreement with those of previous in vivo experiments using the same mutations. Reporter gene RNA 3′-end processing in COS cells was diminished to 30% of the wild-type level by LS mutations between 56 and 93 nucleotides upstream of the AAUAAA (39).

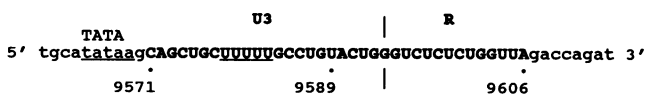

FIG. 6. Sequence of the major USE of the HIV-1 poly(A) signal as defined by in vivo and in vitro LS analyses. The USE sequences defined by LS mutants NXS 9 and NXS 10 (between HIV nucleotides 9571 and 9606) are highlighted in uppercase bold type. The UUUUU sequence (underlined) occurs at positions 9578 to 9582. The USE is indicated in relationship to the TATA box (underlined) and the transcription start site (vertical line at the U3/R junction) of the LTR. The AAUAAA is located 58 nucleotides downstream.

USEs affecting in vivo polyadenylation efficiency have been described in a number of viral systems (7, 11, 33, 39). However, the mechanisms of action of both USEs and DSEs remain unknown. It is likely that these elements act to increase polyadenylation complex binding to the adjacent AAUAAA. This could be accomplished through a number of mechanisms. For example, factor binding to USEs and DSEs might hold pre-mRNA in a conformation in which the adjacent AAUAAA remains open and unobstructed by nonspecific RNA-binding proteins. Alternatively, USEs and DSEs, or secondary structures involving these elements, may allow binding of factors which interact directly with the polyadenylation complex to activate processing.

Whichever mechanism is involved, factor interactions with USEs and DSEs may provide the same function and hence be interchangeable. We have previously shown that the USE of the HIV polyadenylation signal can functional replace the USE of the SV40 late polyadenylation signal (39). The primary sequences of these two elements are quite different. In addition, in data to be presented elsewhere, we have found that the DSE of the SV40 late polyadenylation signal can functionally replace the USE of the HIV polyadenylation signal. Both of these elements contain U-rich elements (five to six consecutive U's) corresponding to heterogeneous nuclear ribonucleoprotein (hnRNP) C protein binding sites. Protein interaction with USEs and DSEs is supported by studies showing that UV cross-linking of hnRNP C protein is mediated by these U-rich sequences (43). Further, polyadenylation substrate RNAs bearing only UUUUU as the DSE are efficiently processed in vitro (43). Clearly not all USEs and DSEs contain definable C protein binding sites, and our data for HIV suggest that the USE is more complex than the C protein binding site. Nevertheless, taken together, these observations suggest the existence of a class of RNA-binding proteins, or hnRNP proteins, which facilitates the function of USEs and DSEs.

The USEs that we have defined for the HIV-1 poly(A) signal lie primarily in the U3 region just upstream of the transcription initiation site in the 5′ long terminal repeat (LTR) (Fig. 6). The AAUAAA and DSE are transcribed in the 5′ LTR; however, a poly(A) signal including the USE would appear in precursor RNA only by transcription through the 3′ LTR. Hence, it has been proposed that the lack of the USE in the 5′ poly(A) signal contributes to its lack of utilization (6, 11, 39). It has also been proposed that the distance between the 5′ cap site and AAUAAA plays a role in the failure to utilize the 5′-LTR poly(A) signal (40). Those studies suggest that a distance greater than 250 nucleotides is required for efficient utilization of the HIV poly(A) signal in vivo. It is likely that the lack of USEs and the proximity of the signal to the promoter each contribute to the inefficiency of utilization of the 5′-LTR polyadenylation signal, providing greater protection against aberrant utilization of this polyadenylation signal. However, we have found that as long as the upstream region contains the USE, a precursor RNA with 140 nucleotides between the 5′ cap and the AAUAAA can be efficiently processed. Thus, the polyadenylation complex has no inherent inability to cleave precursor RNAs with short cap-to-AAUAAA distances, as long as all of the known polyadenylation signal elements are present. Thus, the effects on polyadenylation relating to 5′-end proximity in vivo do not appear to depend on absolute distance but may be mediated by a mechanism which is not reproduced in the in vitro system. For example, proteins interacting with the nearby promoter, tethering of the 5′ end of the nascent transcript to transcription factors or cap-binding proteins, may hinder the polyadenylation reaction in vivo but may not be detected in an in vitro system in which active transcription is not occurring.

Lastly, regulation of poly(A) signal utilization is a means by which gene expression may be controlled. For example, poly(A) site utilization is regulated during viral infection (10, 12, 13, 26) and has been shown to be modulated tissue specifically (15, 16, 31). Given that the AAUAAA is found in all poly(A) signals, specific control at this level seems unlikely. Since USEs and DSEs affect efficiency of signal utilization and have less defined consensus elements, they appear to be good candidates for the sites of regulatory effects.

ACKNOWLEDGMENTS

We thank the members of the Alwine laboratory for helpful discussions and encouragement, Jeff Wilusz for advice on in vitro polyadenylation, and Sherri Adams for critical comments on the manuscript. Cheers to all.

A.V. was supported by the Medical Scientist Training Program of the University of Pennsylvania School of Medicine. This work was supported by Public Health Service grant GM45773 awarded to J.C.A. by the National Institutes of Health.

REFERENCES

1. **Ahmed, Y. F., G. M. Gilmartin, S. M. Hanly, J. R. Nevins, and W. C. Greene.** 1991. The HTLV-I rex response element mediates a novel form of mRNA polyadenylation. Cell **64**:727–737.
2. **Ausubel, F. A., R. Brent, R. E. Lingstrom, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. Greene Publishing and Wiley-Interscience, New York.
3. **Bar-Shira, A., A. Panet, and A. Honigman.** 1991. An RNA secondary structure juxtaposes two remote genetic signals for human T-cell leukemia virus type I RNA 3′-end processing. J. Virol. **65**:5165–5173.
4. **Bhat, B. M., and W. S. M. Wold.** 1985. ATTAAA as well as downstream sequences are required for RNA 3′-end formation in the E3 complex transcription unit of adenovirus. Mol. Cell. Biol. **5**:3183–3193.
5. **Bohnlein, S., J. Hauber, and B. R. Cullen.** 1989. Identification of a U5-specific sequence required for efficient polyadenylation within the human immunodeficiency virus long terminal repeat. J. Virol. **63**:421–424.
6. **Brown, P. H., L. S. Tiley, and B. R. Cullen.** 1991. Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. J. Virol. **65**:3340–3343.
7. **Carswell, S., and J. C. Alwine.** 1989. Efficiency of utilization of the simian virus 40 late polyadenylation signal: effects of upstream sequences. Mol. Cell. Biol. **9**:4248–4258.
8. **Cole, C. N., and T. P. Stacy.** 1985. Identification of sequences in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation. Mol. Cell. Biol. **5**:2104–2113.
9. **Conway, L., and M. Wickens.** 1985. A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3′ termini in frog oocytes. Proc. Natl. Acad. Sci. USA **82**:3949–3953.
10. **DeZazzo, J. D., and M. J. Imperiale.** 1989. Sequences upstream of AAUAAA influence poly(A) site selection in a complex transcription unit. Mol. Cell. Biol. **9**:4951–4961.
11. **DeZazzo, J. D., J. E. Kilpatrick, and M. J. Imperiale.** 1991. Involvement of long terminal repeat U3 sequences overlapping the transcription control region in human immunodeficiency virus type 1 mRNA 3′ end formation. Mol. Cell. Biol. **11**:1624–1630.
12. **Falck-Pederson, E., and J. Logan.** 1989. Regulation of poly(A) site selection in adenovirus. J. Virol. **63**:532–541.
13. **Falck-Pederson, E., J. Logan, T. Shenk, and J. E. Darnell.** 1985. Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse β-major globin terminator element. Cell **40**:897–905.
14. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5′-AAUAAA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell **24**:251–260.
15. **Galli, G., J. W. Guise, M. A. McDevitt, P. W. Tucker, and J. R. Nevins.** 1987. Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted μ-chain expression during B-cell development. Genes Dev. **1**:471–481.
16. **Galli, G., J. Guise, P. W. Tucker, and J. R. Nevins.** 1988. Poly(A) site choice rather than splice site choice governs the regulated production of IgM heavy-chain RNAs. Proc. Natl. Acad. Sci. USA **85**:2439–2443.
17. **Gil, A., and N. J. Proudfoot.** 1984. A sequence downstream of AAUAAA is required for rabbit β-globin mRNA 3′-end forma-

tion. Nature (London) **312**:473–474.
18. **Gil, A., and N. J. Proudfoot.** 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β-globin mRNA formation. Cell **49**:399–406.
19. **Hart, R. P., M. A. McDevitt, H. Ali, and J. R. Nevins.** 1985. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the early simian virus 40 polyadenylation sites. Mol. Cell. Biol. **5**:2975–2983.
20. **Hart, R. P., M. A. McDevitt, and J. R. Nevins.** 1985. Poly(A) site cleavage in a HeLa nuclear extract is dependent on downstream sequences. Cell **43**:677–683.
21. **Heath, C. V., R. M. Denome, and C. N. Cole.** 1990. Spatial constraints on polyadenylation signal function. J. Biol. Chem. **265**:9098–9104.
22. **Iwasaki, K., and H. M. Temin.** 1991. The efficiency of RNA 3′-end formation is determined by the distance between the cap site and the poly(A) site in spleen necrosis virus. Genes Dev. **4**:2299–2307.
23. **McDevitt, M. A., M. J. Imperiale, H. Ali, and J. R. Nevins.** 1984. Requirement of a downstream sequence for generation of a poly(A) addition site. Cell **37**:993–999.
24. **McDevitt, M. A., R. P. Hart, W. W. Wong, and J. R. Nevins.** 1986. Sequences capable of restoring poly(A) site function define two distinct downstream elements. EMBO J. **5**:2907–2913.
25. **McLauchlan, J., D. Gaffney, J. L. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**:1347–1368.
26. **McLauchlan, J., S. Simpson, and J. B. Clements.** 1989. Herpes simplex virus induces a processing factor that stimulates poly(A) site usage. Cell **59**:1093–1105.
27. **Montell, C., E. F. Fisher, M. H. Caruthers, and A. J. Berk.** 1983. Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3′ consensus sequence AAUAAA. Nature (London) **305**:600–606.
28. **Moore, C. L.** 1990. Preparation of mammalian extracts active in polyadenylation. Methods Enzymol. **181**:49–74.
29. **Moore, C. L., and P. A. Sharp.** 1984. Site-specific polyadenylation in a cell-free reaction. Cell **36**:581–591.
30. **Moore, C. L., and P. A. Sharp.** 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell **41**:845–855.
31. **Peterson, M. L., and R. P. Perry.** 1989. The regulated production of μ_m and μ_s mRNA is dependent on the relative efficiencies of μ_s poly(A) site usage and the Cμ4-to-M1 splice. Mol. Cell. Biol. **9**:726–738.
32. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′ noncoding region sequences in eukaryotic messenger RNA. Nature (London) **236**:211–214.
33. **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.
34. **Ryner, L. C., Y. Takagaki, and J. L. Manley.** 1989. Sequences downstream of AAUAAA signals affect pre-mRNA cleavage and polyadenylation in vitro both directly and indirectly. Mol. Cell. Biol. **9**:1759–1771.
35. **Sadofsky, M., and J. C. Alwine.** 1984. Sequences on the 3′ side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. Mol. Cell. Biol. **4**:1460–1468.
36. **Sadofsky, M., S. Connelly, J. L. Manley, and J. C. Alwine.** 1985. Identification of a sequence element on the 3′ side of AAUAAA which is necessary for simian virus 40 late mRNA 3′-end processing. Mol. Cell. Biol. **5**:2713–2719.
37. **Sanfacon, H., P. Brodmann, and T. Hohn.** 1991. A dissection of the cauliflower mosaic virus polyadenylation signal. Genes DEv. **5**:141–149.
38. **Sperry, A. O., and S. M. Berget.** 1986. In vitro cleavage of the simian virus 40 early polyadenylation site adjacent to a required downstream TG sequence. Mol. Cell. Biol. **6**:4734–4741.
39. **Valsamakis, A., S. Zeichner, S. Carswell, and J. C. Alwine.** 1991. The human immunodeficiency virus type 1 polyadenylation signal: a 3′ long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylation. Proc. Natl. Acad. Sci. USA **88**:2108–2112.
40. **Weichs an der Glon, C., J. Monks, and N. J. Proudfoot.** 1991. Occlusion of the HIV poly(A) site. Genes Dev. **5**:244–253.
41. **Wickens, M.** 1990. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. **15**:271–281.
42. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3′ end formation. Science **226**:1045–1051.
43. **Wilusz, J., and T. Shenk.** 1990. A uridylate tract mediates efficient heterogeneous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. Mol. Cell. Biol. **10**:6397–6407.
44. **Zarkower, D., P. Stephenson, M. Sheets, and M. Wickens.** 1986. The AAUAAA sequence is required both for cleavage and for polyadenylation of simian virus 40 pre-mRNA in vitro. Mol. Cell. Biol. **6**:2317–2323.
45. **Zeichner, S. L., J. Y. Kim, and J. C. Alwine.** 1991. Linker-scanning mutational analysis of the transcriptional activity of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. **65**:2436–2444.
46. **Zhang, F., and C. N. Cole.** 1987. Identification of a complex associated with processing and polyadenylation in vitro of herpes simplex virus type 1 thymidine kinase precursor RNA. Mol. Cell. Biol. **7**:3277–3286.
47. **Zhang, F., R. M. Denome, and C. N. Cole.** 1986. Fine-structure analysis of the processing and polyadenylation region of the herpes simplex virus type 1 thymidine kinase gene by using linker scanning, internal deletion and insertion mutations. Mol. Cell. Biol. **6**:4611–4623.