The EMBO Journal vol.7 no.8 pp.2515-2522, 1988

A rabbit β-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA

Jacqueline Lanoix and Nicholas H.Acheson

Department of Microbiology and Immunology, McGill University, Montreal, Quebec, H3A 2B4, Canada

Communicated by J.R.Tata

We constructed a viable insertion mutant (ins 5) of polyomavirus which contains, upstream of the L-strand polyadenylation signal, a 94-nt fragment of rabbit β-globin DNA. Included in this fragment are all of the sequence elements required for efficient cleavage and polyadenylation of rabbit β-globin RNA. The β-globin signal was efficiently recognized by the cleavage/polyadenylation machinery in mouse 3T6 cells infected with ins 5, signalling >90% of the polyadenylation events on L-strand RNAs. Furthermore, the presence of this efficient polyadenylation signal resulted in a 1.4- to 2.5-fold increase in the fraction of virus-specific RNAs that were polyadenylated. Most importantly, termination of transcription by RNA polymerase II on ins 5 DNA was also increased compared with wild-type virus; nearly 100% of polymerases terminated per traverse of the ins 5 genome. These findings demonstrate that the rabbit β-globin insert, which contains a strong polyadenylation signal, also contains at least part of a signal for termination of transcription by RNA polymerase II. These results also show that the multiple, spliced leaders on polyomavirus L-strand mRNAs, which arise as a result of inefficient termination and polyadenylation, are not necessary for efficient virus replication.

Key words: insertion mutants/nascent RNA/ribonuclease mapping/S1 nuclease mapping/transcription termination

Introduction

The 3′ termini of most mRNAs in higher eukaryotes are formed by endonucleolytic cleavage of precursor RNAs, followed by polyadenylation of the newly-created 3′ end. This process is signalled by the sequence AATAAA (or slight variants) in conjunction with other, less conserved downstream sequence elements (Gil and Proudfoot, 1984, 1987; McDevitt *et al.*, 1984; McLauchlan *et al.*, 1985; Lanoix *et al.*, 1986; Mason *et al.*, 1986, and others). RNA polymerase II continues transcription beyond polyadenylation signals, terminating hundreds or thousands of nucleotides downstream, at sites that are still poorly defined (Citron *et al.*, 1984; Hagenbuchle *et al.*, 1984; LeMeur *et al.*, 1984). The nature of the sequences in DNA (or RNA) that signal termination of transcription by RNA polymerase II remains obscure.

Termination of transcription by the more specialized RNA polymerases I (ribosomal RNA genes) and III (5S and tRNA genes) is signalled by relatively simple sequence elements which lie near termination sites in the DNA template (Bogenhagen and Brown, 1981; Grummt *et al.*, 1986). Unlike these enzymes, RNA polymerase II must transcribe many thousands of different genes varying in length from 1 to >1000 kb (Koenig *et al.*, 1987). Clearly, RNA polymerase II must be designed to transcribe these genes without interruption until the functional end of the gene has been reached. Since this functional end is the site where cleavage and polyadenylation take place, it has occurred to us (and others: Acheson, 1984; Citron *et al.*, 1984; Lanoix *et al.*, 1986; Tseng and Acheson, 1986; Whitelaw and Proudfoot, 1986) that polyadenylation signals may themselves transmit to RNA polymerase II the information required to terminate transcription.

Falck-Pedersen *et al.* (1985) showed that insertion of a 1.6-kb fragment of the mouse β-globin gene, including the polyadenylation signal, within the adenovirus E1A gene caused efficient termination during the early phase of adenovirus infection. However, a smaller fragment that did not contain the polyadenylation signal failed to terminate transcription. Furthermore, Whitelaw and Proudfoot (1986) showed that transcription of the human α-globin gene failed to terminate correctly in a thalassaemic mutant which lacks a functional polyadenylation signal.

We have previously noted a correlation between the efficiency with which polyadenylation signals are recognized and the efficiency of transcription termination in the papovaviruses, polyomavirus and simian virus 40 (Ford and Hsu, 1978; Acheson, 1981, 1984; Miller *et al.*, 1982; Lanoix *et al.*, 1986). This led us to construct a polyomavirus mutant (ins 4) in which the late polyadenylation signal was duplicated (Lanoix *et al.*, 1986). Although 3′ end formation was correctly signalled at both polyadenylation sites, neither the overall efficiency of polyadenylation of late viral RNA nor the efficiency of termination of transcription was increased in this mutant. Two alternative hypotheses could explain these results: either polyadenylation and termination are both inefficient in polyomavirus-infected cells because of the scarcity of essential factors involved in these processes; or the polyomavirus late polyadenylation signal is inherently weak and even when duplicated cannot be efficiently recognized by the factors responsible for cleavage and polyadenylation and possibly also termination of transcription.

To test these hypotheses, we have made use of a tandem duplication of the rabbit β-globin polyadenylation signal, constructed by Gil and Proudfoot (1984). When plasmids containing this duplicated signal were transfected into HeLa cells, all mRNAs produced were found to have 3′ ends directed by the upstream signal, even though the downstream signal was shown to be fully functional when the upstream signal was deleted (Gil and Proudfoot, 1984). The rabbit β-globin polyadenylation signal is therefore very efficiently recognized by the cleavage/polyadenylation machinery in

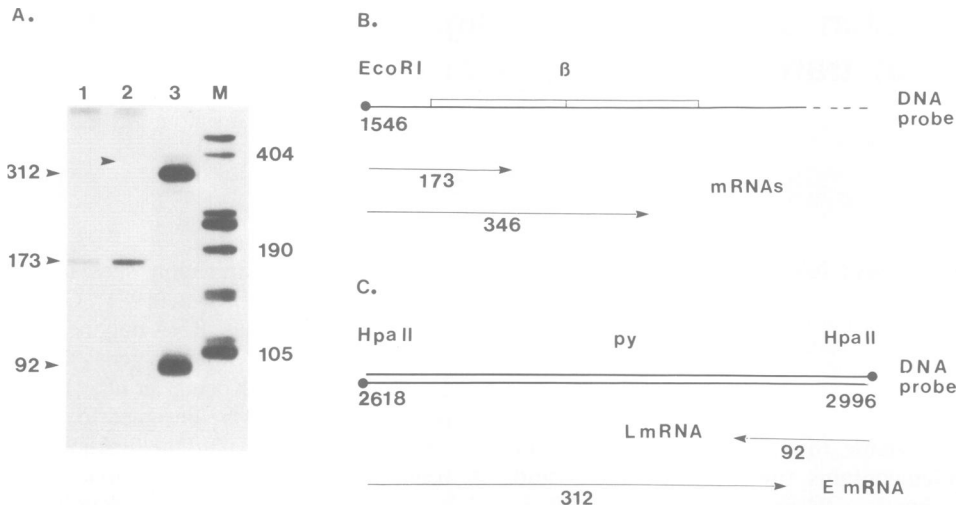

Fig. 1. S1 nuclease mapping of mRNAs from 3T6 cells transfected with a plasmid containing a duplicated rabbit β-globin polyadenylation signal. **(A)** Autoradiogram of S1 mapping gel. Each hybridization reaction received 5 ng of 3′ end-labeled DNA probe. The probe diagrammed in B was hybridized to cytoplasmic poly(A)$^+$ RNA prepared from six 100-mm Petri dishes of uninfected cells transfected with the rabbit β-globin plasmid (see Materials and methods, and lane 1), or from 12 Petri dishes of polyomavirus-infected cells transfected with the plasmid (lane 2). Hybridization was carried out at 53°C overnight. In lane 3, the probe diagrammed in C was hybridized to RNA from six Petri dishes of infected, transfected cells (same sample as in lane 2). This hybridization was at 49°C overnight. Hybrids were treated with 100 U/ml S1 nuclease, and S1-resistant DNA fragments were analyzed on a 7% polyacrylamide−8 M urea gel. The mol. wt markers (lane M) are an *Msp*I digest of ^{32}P-labeled plasmid pSP65. The arrowhead indicates the expected position of the DNA fragment that would be protected by mRNAs with 3′ ends at the downstream β-globin polyadenylation site. **(B)** Structure of the 3′ end-labeled β-globin DNA fragment used as probe in lanes 1 and 2. The region duplicated in the β-globin plasmid is shown as two elongated rectangles. The labeled end of the probe is marked by a filled circle. The length of the expected S1-resistant DNA fragments is shown below the lines. Arrowheads designate positions of the polyadenylated 3′ ends of mRNAs. **(C)** Structure of the 3′ end-labeled polyomavirus *Hpa*II DNA fragment used as probe in lane 3. Arrowheads designate expected positions of polyadenylated 3′ ends of early and late mRNAs.

HeLa cells. We show here that this signal is also efficiently recognized in polyomavirus-infected mouse 3T6 cells, which indicates that infected cells do not lack factors necessary for polyadenylation. Furthermore, insertion of a single copy of the rabbit β-globin sequence, 94 nt long, into the late transcription unit of polyomavirus increased the overall efficiency of polyadenylation of late viral RNA and simultaneously increased the efficiency of termination of transcription on the L DNA strand. These findings therefore demonstrate that a functional polyadenylation signal is also part of a signal for termination of transcription.

Results

The rabbit β-globin polyadenylation signal is efficiently recognized in polyomavirus-infected mouse cells

We first wanted to verify whether the cleavage/polyadenylation machinery was intact in polyomavirus-infected mouse 3T6 cells. Our approach was to test the efficiency with which the rabbit β-globin polyadenylation signal is recognized in these cells. The plasmid we used (Gil and Proudfoot, 1984, see Materials and methods) contains a tandem duplication of this signal. When transfected into HeLa cells, this rabbit β-globin plasmid generates mRNAs polyadenylated only at the upstream site (Gil and Proudfoot, 1984). We reasoned that if the polyadenylation machinery were intact in 3T6 cells during the late phase of polyomavirus infection, transfection of the β-globin plasmid into these cells should also give rise to mRNAs polyadenylated only at the upstream site. If, however, the polyadenylation machinery were deficient, the β-globin signal would fail to be efficiently recognized, and therefore some RNAs would be polyadenylated at the upstream site, and some at the downstream site.

To ensure that most β-globin-specific mRNAs were synthesized and processed during the late phase of infection (after the onset of viral DNA replication), we first determined the time-course of expression of the plasmid after transfection of 3T6 cells. This experiment (not shown) revealed no significant accumulation of β-globin mRNA during the first 12 h after transfection. Since viral DNA replication begins at 12−15 h after infection (Tooze, 1980), we infected 3T6 cells with polyomavirus and transfected the same cells 6 h later with the β-globin plasmid. β-globin mRNA will therefore only begin to accumulate at 18 h after infection, well into the late phase. We extracted RNA from these cells 30 h after infection (24 h after transfection), and analyzed the 3′ ends of cytoplasmic poly(A)$^+$ RNA by S1 nuclease mapping (Figure 1). The probe was a 3′ end-labeled *Eco*RI fragment that spans the entire duplicated region of the rabbit β-globin plasmid (Figure 1B). RNAs were hybridized to this probe, and S1 nuclease-resistant DNA was analyzed on a denaturing polyacrylamide−urea gel (Figure 1A). As shown in lane 1, RNA extracted from uninfected 3T6 cells that were transfected with the plasmid gave rise to an S1-resistant fragment 173 nt long, as expected for RNA molecules polyadenylated within the upstream repeat. RNA extracted from cells that were first infected with polyomavirus and subsequently transfected with the plasmid (lane 2) gave the same result. No RNA was polyadenylated at the downstream site; the expected position of the DNA fragment that would be protected by such an RNA species (346 nt long) is indicated by an arrowhead adjacent to lane 2. RNA from the same cells was also hybridized to a labeled fragment of

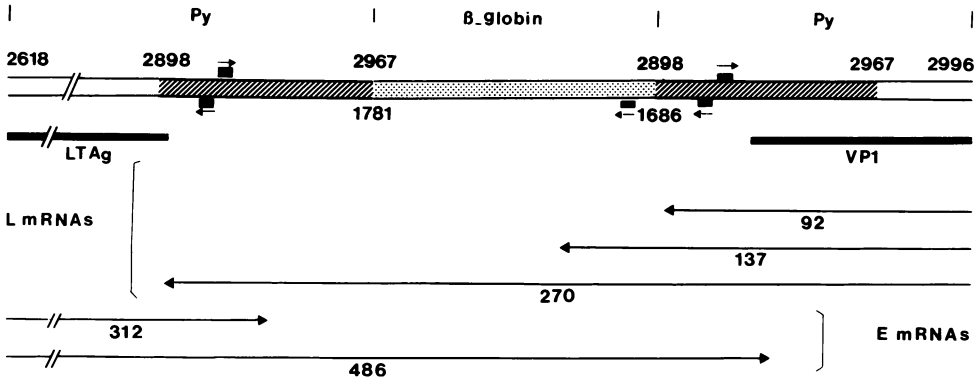

Fig. 2. Structure of the region surrounding the polyadenylation sites in ins 5. The fragment shown (top line) is the *Hpa*II fragment 6 of polyomavirus DNA, into which the *Sfa*nI (nt 1686)−*Aha*III (nt 1780) fragment of rabbit β-globin DNA (dotted box) was inserted (junctions 2967/1781 and 1686/2898). Numbers above the map are polyomavirus nucleotide positions. The 69-nt duplication (nt 2898−2967) is represented by hatched boxes. Numbers below the map are rabbit β-globin nucleotide positions. Black rectangles represent AAUAAA motifs and small arrows indicate the orientation (5′ to 3′) of each motif in RNA transcripts. The C-terminal portions of polyomavirus large T antigen (LTAg) and viral capsid protein VP1 are indicated by black elongated rectangles below the DNA fragment. Expected lengths of S1-resistant DNA fragments are shown below the thin lines, and arrowheads indicate the expected positions of 3′ ends of polyadenylated mRNAs.

polyomavirus DNA (Figure 1C) to monitor the level of expression of the polyomavirus genome in this experiment. Lane 3 shows that polyomavirus-specific early RNA protected a 312-nt fragment, and late RNA protected a 92-nt fragment, as expected.

This experiment indicates that although substantial amounts of polyomavirus early and late RNAs were synthesized in these cells, the rabbit β-globin polyadenylation signal was nevertheless efficiently recognized by the cleavage/polyadenylation machinery during the late phase of infection. We therefore conclude that the cleavage/polyadenylation machinery is intact during the late phase in 3T6 cells infected with wild-type polyomavirus.

An inserted rabbit β-globin signal efficiently directs polyadenylation of polyomavirus RNAs

We next wanted to test whether the rabbit β-globin polyadenylation signal would also be recognized efficiently when inserted into the polyomavirus genome. We therefore isolated a 94-nt *Sfa*nI−*Aha*III fragment of β-globin DNA which extends from nt 1688 to 1781 (see Materials and methods and legend to Figure 2), and contains all four elements shown by Gil and Proudfoot (1987) to be required for efficient polyadenylation. We inserted this fragment into the DNA of a polyomavirus mutant (ins 3) which we had previously created (Lanoix *et al.*, 1986). Ins 3 contains a tandem duplication of 69 nt which includes the C-terminal portion of the coding region for the major viral capsid protein VP1, as well as two AATAAA motifs which are part of the E- and L-strand polyadenylation signals (Figure 2). Foreign DNA sequences can be inserted at a unique *Sal*I site located at the junction between the tandem repeats, without interrupting either E- or L-strand protein coding regions and without disrupting the polyadenylation signals for E- or L-strand mRNAs. We isolated viable recombinant virus after transfection of mouse cells with viral DNA released from the parent plasmids. The virus that contains the β-globin cleavage/polyadenylation signal inserted in the direction of L-strand transcription was called ins 5, and the virus with the insert in the reverse orientation was called ins 6. Both mutants multiplied as well as wild-type virus in primary baby

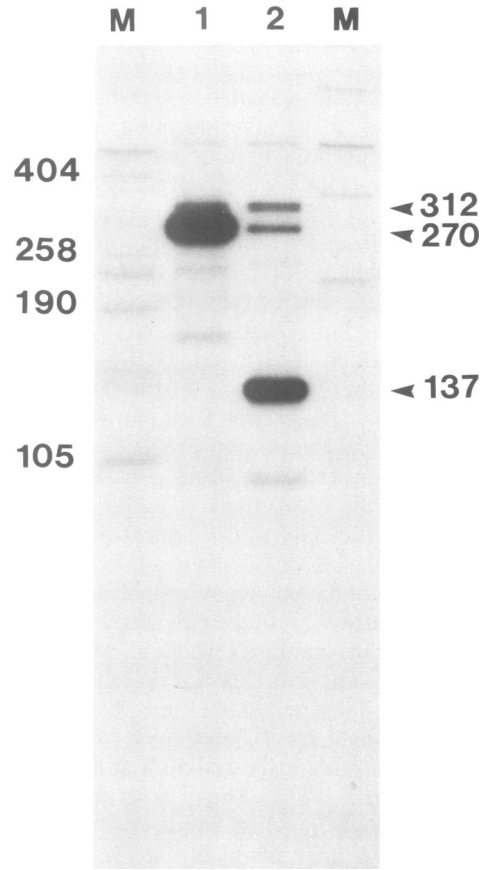

Fig. 3. S1 nuclease analysis of 3′ ends of L mRNAs from 3T6 cells infected with ins 5 or ins 6 virus. Each hybridization reaction received 45 ng of 3′ end-labeled probe (*Hpa*II fragment 6, see Figure 2, uniquely labeled at nt 2992, after cleavage of the other end with *Hph*I at position 2638) from either ins 6 (lane 1) or ins 5 (lane 2) DNA, and 10 μg of the appropriate cytoplasmic poly(A)$^+$ RNA. Hybridization was carried out at 49°C overnight. Hybrids were treated with 200 U/ml of S1 nuclease for 1 h at 30°C. S1-resistant DNA fragments were analyzed on a 7% polyacrylamide−8 M urea gel. The mol. wt markers (lanes M) are, from left to right: *Msp*I and *Hin*fI digests of pSP65 DNA.

mouse kidney cells, despite the presence of an insert which contains a total of 174 extra nucleotides, resulting in viral DNA 5484 nt long.

We infected 3T6 cells with ins 5 or ins 6, extracted RNA 28 h p.i., and analyzed cytoplasmic poly(A)$^+$ RNA by S1 nuclease mapping, as described in the legend to Figure 3. The probe used for this analysis was a 3′ end-labeled *Hpa*II fragment of ins 5 or ins 6 DNA that spans the entire duplication (Figure 2). As shown in Figure 3 (lane 2), the major protected DNA fragment was 137 nt long, and corresponded to RNAs whose 3′ ends lie within the β-globin insert at the position expected (Gil and Proudfoot, 1984), 19 nt downstream of AATAAA. A minor band 270 nt long represented RNAs with 3′ ends at the downstream viral polyadenylation site. This band was determined by densitometry to contain only 7% as much labeled probe as the major band. The 312-nt protected fragment in lane 2 probably resulted from hybridization between E-strand RNA polyadenylated at the normal site within viral sequences, and residual uncleaved probe labeled at the *Hpa*II site at nucleotide 2618 (see Figure 2). Interestingly, the shorter 92-nt protected fragment in lane 2 probably corresponds to L-strand RNAs polyadenylated at the upstream duplicated copy of the polyomavirus polyadenylation site. This insert did not function as a polyadenylation signal in ins 3 (Lanoix *et al.*, 1986) and may have acquired some essential downstream elements brought in by addition of the β-globin sequences. However, the level of polyadenylation at this site accounted for $<1\%$ of total polyadenylation events.

RNA extracted from ins 6-infected cells protected a 270-nt DNA fragment which corresponds to L-strand mRNAs polyadenylated at the downstream viral polyadenylation site (Figure 3, lane 1). There was no L-strand RNA polyadenylated within the (inverted) β-globin insert, as expected. As in the case of ins 5, the 312-nt protected fragment probably represents E-strand RNA that was hybridized to residual uncleaved probe.

These results indicate that the strong polyadenylation signal contained within the 94-nt insert of rabbit β-globin DNA was recognized efficiently in the context of polyomavirus sequences during the late phase, in virus-infected cells. When inserted upstream of the polyomavirus polyadenylation site, this signal directed ~92% of the polyadenylation events on L-strand RNA. In contrast, when a 144-nt fragment containing the polyomavirus cleavage/polyadenylation signal was duplicated and inserted at the same position within the viral genome (mutant ins 4), only 50% of the RNA was processed at the upstream site and 50% at the downstream site (Lanoix *et al.*, 1986). These experiments therefore suggest that the polyomavirus signal was not recognized efficiently because it is inherently weak (see Cole and Santangelo, 1983).

The overall level of virus-specific polyadenylated RNA is increased in ins 5

Only a small proportion of viral RNA is polyadenylated during the late phase of polyomavirus infection (Acheson, 1984). We have previously shown that duplication of the polyomavirus L-strand polyadenylation signal in a viable mutant (ins 4) did not significantly increase the overall level of L-strand RNA that is polyadenylated (Lanoix *et al.*, 1986). Having determined that the rabbit β-globin polyadenylation signal was efficiently recognized when inserted into the L strand of the polyomavirus genome, we next asked whether the presence of this signal would increase the proportion of virus-specific RNA that is polyadenylated.

Table I. Proportion of viral L-strand nuclear RNA that is polyadenylated

Experiment	Virus	Virus specific RNA (10^3 c.p.m./Petri dish)[a]		% of RNA polyadenylated[b]
		A^+	A^-	
1	wild-type	2.5	10.5	19
	ins 5	4.5	5.5	45
2	wild-type	2.2	8.1	22
	ins 5	4.6	10.1	31

[a]Poly(A)$^+$ or poly(A)$^-$ RNAs from nuclei of infected cells labeled for 10 min with [^{3}H]uridine were hybridized to excess *Hpa*II fragment IL in solution. Hybrids were collected on nitrocellulose filters, treated with ribonuclease, washed and counted as described (Lanoix *et al.*, 1986). Each number represents the average of data from three or four replicate samples.
[b]$A^+/(A^+ + A^-) \times 100$.

We infected 3T6 cells with either wild-type or ins 5 virus and labeled RNA with [^{3}H]uridine for 10 min at 28 h after infection. We subsequently hybridized poly(A)$^+$ and poly(A)$^-$ fractions of nuclear RNA in solution to an excess of single-stranded M13 DNA containing the L strand of polyomavirus *Hpa*II fragment 1, which encompasses most of the late mRNA coding region (Tooze, 1980). Hybrids were treated with ribonuclease to trim off single-stranded RNA tails, and collected on nitrocellulose filters. This method allowed us to quantitate the proportion of L-strand transcripts that were polyadenylated during the 10 min labeling period (Lanoix *et al.*, 1986). As shown in Table I, the proportion of virus-specific nuclear RNA that was polyadenylated increased from 19% in cells infected with wild-type virus to 45% (a 2.5-fold increase) in cells infected with ins 5 in one experiment, and from 22 to 31% (a 1.4-fold increase) in another experiment. In numerous previous experiments with wild-type or ins 4 virus, the proportion of RNA polyadenylated varied between 13 and 26% (Acheson, 1984; Lanoix *et al.*, 1986; our unpublished results), but was constant for multiple RNA samples isolated from the same batch of cells infected simultaneously. Therefore we conclude that the presence of the strong β-globin polyadenylation signal in the L DNA strand led to a significant increase in the overall polyadenylation efficiency of L-strand RNA. These experiments demonstrate that polyomavirus L-strand RNAs were inefficiently polyadenylated during infection with wild-type virus because the L-strand polyadenylation signal is 'weak' and therefore poorly recognized by factors involved in the cleavage/polyadenylation reaction.

The efficiency of termination of transcription is increased in ins 5

We have shown previously (Lanoix *et al.*, 1986) that duplication of the polyomavirus L-strand polyadenylation signal had no effect on termination of L-strand transcription. We wanted to test whether the presence of the strong β-globin polyadenylation signal in the L strand would result in increased efficiency of termination of L-strand transcription. The strategy we used has been described fully

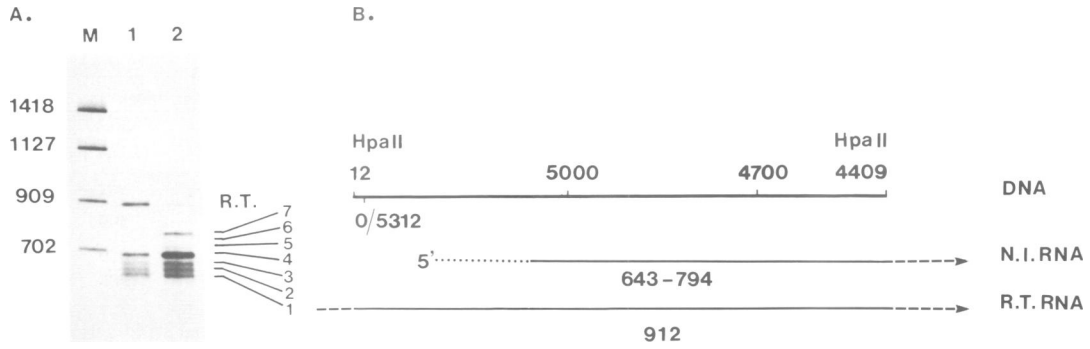

Fig. 4. Analysis of *in vivo* pulse-labeled nuclear RNA from cells infected with polyomavirus wild type or ins 5. (**A**) Nuclear RNA was labeled *in vivo* with [^{3}H]uridine for 2 min, at 28 h after infection with wild-type (lane 1) or ins 5 (lane 2) virus. [^{3}H]RNA (7.4×10^5 and 4.4×10^5 c.p.m., respectively) was hybridized to 10 μg of single-stranded bacteriophage M13mp8 DNA containing cloned polyomavirus DNA fragment *Hpa*II-3L. Hybrids were treated with ribonuclease, bound to nitrocellulose filters, eluted by incubation with S1 nuclease and subjected to electrophoresis on a mixed 0.5% agarose−2.5% polyacrylamide gel which was dried and subjected to fluorography. Band labeled RT (read-through) represents hybrids between the DNA fragment (912 nt) and RNAs made by polymerases that read through the initiation region after having traversed the entire genome at least once; bands labeled 1 to 7 correspond to hybrids with RNAs that are newly-initiated (NI) during the labeling period at sites 1 to 7, respectively (Tseng and Acheson, 1986). The mol. wt markers (lane M) are an *Hpa*II digest of polyomavirus DNA. (**B**) Diagram showing polyomavirus DNA fragment *Hpa*II-3L. Numbers above the line indicate the positions of the ends of the fragment. Below are the expected lengths of hybrids with newly-initiated RNA (NI) whose 5′ ends are at initiation sites 1 to 7 (shown by a dotted line) within the fragment, and hybrids with readthrough RNA (RT) that is protected along the full length of the DNA fragment.

elsewhere (Tseng and Acheson, 1986). Pulse-labeled nuclear RNA was hybridized to a DNA fragment that encompasses the region where L-strand transcription initiates. RNA polymerases that have just begun transcription will transcribe only the region downstream of the initiation sites; in contrast, polymerases that have made at least one full (circular) genome traverse before passing through the initiation region again will transcribe through the entire length represented by the DNA fragment. It is therefore possible to distinguish RNAs produced by newly initiated polymerases (giving rise to less-than-full-length DNA−RNA hybrids) from RNAs produced by polymerases which have read through the entire genome without terminating (giving rise to full-length DNA−RNA hybrids). Quantitation of the radioactivity in these different classes of RNA allows us to calculate the efficiency of termination on the L DNA strand.

To compare the efficiency of termination of transcription on ins 5 DNA to that on wild-type DNA, we infected the same batch of 3T6 cells with either ins 5 or wild-type virus and labeled these cells at 28 h p.i. with [^{3}H]uridine for 2 min. Nuclear RNA was hybridized to an excess of single-stranded DNA from an M13 bacteriophage clone of polyomavirus *Hpa*II fragment 3L. This fragment extends between nt 12 and 4409 on our strain of polyomavirus DNA (see Figure 4B), and includes all of the closely-spaced initiation sites for L-strand transcription, which lie between nt 5050 and 5205 (Tseng and Acheson, 1986). After RNase T1 and S1 nuclease digestion, hybrids were analyzed by gel electrophoresis followed by fluorography. Figure 4A shows the result of hybridizing nuclear RNA labeled for 2 min in cells infected with wild-type virus (lane 1) or with ins 5 virus (lane 2) to *Hpa*II fragment 3L. Both samples gave rise to hybrid bands corresponding to labeled RNAs whose 5′ ends lie at initiation sites 1 to 7, the major hybrid being band 4 (Tseng and Acheson, 1986). In addition, wild-type virus gave rise to a full-length read-through (RT) hybrid band, which was as prominent as band 4 (lane 1). This full-length hybrid arises exclusively from RNA made by polymerases which did not terminate transcription during a traverse of the viral DNA. In contrast, RNA from ins 5-infected cells gave rise to a RT band which contained a negligible amount of radioactivity compared to that in band 4 (lane 2). Clearly, then, termination on ins 5 DNA is much more efficient than on wild-type DNA.

To determine the frequency of transcription termination, gel lanes similar to lanes 1 and 2 of Figure 4A (but from an independent experiment) were sliced and the radioactivity in each 1-mm slice was measured. Figure 5 shows the profile obtained from such measurements. The results confirm the observations made from the autoradiogram; the ratio of radioactivity in full-length hybrids (RT RNA) to that in less-than-full-length hybrids (newly initiated RNA) was much greater in wild-type than in ins 5-infected cells.

To compare molar amounts of hybrids of different lengths, these results were normalized by dividing the radioactivity associated with each band by the number of uridine residues calculated to be present in each RNA, and multiplying by 100. The results are shown in Table II. Column A represents the radioactivity associated with full-length hybrids, generated by polymerases that completed at least one genome traverse before pasing through the initiation region again. Column B represents the sum of radioactivity associated with less-than-full-length hybrids (faster-migrating bands generated by polymerases that initiated transcription at sites 1−7 during the labeling period). As shown for two independent experiments, the efficiency of transcription termination on ins 5 DNA was 94−98% per genome traverse, compared to 78−81% on wild-type DNA. This means that only 2−6% of RNA polymerases that initiated transcription on the L strand of ins 5 DNA read through the entire viral genome without terminating, compared to 19−22% on wild-type DNA.

Therefore, insertion of a 94-nt fragment containing the efficient rabbit β-globin polyadenylation signal into the L strand of the polyomavirus genome significantly increased the efficiency of termination of L-strand transcription by RNA polymerase II. We conclude that this 94-nt fragment, which contains all the sequences required for efficient

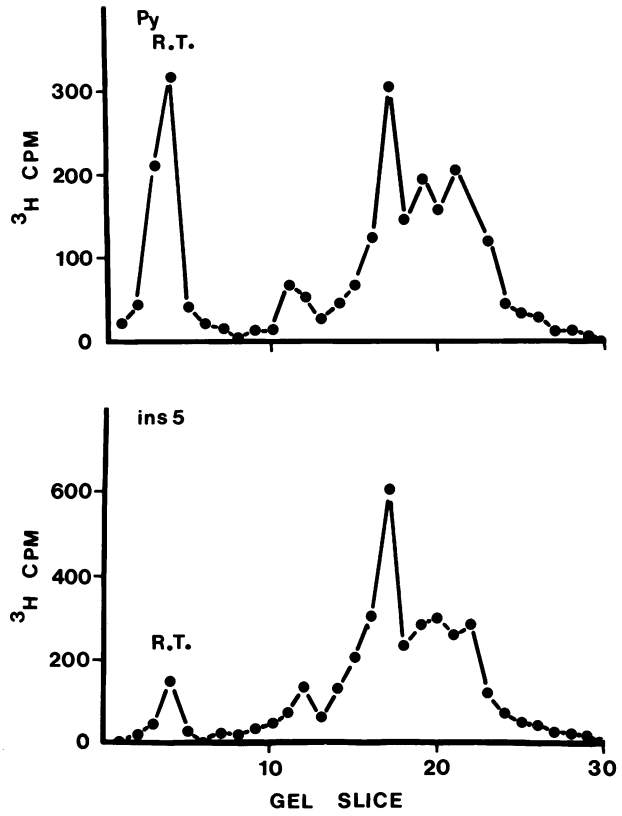

Fig. 5. Radioactivity profile from sliced gel: nuclear RNA from cells infected with wild-type or ins 5 hybridized to *Hpa*II fragment 3L. Nuclear RNA labeled for 2 min from cells infected with wild-type virus (1.2×10^6 c.p.m.) or ins 5 virus (5.8×10^5 c.p.m.) was hybridized to 10 μg of DNA fragment *Hpa*II-3L. Hybrids were treated as described in legend to Figure 4. After electrophoresis, a gel lane from each sample was sliced and radioactivity in each slice was determined. The position of full-length or read-through RNA (RT) is indicated. Newly-initiated RNA (NI) is represented by the sum of radioactivity in all the other peaks.

cleavage and polyadenylation of mRNAs, also contains a signal which leads to efficient termination of transcription by RNA polymerase II.

Discussion

This paper provides evidence for a direct relationship between polyadenylation and termination of transcription. Insertion of a 94-nt fragment containing the strong rabbit β-globin polyadenylation signal into the L strand of the polyomavirus genome simultaneously increased the efficiency of polyadenylation of L-strand RNA and that of transcription termination on the L DNA strand. Our experiments demonstrate that the inserted 94-nt fragment that contains all the sequences required for polyadenylation (Gil and Proudfoot, 1987), can also signal termination of transcription.

Two recent papers (Whitelaw and Proudfoot, 1986; Logan *et al.*, 1987) have also concluded that termination of transcription requires a functional polyadenylation signal. In both cases, efficient termination on the wild-type gene was reduced by mutations within the AATAAA motif of the polyadenylation signal. In contrast, we have introduced an efficient polyadenylation signal into a gene where both

Table II. Efficiency of transcription termination on L strand of wild-type or ins 5 DNA

Experiment	Virus	c.p.m./100 U residues[a] in:		Percent termination[b]
		RT RNA hybrids	Newly initiated RNA hybrids	
1	wild-type	65	275	80.9
	ins 5	14	599	97.8
2	wild-type	295	1070	78.4
	ins 5	110	1808	94.3

[a]Counts per minute in each peak from sliced gels such as shown in Figure 4 were divided by the number of U residues present in each RNA class, determined from the sequence of polyomavirus DNA and the known 5′ and 3′ ends of each RNA, and then multipied by 100.
[b]Percent termination per genome traverse was calculated as 100 × (c.p.m./100 U residues in newly-initiated RNA hybrids)/(c.p.m./100 U residues in read-through + newly initiated RNA hybrids) (Tseng and Acheson, 1986).

polyadenylation and termination were known to be poor, and have corrected these defects.

What makes one polyadenylation (termination) signal stronger than another? It has been shown that AATAAA alone is not sufficient to direct efficient cleavage and polyadenylation either *in vivo* or *in vitro*, but that other sequence elements downstream of this motif are required (Gil and Proudfoot, 1984; 1987; McDevitt *et al.*, 1984; McLauchlan *et al.*, 1985; Lanoix *et al.*, 1986; Mason *et al.*, 1986, and others). However, the exact nature of these downstream elements is not yet known (Mason *et al.*, 1986; McDevitt *et al.*, 1986; Conway and Wickens, 1987; Gil and Proudfoot, 1987). Specific complexes have been shown to protect from ribonuclease T1 digestion a 67-nt region of the adenovirus L3 pre-mRNA, including the AATAAA and a GT-repeat (Humphrey *et al.*, 1987). Mutations downstream of the polyadenylation site which prevent cleavage also prevent formation of this complex (Conway and Wickens, 1987; Zarkower and Wickens, 1987). It is probable that a number of more-or-less conserved sequence elements downstream of AATAAA contribute to the efficiency of recognition of polyadenylation signals by providing a better binding site for the factors involved.

However, it is clear that elements other than nucleotide sequence can affect the efficiency of transcription termination. Termination of simian virus 40 transcription is efficient in monkey cells (Ford and Hsu, 1978), but not in *Xenopus* oocytes (Miller *et al.*, 1982). Termination is less efficient during the late than during the early phase of infection in the adenovirus major late transcription unit and also in inserts whose transcription is directed by the major late promoter (Falck-Pedersen *et al.*, 1985). In our own experiments, we previously found that ~50% of RNA polymerases terminated during each traverse of wild-type polyomavirus DNA (Tseng and Acheson, 1986; Lanoix *et al.*, 1986). However, we now obtain values of nearly 80% (Table II). We do not presently know the reason for this discrepency, but we have recently changed to a new, low passage line of 3T6 cells, and it could be that the physiological state of these cells is different from that of our previous line. These differences do not affect the present results, however, because in all cases we have infected the same batch of cells with mutant and wild-type virus, to be able to make direct comparisons.

It remains to be seen whether other sequence elements besides the polyadenylation signal are required for transcription termination. Logan *et al.* (1987) showed that insertion of a fragment containing the mouse β-globin polyadenylation signal into the adenovirus E1A gene was not sufficient to cause efficient termination. They concluded that sequence elements in the region where termination takes place, ~1 kb downstream from the polyadenylation site, are also required for efficient termination. This finding would appear to conflict with our results which show that the sequences present in the 94-nt fragment of rabbit β-globin DNA were sufficient to increase the efficiency of termination of transcription. However, it is possible that additional termination signals are already present downstream of the polyadenylation signal in polyomavirus DNA, and that their use is simply increased in the presence of the inserted rabbit β-globin polyadenylation signal.

On the other hand, termination may take place in regions defined not by sequence on the DNA template, but rather by distance from the polyadenylation signal. We measured termination by quantitating read-through transcription *in vivo* in the initiation region of polyomavirus DNA, ~2.5 kb downstream of the polyadenylation signals. In our insertion mutants, downstream sequences were not changed from the wild-type sequence. Logan *et al.* (1987) measured termination by quantitating transcription in isolated nuclei from a region of the E1A gene located immediately downstream of their insert. The distance from the polyadenylation signal therefore varied depending on the size of the insert. Only further experiments in which different downstream sequence elements are tested for their effect on termination will elucidate the requirement for a second downstream element.

Two models may explain how a polyadenylation signal causes termination of transcription (see Acheson, 1984; Lanoix *et al.*, 1986; Platt, 1986; Logan *et al.*, 1987). Cleavage of the mRNA precursor at the polyadenylation site may destabilize the complex between DNA, RNA polymerase and RNA, and cause it to dissociate at positions downstream of the polyadenylation site. Alternatively, RNA polymerase may undergo a transition from an elongation mode to a termination mode upon passage through a polyadenylation signal. This transition could occur by the acquisition of a termination factor or by the loss of an elongation factor as the polymerase progresses through the polyadenylation signal.

The first model postulates that cleavage of the RNA chain is a prerequisite for subsequent termination. The second model would not require that cleavage precede termination; in addition, sequence elements required for cleavage/polyadenylation might not be identical to elements required for termination. Further experiments are required to define the sequences within polyadenylation signals that are essential for termination of transcription. The development of an *in vitro* system that carries out authentic termination by RNA polymerase II will help resolve questions concerning these mechanisms.

Because both termination of transcription and cleavage/polyadenylation of RNA are inefficient, wild-type polyomavirus produces 'giant' RNAs which contain multiple, tandem copies of the L strand (Acheson, 1978, 1984). These multigenome-length transcripts are spliced in the cell nucleus to produce L-strand mRNAs that code for viral capsid proteins VP1, VP2 and VP3. L-strand mRNAs contain, near their 5′ ends, tandem copies of a 57-nt 'leader' (Legon *et al.*, 1979), which are spliced together from regions originally spaced 5.3 kb apart in multigenome-length transcripts. It has been proposed that these multiple leaders may play a role in increasing translation of late mRNAs (Legon, 1979), although others have shown that the sequence of the leader region can be altered without affecting late gene expression (Adami and Carmichael, 1986). We have now constructed a virus (ins 5) in which the efficiency of termination of transcription on the L DNA is nearly 100%. This virus therefore should not make the multigenome-length transcripts that give rise to mRNAs with multiple leaders. Yet ins 5 replicates as well as wild-type virus in tissue culture cells, yielding equivalent titers of infectious virus and producing equivalent amounts of virus-specific RNA during a productive infection (data not shown). We thus have been able to short-circuit, with no ill effect on the virus, an intricate system involving transcription, polyadenylation and splicing whose end result was the production of mRNAs with multiple leaders. In this light, ins 5 can be looked upon as a virus which has evolved, in our hands, into a more tightly run and perhaps a more efficient organism.

Materials and methods

Virus, cells and materials

These were described elsewhere (Lanoix *et al.*, 1986). Polyomavirus strain AT3 (Skarnes *et al.*, 1988) was used in all experiments. 3T6 Swiss albino cells were obtained from American Type Culture Collection (Maryland, USA). S1 nuclease was from Boehringer Mannheim Biochemicals (Dorval, Quebec, Canada). ^{32}P-labeled deoxynucleoside triphosphates and [5-^{3}H]uridine were purchased from Amersham Corp. (Arlington Heights, IL).

RNA preparation and DNA – RNA hybridization

Cell fractionation and RNA extraction were as previously described (Acheson, 1978). Poly(A)$^+$ RNA was selected by oligo(dT) cellulose chromatography (Aviv and Leder, 1972). Nuclear RNA was pulse-labeled with [5-^{3}H]uridine, hybridized to single-stranded DNA, and treated with ribonuclease T1 and S1 nuclease, as described in Tseng and Acheson (1986).

S1 nuclease mapping of steady-state mRNA

This was essentially as described in Lanoix *et al.* (1986). Conditions for each experiment are given in the appropriate figure legend.

Transfection

'Clone 4', derived from the plasmid pβ5′SV (Gil and Proudfoot, 1984), was a gift from N.Proudfoot. It is a recombinant plasmid that contains the entire rabbit β-globin gene and a duplicated polyadenylation signal. The duplication extends between nt 1626 and 1799, and contains 100 nt downstream of AATAAA. This construct is referred to in the text as the rabbit β-globin plasmid.

Transfection of the rabbit β-globin plasmid into 3T6 cells was carried out by the calcium phosphate precipitation method of Graham and van der Eb (1973). Plasmid DNA (10 μg) and calf thymus DNA (10 μg) were used to transfect each 100-mm Petri dish of 3T6 cells that were seeded 24 h earlier at 5×10^5 cells/plate. The precipitate was allowed to form at 37°C and 5% CO_2 for 8 h. The DNA solution was removed, cells were washed with phosphate-buffered saline containing 1 mM EGTA and a second time with saline alone. Cells were then covered with 10 ml of fresh Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, and incubated at 37°C and 5% CO_2. RNA was extracted 24 h after transfection, as described previously (Acheson, 1978).

Construction of ins 5 and ins 6

The rabbit β-globin plasmid was digested with *Sfa*NI and *Aha*III to release a 94-nt fragment that contains a functional rabbit β-globin polyadenylation signal: the AATAAA motif, 6 nt downstream of the *Sfa*NI cleavage site, plus 82 nt downstream of AATAAA. The *Sfa*NI end of this fragment was blunt-ended with the Klenow fragment of DNA polymerase I, and *Sal*I linkers (New England BioLabs) were ligated to both ends. Small linker oligomers

were removed by passsage through a NACS ion exchange column (Bethesda Research Laboratories, Inc., Gaitherburg, MD), and the remaining DNA fragments were treated with *SalI* to remove multiple linkers and to create *SalI* sticky ends. The resulting fragment was cloned into a unique *SalI* site located between the 69-bp duplication in the plasmid parent of polyomavirus ins 3, which we had previously constructed (Lanoix *et al.*, 1986). After transformation of DH-5 *Escherichia coli* cells, colonies on kanamycin plates were screened for the presence of recombinant plasmids by colony hybridization. Positive colonies were further characterized by restriction analysis to determine the orientation of the β-globin insert. In ins 5, the β-globin polyadenylation signal is inserted in the direction of L-strand transcription, and in ins 6, it is inserted in the direction of E-strand transcription.

Acknowledgements

We would like to thank Harvey Miller for excellent technical assistance, and Nicholas J.Proudfoot for providing the β-globin plasmids. This work was supported by the Medical Research Council of Canada (grant MA-7281). JL was supported partly by a studentship from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (Quebec), and by a McGill University McConnell studentship, and is presently a recipient of a McGill University Faculty of Medicine studentship.

References

Acheson,N.H. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 4754–4758.
Acheson,N.H. (1981) *J. Virol.*, **37**, 628–635.
Acheson,N.H. (1984) *Mol. Cell. Biol.*, **4**, 722–729.
Adami,G.R. and Carmichael,G.G. (1986) *J. Virol.*, **58**, 417–425.
Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
Bogenhagen,D.F. and Brown,D.D. (1981) *Cell*, **24**, 261–270.
Citron,B., Falck-Pedersen,E., Salditt-Georgieff,M. and Darnell,J.E. Jr (1984) *Nucleic Acids Res.*, **22**, 8723–8731.
Cole,C.N. and Santagelo,G.M. (1983) *Mol. Cell. Biol.*, **3**, 267–279.
Conway,L. and Wickens,M. (1987) *EMBO J.*, **6**, 4177–4184.
Falck-Pedersen,E., Logan,J., Shenk,T. and Darnell,J.E. Jr (1985) *Cell*, **40**, 897–905.
Ford,J.P. and Hsu,M.T. (1978) *J. Virol.*, **28**, 795–801.
Gil,A. and Proudfoot,N.J. (1984) *Nature*, **312**, 473–474.
Gil,A. and Proudfoot,N.J. (1987) *Cell*, **49**, 399–406.
Graham,F.L. and van der Eb,A.J. (1973) *Virology*, **52**, 456–476.
Grummt,I., Rosenbauer,H., Niedermeyer,I., Maier,U. and Ohrlein,A. (1986) *Cell*, **45**, 837–846.
Hagenbuchle,O., Wellauer,P.R., Cribbs,D.L. and Schibler,U. (1984) *Cell*, **38**, 737–744.
Humphrey,T., Christofori,G., Lucijanic,V. and Keller,W. (1987) *EMBO J.*, **6**, 4159–4168.
Koenig,E.P., Hoffman,E.P., Bertelson,C.J., Monaco,A.P., Feener,C. and Kunkel,L.M. (1987) *Cell*, **50**, 509–517.
Lanoix,J., Tseng,R.W. and Acheson,N.H. (1986) *J. Virol.*, **58**, 733–742.
Legon,S. (1979) *J. Mol. Biol.*, **134**, 219–240.
Legon,S., Flavell,A., Cowie,A. and Kamen,R. (1979) *Cell*, **16**, 373–378.
LeMeur,M.A., Galliot,B. and Gerlinger,P. (1984) *EMBO J.*, **3**, 2779–2786.
Logan,J., Falck-Pedersen,E., Darnell,E. Jr and Shenk,T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8306–8310.
Mason,P.J., Elkington,J.A., Lloyd,M.M., Jones,M.B. and Williams,J.G. (1986) *Cell*, **46**, 263–270.
McDevitt,M.A., Imperiale,M.J., Ali,H. and Nevins,J.R. (1984) *Cell*, **37**, 993–999.
McDevitt,M.A., Hart,R.P., Wong,W.W. and Nevins,J.R. (1986) *EMBO J.*, **5**, 2907–2913.
McLauchlan,J., Gaffney,D., Whitton,J.L. and Clements,J.B. (1985) *Nucleic Acids Res.*, **13**, 1347–1368.
Miller,T.J., Stephens,D.L. and Mertz,J.E. (1982) *Mol. Cell. Biol.*, **2**, 1581–1594.
Platt,T. (1986) *Annu. Rev. Biochem.*, **55**, 339–372.
Skarnes,W.C., Tessier,D.C. and Acheson,N.H. (1988) *J. Mol. Biol.*, in press.
Tooze,J. (ed.) (1980) *DNA Tumor Viruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Tseng,R.W. and Acheson,N.H. (1986) *Mol. Cell. Biol.*, **6**, 1624–1632.
Whitelaw,E. and Proudfoot,N. (1986) *EMBO J.*, **5**, 2915–2922.
Zarkower,D. and Wickens,M. (1987) *EMBO J.*, **6**, 4185–4192.

Received on March 28, 1988; revised on May 17, 1988