Volume 16 Number 8 1988 Nucleic Acids Research

# Cis-acting sequences affecting the length of the poly(A) head of vaccinia virus late transcripts

Luisa de Magistris and Hendrik G.Stunnenberg*

European Molecular Biology Laboratories, Meyerhofstrasse 1, PO Box 10.2209, Heidelberg, FRG

Received February 19, 1988; Revised and Accepted March 22, 1988

ABSTRACT

Mutations in the sequences flanking the conserved and essential TAAAT motif of vaccinia late gene promoters (consensus: T/A T/A **TAAAT** G Pu Pu) affect the level of expression. Introduction of a pyrimidine in the purine stretch downstream of the TAAAT motif reduces the level of RNA synthesis. Mature transcripts from the wild-type 11K late promoter have a non-contiguous 5' poly(A) leader of approximately 35 A-residues (referred to as a poly(A) head). We show here by RNA sequencing, primer extension and subsequent $m^7G$ cap selection of cDNA/RNA hybrids that the mutations affect the length of the poly(A) head but not the location of the junction between the poly(A) leader and sequences encoded in the genome. These results are consistent with a slippage mechanism underlying the process of 5' poly(A) addition, but are not in agreement with a splicing event.

## INTRODUCTION

The genes encoded in the vaccinia virus genome can be divided into two classes: immediate early/early genes which are expressed before DNA replication and late genes transcribed post-replicationally. Early and late promoters do not share sequence elements and both are specifically recognized by the virus encoded RNA polymerase and not by the cellular polymerases (1). Late class promoters are characterized by a highly conserved TAAAT motif which is an essential cis-acting regulatory element (2, 3). The sequences surrounding this motif are partially conserved as reflected in the consensus sequence T/A T/A **TAAATG** Pu Pu (2); the translation start codon is frequently part of the TAAAT motif.

Transcription initiation was initially mapped within the TAAAT motif based on S1 protection experiments (2,3). This suggested that a putative untranslated leader sequence of late messengers was not encoded for in the genome. Structural analysis of the 5' termini of late messengers

revealed that additional poly(A) sequences were present at the 5' end of mature vaccinia late transcripts which were not encoded in the viral genome (4, 5). Poly(A) sequences were also shown to be present at the 5' ends of late transcripts encoding the major protein component of the A-type inclusion bodies produced by the closely related cowpox virus (6). It has also been shown that the poly(A) stretch is relatively discrete in length (~ 35 nt) and is preceded by a $m^7G$ cap structure (5). The capped poly(A) leader RNA will be referred to as a poly(A) head. Analysis of transcripts obtained in an in vitro cell-free transcription system derived from virus infected HeLa cells indicates that the poly(A) head is made de novo and its synthesis is directly coupled with the transcription of the coding body of the messenger (7). Our in vitro results are consistent with a slippage mechanism as postulated by Kassavetis and collaborators for transcription in the late phase of bacteriophage T4 infection (8). If transcription initiation starts at the first A-residue of the TAAAT motif, the primary RNA transcript pppApApA slips in a 3'->5' direction. As a consequence the proximal T-residue on the coding strand becomes unpaired and can subsequently be transcribed. Several rounds of slippage and addition of an A-residue would result in a 5' poly(A) stretch.

The mutational analysis presented here was undertaken to genetically discriminate between a slippage mechanism as proposed on the basis of our biochemical data and alternative processes such as splicing. The putative role of the conserved TAAAT sequence motif in the process of A-addition cannot be investigated directly because mutations in this sequence motif result in the inactivation of the promoter (2,9). We have, therefore, investigated whether sequences flanking the TAAAT motif affect the process of poly(A) addition. We have analyzed the effect of the mutations on the level of RNA and protein synthesis and the structure of the putative poly(A) heads added to the 5' termini of the transcripts.

## MATERIALS AND METHODS

### Virus and Cells

Vaccinia virus (strain WR) was grown in RK-13 monolayers or HeLa S3 suspension cultures that were maintained in Eagle medium containing 10% fetal calf serum (RK-13 and Human 143) or 5% new born calf serum (HeLa).

Preparation of recombinant viruses

Recombinant viruses were prepared using the temperature sensitive mutant ts7 (10) essentially according to the procedure of Kieny et al. (11).

Construction of promoter mutants

The construction of the mutants Δ6 and F was described previously (2). Synthetic double stranded oligonucleotides, containing the TAAAT sequence motif and an insertion or deletion at position +3, were inserted downstream of the Cla I restriction site in the mutant F and the Eco RI site of the vaccinia thymidine kinase (tk) locus (2). The mutant 11-stop was constructed by digestion of the plasmid 111-dhfr (2) with the restriction enzymes Eco RI and Bam HI, 5' overhangs were filled in by Klenow enzyme and the plasmid was re-ligated. The promoter mutants were fused to the mouse dehydrofolate reductase (dhfr) gene (12).

Purification of RNA and S1 mapping.

RNA from Hela or RK-13 cells infected wild-type or recombinant vaccinia virus (m.o.i. of 5) was extracted 16 hours post-infection. Purification of RNA was done as described (13). Two μg of total RNA were hybridized at 44 $^{o}$C to asymmetrically end-labelled DNA probes as described (2). After overnight hybridization single stranded DNA was digested with S1 nuclease for 1 h at room temperature (genomic mapping) or at 14 $^{o}$C (mapping of the A-stretch). In vitro synthesized dhfr RNA was obtained from a Gem vector-dhfr plasmid using T7 RNA polymerase as indicated by the manufacturer (Boehringer).

Primer extension and RNA sequencing

Primer extensions were performed using 5' end-labelled synthetic oligonucleotides as described previously (5). In RNA sequencing experiments, the deoxy- to dideoxynucleotide ratios were 5.5, 1.4, 1.4 and 2.7 for T-, A-, G- and C-reactions, respectively

Cap selection

The cDNA-RNA products obtained by primer elongation were phenol extracted and precipitated. RNase treatment was performed for 15 minutes at room temperature using 1 μg of RNase A in 10 mM Tris-HCl (pH 7.5) - 1 mM EDTA and 5 Units of RNase T2 in 50 mM NaAc (pH 4.5) - 2 mM EDTA - 150 mM NaCl per 5 μg of total RNA. Binding of the cDNA-RNA hybrids to rabbit polyclonal anti-m$^{7}$G serum (14) were done as described previously (5)

Pulse labeling of newly synthesized polypeptides

RK-13 cell monolayers (3 cm dishes) were either mock-infected or infected with wild-type or recombinant virus (m.o.i. of 5) and incubated at 37 °C for 16 hours. Subsequently, cells were incubated in the absence of methionine for 1 hour and the polypeptides were pulse-labelled using 100 µCi/ml of L-[$^{35}$S]-methionine for one hour in methionine free medium. Cells were harvested, resuspended in 0.5 ml of 50 mM Tris-HCl (pH 7.5) - 150 mM NaCl - 1 % Triton-X-100 and 4 µg/ml of phenylmethylsulfonylfluoride (PMSF) and incubated for 30' at 0 °C. Nuclei and cellular debris were removed by low speed centrifugation and an aliquot was TCA precipitated and applied to a 10-15 % polyacrylamide-SDS gradient gel (15).

Immune precipitation.

An aliquot of the cell lysate was incubated for one hour at 4 °C with a rabbit polyclonal anti-dhfr serum. Protein-A Sepharose CI-4B equilibrated in buffer A (10 mM Tris-HCl (pH 7.5) - 150 mM NaCl - 0.2 % Triton-X-100 - 2 mM EDTA) was added. and the incubation was continued for an additional hour. The beads were collected by low speed centrifugation and consecutively washed three times with buffer A, twice with the buffer A + 500 mM NaCl and once with 10 mM Tris-HCl (pH 7.5). The beads were taken up into SDS sample buffer, boiled for 10', 0.25 vol of 0.5 M Iodine Acetate Amide and loaded onto a SDS-PAA gel.

Sources of materials.

Restriction enzymes and RNase A were purchased from Boehringer, Mannheim and AMV reverse transcriptase and RNase inhibitor from Genofit, Geneva. Protein A-sepharose CL-4B and nuclease S1 were obtained from Pharmacia and Heparin (174.000 IU/g) from Serva, Heidelberg. RNases T2 was obtained from BRL. Radioactive nucleotides and methionine as well as [$^{14}$C]-labeled molecular weight protein marker were purchased from Amersham and the media for cell culture from Gibco.

## RESULTS

Comparison of several late promoters revealed a highly conserved and late specific TAAAT sequence motif as well as flanking sequences which are partially conserved: A- and T-residues are frequently preceding the TAAAT element, whereas purines are preferred

downstream of this motif (2). The consensus and the sequence of the different promoter mutants are listed in Table 1.

S1 analysis

The promoter mutant-dhfr gene constructs were stably inserted into the thymidine kinase (tk) gene of vaccinia virus by homologuos recombination (16, 17). RNA was prepared 16 hrs post-infection and subjected to S1 analysis using genomic DNA probes. Transcripts originating from the gene coding for the 11 kDa polypeptide (located at the junction between the HindIII-E and -F fragments) were mapped simultaneously and regarded as an internal control (Fig. 1). The transcriptional activity of the wild type (wt) translocated 11K promoter was slightly reduced compared to the authentic 11K promoter (Fig. 1). The reduction was less prominent then previously reported (Hänggi et al). Deletion of three nucleotides from position +8 to +10 (mutant 11-stop), introducing a stop codon between the ATG present in the 11K promoter fragment and the ATG preceding the dhfr gene did not affect the transcriptional activity. The substitution of nucleotides in the region from +3 to +9 (Δ6 promoter) and the additional conversion of the nucleotides from -7 to -4 into their complement (mutant F) did not have a significant effect on the level of RNA synthesis (Fig. 1). The Δ6 promoter strength was previously reported to be reduced compared to the wt translocated promoter (2). This reduced level of transcription is only obtained with that particular recombinant virus preparation and

Table 1: promoter sequences

| Consensus: | A/T A/T **TAAAT** G Pu Pu | | | |
|---|---|---|---|---|
| | -10 | +1 | +11 | +21 | |
| wt | TATGCTATAA | ATGAATTCCG | GATCCGGCAT | CATG | dhfr |
| 11-stop | TATGCTATAA | ATGAATT--- | GATCCGGCAT | CATG | dhfr |
| Δ6 | TATGCTATAA | AT--AAA--G | GATCCGGCAT | CATG | dhfr |
| F | TATCGATTAA | AT--AAA--- | GATCCGGCAT | CATG | dhfr |
| $ATA^+$ | TATCGATTAA | ATaGAATTCC | GGATCCGGCA | TCATG | dhfr |
| $ATT^+$ | TATCGATTAA | ATtGAATTCC | GGATCCGGCA | TCATG | dhfr |
| $ATC^+$ | TATCGATTAA | ATcGAATTCC | GGATCCGGCA | TCATG | dhfr |
| $ATT^-$ | TATCGATTAA | ATtAATTCCG | GATCCGGCAT | CATG | dhfr |
| $ATC^-$ | TATCGATTAA | ATcAATTCCG | GATCCGGCAT | CATG | dhfr |

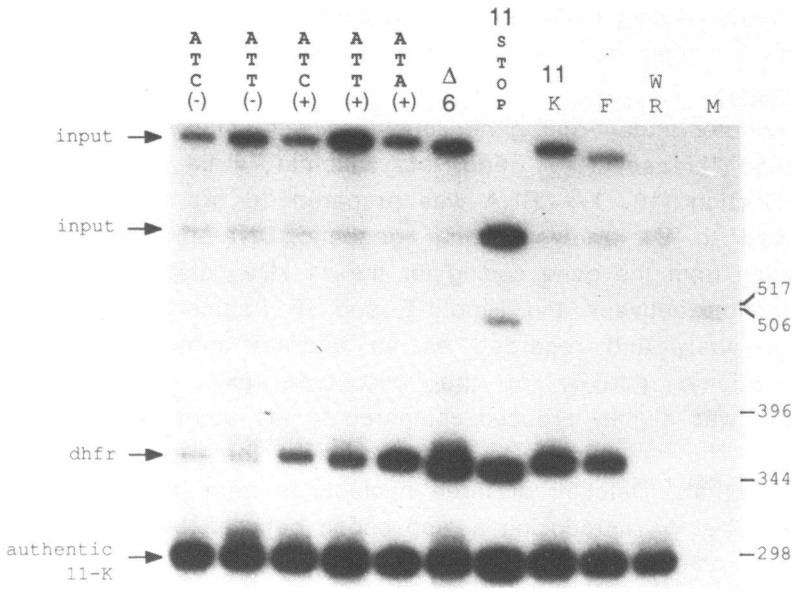


Figure 1: S1 analysis. RNA was prepared from wild-type or recombinant virus infected RK-13 cells 18 hours post infection. S1 mapping was performed as described using asymmetrically labeled genomic DNA fragments. Fragments resistant against nuclease S1 were analyzed on a sequencing polyacrylamide gel. (M): [$^{32}$P]-labeled Hinf I-digested pBR-322 DNA size markers.

might be due to an additional mutation in the promoter region. The insertion of an A-residue at position +3 within the conserved purine stretch (ATA+) did not significantly affect the promoter strength either, whereas insertion of a pyrimidine (ATT+ and ATC+, respectively) at the same position resulted in a 2-3 fold reduction of the level of RNA synthesis as compared to ATA+ and wt (Fig.1). Substitution of the G-residue at position +3 by a pyrimidine (ATT$^-$ and ATC$^-$) results in an 5-10 fold reduction of the promoter activity as compared to the ATA+ mutant (Fig. 1). These results indicate that a pyrimidine at position +3 strongly reduces the promoter strength. This effect is less pronounced if a pyrimidine is inserted at position +3 leaving three purines (and the G-residue) in close proximity of the TAAAT motif (mutants ATT+ and ATC+ compared to ATT$^-$ and ATC$^-$, respectively (Fig. 1).

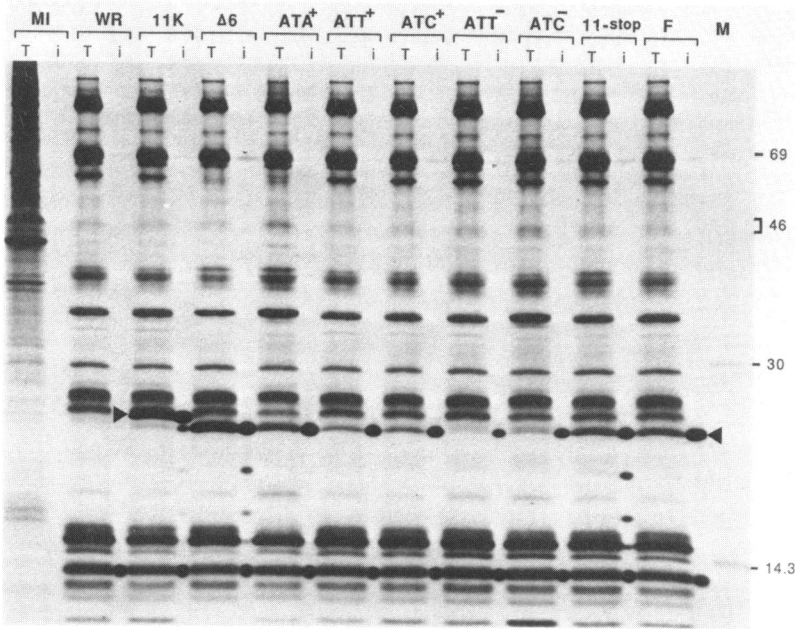


Figure 2: Pulse labeling and immunoprecipitation. Hela S3 cells were mock-infected (MI) or infected with wild-type (WR) or recombinant virus at a m.o.i. of 5. Newly synthesized polypeptides were labeled for one hour with [$^{35}$S]-methionine after starvation. Aliquots of total cell lysates were TCA precipitated (T) or immunoprecipitated using anti-dhfr antiserum (i) as described and applied to a SDS-polyacrylamide gradient (10-15 %) gel. The arrows indicate the position of authentic and dhfr fusion protein. (M): [$^{14}$C]-labeled high molecular weight standard protein markers.

Protein synthesis.

The level of protein synthesis in cells infected with wt virus (strain WR) or recombinant viruses is determined in a pulse labeling experiment using [$^{35}$S]-methionine and fluorography after gel electrophoresis (Fig. 2). Immune precipitation of the newly synthesized dhfr polypeptides have been performed solely to determine the position of the polypeptides within the gel (indicated by arrow heads). The major new polypeptide synthesized in cells infected with the recombinant virus containing a wt 11K promoter fragment is a fusion protein of 25.5 kDa (Fig. 2). A low level of authentic dhfr protein (24.5 kDa) is detectable

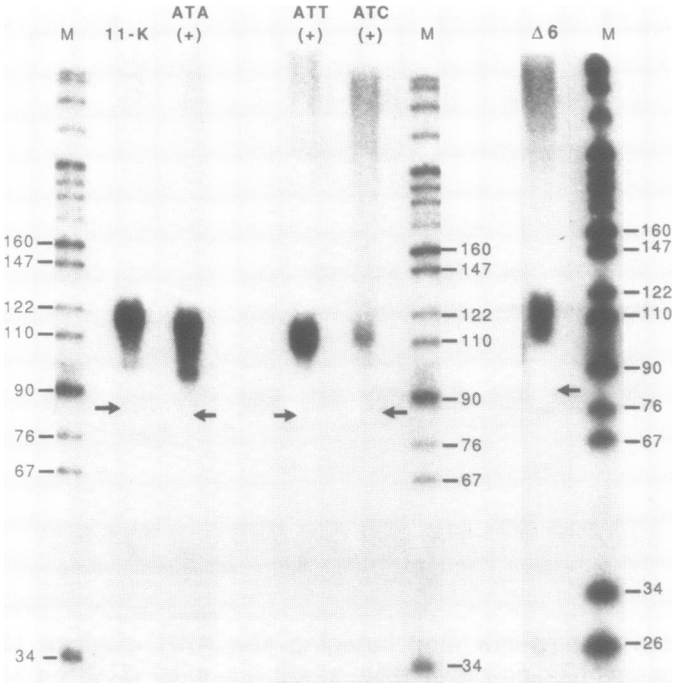


Figure 3: Primer elongation and cap-selection. RNA from cells infected with recombinant viruses containing 11K promoter mutants-dhfr chimeric genes were primer extended using a 5'-labeled oligonucleotide complementary to dhfr RNA sequences (+42 to +66 from the A-residue of the dhfr AUG)) as described. cDNA-RNA hybrids were incubated in the presence of RNase A and $T_2$ and were cap-selected by immunoprecipitation using a rabbit anti-$m^7G$ antiserum and analyzed on a sequencing polyacrylamide gel as described. The arrows indicate the position of the S1 start site as revealed by genomic S1 mapping. (M): [$^{32}$P]-labeled Hpa II-digested pBR-322 DNA size markers.

indicating that the second ATG is used at low frequency for translation initiation. The introduction of a stop codon (11-stop) between the ATG in the promoter fragment and the ATG preceding dhfr coding sequences results in the synthesis of slightly reduced levels of authentic dhfr protein (Fig. 2). This indicates that the ribosomes are able to efficiently re-initiate at the second AUG. High levels of authentic dhfr protein are produced in cells infected with Δ6 recombinant virus (Fig. 2). The lower levels of dhfr protein synthesized from the ATT+, ATC+, ATT$^-$ and ATC$^-$

chimeric constructs can be explained by the reduced RNA levels as determined by S1 mapping (Fig. 1). Slightly reduced protein levels are obtained with the mutants ATA+ and F compared to the wt and Δ6 promoters (Fig. 2). This effect is reproducible and cannot be explained by a reduced transcription level (Fig. 1).

Primer extension and cap selection.

The structure of the 5' termini have been further analyzed by primer extension and subsequent cap selection as described previously (5). An end-labeled oligonucleotide is extended by reverse transcriptase and full-length cDNA-RNA hybrids are selected using antibodies directed against the $m^7G$ cap structure found at the 5' end of mature mRNAs (14). RNase A and $T_2$ treatment prior to cap selection of the cDNA-RNA hybrids removes the cap structure at the 5' end of RNA which is not full-length protected by complementary DNA (Fig. 3). The vast majority of the selected cDNAs complementary to RNA transcripts obtained from the wt translocated promoter are approximately 35 nt longer than is anticipated based on S1 mapping (Fig. 3). We have shown that this leader RNA consists of a capped homopolymeric A stretch (5). The leader RNAs at the 5' end of Δ6 transcripts appear to be more variable in size banding in two size classes of approximately 25 and 35 nt in length (Fig. 3). The primer extension products obtained with transcripts from the mutants F (not shown) and ATA+ appear to be less heterogeneous in size than the Δ6 transcripts, whereas.significantly shorter (15-25 nt) primer extension products are obtained with ATT+ and ATC+ transcripts (Fig. 3). The selected cDNAs migrating in the upper part of the gel appear to have very long leader sequences at the 5' termini. We have not been able to characterize these cDNAs in more detail because they are only obtained at low levels.

RNA sequencing

RNA transcripts originating from the different promoter mutants have been sequenced to determine whether poly(A) sequences are present at their 5' ends. Furthermore, the location of the junction between the putative leader RNA and sequences which are encoded in the DNA can be established. RNA sequencing is complicated by the presence of readthrough transcripts in the late phase of infection due to the absence of termination signals (1). Transcripts initiated at upstream promoters are synthesized which contain downstream promoter-gene sequences. The presence of such transcripts can be demonstrated by insertion of an

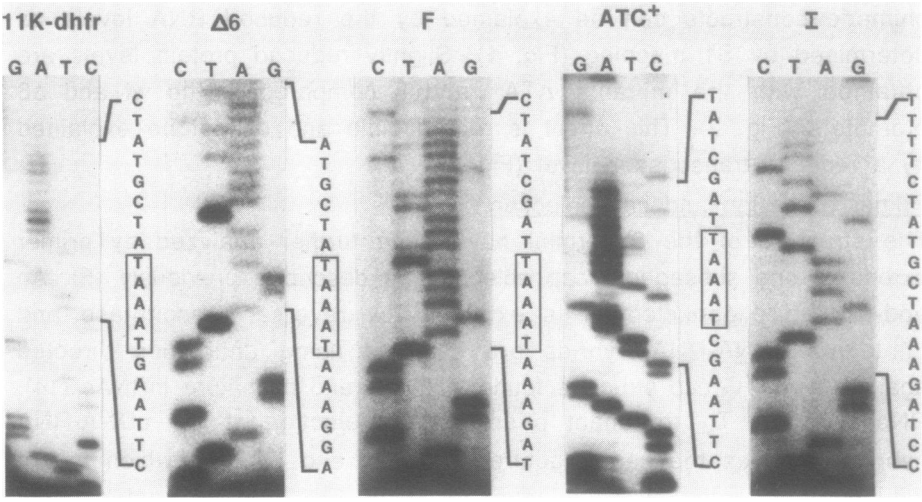


Figure 4: RNA sequencing. RNA from cells infected with 11K mutant-dhfr recombinant virus were primer elongated in the presence of dideoxy-nucleotides using a 5'-labeled synthetic oligonucleotides complementary to dhfr RNA sequences (+3 to +25 from the A-residue of the dhfr AUG). The promoter-dhfr gene sequences surrounding the TAAAT motif are indicated. (I) an inactive 11K promoter mutant dhfr chimeric gene.

inactive promoter mutant-dhfr construct containing a deletion/substitution of the TAAAT motif and flanking sequences (I): only readthrough transcripts containing dhfr sequences are detectable (Fig. 4).

Poly(A) tracts which are not encoded in the genome are present at the 5' termini of transcripts from all different promoter mutants (only wt, Δ6, F and ATC+ are shown Fig. 4) Furthermore, the location of the junction between the discontinuous leader and sequences co-linearly encoded in the genome is in all cases within the TAAAT motif. The presence of a TAAAG motif in the mutants Δ6 and F immediately following the TAAAT box does not affect the location of the RNA junction site (Fig. 4).

S1 mapping of the A-heads.

The size distribution of the poly(A) tracts at the 5' end of the transcripts can be further analyzed by nuclease S1 analysis as described previously (5). Artificial DNA fragments are constructed

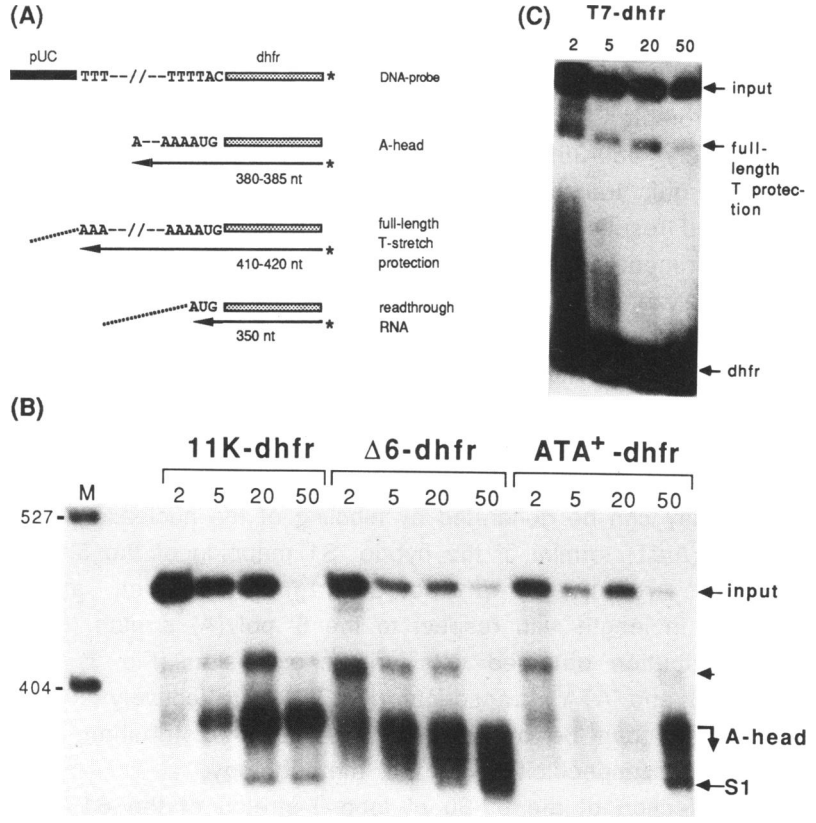


Figure 5: S1 analysis of the 5' poly(A) head. (**A**): Schematic presentation of the S1 mapping using an end-labeled DNA fragment containing a homopolymeric T-stretch (50-60 nt in length) replacing 11K promoter sequences upstream of the three A-residues of the TAAAT motif.(**B**): S1 analysis of RNA transcripts from recombinant viruses using increasing amounts of nuclease S1 as indicated above the lanes. The arrow head indicates the position of DNA fragments corresponding to full-length protection of the homopolymeric T-stretch. (S1): indicates the position of dhfr transcripts lacking a poly(A) stretch. (**C**): S1 analysis of dhfr transcripts synthesized in vitro from a T7 promoter-dhfr construct in the presence of RNA from cells infected with the wild-type (WR) virus using increasing amounts of nuclease S1 as indicated. (M): [$^{32}$P]-labeled Hpa II-digested pBR-322 DNA size markers.

containing a stretch of 50 to 60 A-residues upstream of the co-linearly transcribed DNA sequences. S1 probes derived from these artificial DNA constructs are complementary to the RNA transcripts including their

discontinuous 5' poly(A) stretch (Fig. 5A). The results of S1 mapping of 11K, Δ6 and ATA+ transcripts using increasing concentrations of nuclease S1 are shown in Fig. 5B. Analysis of transcripts from the wt translocated promoter results in the generation of protected fragments of approximately 380-385 nt in length (Fig. 5B). These fragments correspond to dhfr transcripts including a 5' leader RNA of 30-35 A-residues. These results are in agreement with the primer extension/cap selection experiments indicating that the cap selection procedure can be used to analyze mature mRNAs (Fig. 3). Increasing concentrations of nuclease S1 results in the generation of protected fragments migrating at the position of the S1 "start site" as revealed by S1 mapping using genomic probes (Fig. 1). These fragments represent readthrough transcripts which are initiated at promoter sequences upstream of the inserted 11K-dhfr chimeric gene and lacking a poly(A) head Furthermore, they can be generated by nibbling of the nuclease S1 at the homopolymeric(A/dT) termini of the hybrid. S1 mapping of the 5' ends of Δ6 transcripts results in protected fragments which are very heterogeneous in length with respect to the 5' poly(A) stretch (Fig. 5B). The size distribution obtained with primer extension (Fig 3) and S1 mapping of Δ6 and ATA+ transcripts (fig. 5B) is effectively equivalent. Minor differences can be observed in the length distribution probably reflecting different specificities of the methodology.

Full-length protection of the 50-60 nt long T-stretch of the S1 probe is also obtained (indicated by the arrow head). These protected fragments might reflect the presence of transcripts with long and heterogeneous leader sequences as obtained in primer extension experiments (Fig. 3). Alternatively, the T-stretch might be protected non-specifically by poly(A) sequences present within the RNA preparation. We have tested this hypothesis by S1 mapping of in vitro synthesized dhfr transcripts (lacking 5' poly(A) sequences) in the presence of RNA from wild-type virus infected cells (lacking dhfr sequences) (Fig. 5C). A smear is obtained at low nuclease S1 concentrations; addition of more enzyme results in the generation of a fragment of discrete size corresponding to protection of the in vitro synthesized dhfr RNA. Full-length protection of the T-stretch (indicated by the arrow head) is obtained even at high enzyme concentrations. This full-length protection can only be generated if heteroduplexes formed between the T-stretch of the S1 probe and poly(A) sequences present at the 5' and/or 3' ends of wild-

type vaccinia RNA. This does not rule out, however, the possibility that transcripts with very long and heterogeneous poly(A) stretches are present at very low levels in the infected cell.

## DISCUSSION

The conserved TAAAT sequence motif plays a crucial role in the transcription regulation of vaccinia virus late genes. Mutations within this sequence motif result in the inactivation of the promoter (2, 9). Furthermore, the motif is at the junction between a non-contiguously encoded poly(A) leader RNA and sequences which are co-linearly transcribed from the viral genome (4, 5, 6). Finally, analysis of late transcripts synthesized in vitro in a cell-free transcription system derived from virus infected cells strongly suggests that transcription initiation is occurring on the A-residues of the TAAAT motif (7).

We have shown that the insertion of an A-residue downstream of the sequence motif at position +3 within the purine stretch (ATA+) did not significantly affect the promoter strength, whereas insertion of a pyrimidine at the same position (ATT+ and ATC+, respectively) resulted in a 2-3 fold reduction of RNA synthesis. The promoter activity is further reduced in the mutants ATT- and ATC- in which the G-residue at position +4 is substituted by a pyrimidine (Fig. 1). It can be concluded that the conserved purine stretch downstream of the TAAAT motif (with a preference for a G-residue immediately following or in close proximity) is important to retain a high level of transcription. The mutations do not seem to have a significant effect on the translation efficiency because the amount of protein obtained with the different mutants correlates very well with the RNA level as determined by S1 mapping (Fig. 1) with the exception of ATA+ and F which express slightly lower amounts of protein than anticipated from the level of transcription. The promoter mutants Δ6 and ATA+ (followed by multiple cloning sites within the final insertion vector) are succesfully used in our laboratory to express high levels of authentic polypeptides (manuscript in preparation).

RNA transcripts have been analyzed further by primer extension and binding of the cDNA/RNA hybrids to antibodies directed against the $m^7G$ cap structure (5, 14). Pre-incubation of the hybrids in the presence of RNases A and T2 prior to antibody binding ensures that the selected cDNAs reflect full-length copies of the mature, capped mRNAs. The

transcripts from the different promoter mutants appear to have 5' termini which are variable in size depending on the mutation. The most pronounced effect is obtained with transcripts from the Δ6 promoter (Fig. 3). The variation in the length of Δ6 transcripts can be attributed to a heterogeneity in the length of the 5' terminal poly(A) head. In fact, RNA sequencing reveals that poly(A) sequences are present at the 5' ends of the transcripts and the location of the junction between the poly(A) head and the encoded sequences is not affected (Fig. 4) Furthermore, S1 mapping using artificial DNA probes (containing a homopolymeric T-stretch in the transcribed strand) shows that the poly(A) heads are heterogeneous in size (Fig. 5). The fact that the data obtained with S1 mapping of the A-heads and primer extension are effectively identical implies that the vast majority and not a minor population of the mature transcripts present within the virus infected cell are heterogeneous in size with respect to the poly(A) head. Furthermore, it shows that the antibody mediated cap selection is reliable and that the procedure can be used to analyze mature, capped mRNAs.

The mutational analysis is complicated by the fact that some of the analyzed promoter mutants contain mutations upstream of the TAAAT motif (the Cla I site) as well as downstream of the motif within the purine stretch. For example, the primer extension pattern obtained with ATA+ transcripts is more discrete than that obtained with Δ6 transcripts (Fig. 3). It is not possible to determine whether this is due to alterations in the purine stretch or to a compensatory effect due to the upstream transversions. Comparing Δ6 with wt or alternatively ATA+ with ATT+ and ATC+ it is evident, however, that the purine stretch downstream of the TAAAT is involved in the process of poly(A) head formation without affecting the location of the RNA junction. This finding implies that the poly(A) head at the 5' termini of transcripts originating from different late genes do not necessarily have to be identical in size depending on the sequence composition.

The observed variation in the length of the leader RNA is not compatible with poly(A) head formation occurring as the result of a splicing event. In such a model it would be anticipated that mutations affecting the putative splice acceptor site would either result in a reduced efficiency of splicing or alternatively in the use of "cryptic" acceptor sites. It would certainly not affect the location of the 5' splice site resulting in

poly(A) stretches that differ in size. We cannot rule out the possibility that the reduced level of RNA synthesis obtained with some of the promoter mutants is due to a reduced efficiency of poly(A) head addition. We have not observed, however, addition of poly(A) sequences at cryptic acceptor splice sites.

Analysis of RNA transcripts synthesized in vitro in a cell-free transcription system derived from vaccinia infected cells (7) suggests that the poly(A) head might be generated by a slippage mechanism (8). The results described in this study would be in agreement with a slippage mechanism if the degree of slippage is affected by strength of the RNA polymerase-promoter interaction. Mechanistic restraints are believed to play an important role during transcription initiation at prokaryotic promoters (18, 19). The latter proposed a model for early initiation in which competition between open complex polymerase-DNA contacts on one hand and initiated complex polymerase-DNA-RNA interactions on the other produces a stressed intermediate during formation of a short RNA-DNA duplex. In this model the strain energy can be relieved either by ejecting the short RNA, resulting in abortive initiation, or by eliminating the sigma subunit and thus escaping into productive transcription. They could show that destabilization of interactions specific for either the open or initiated complex had the predicted effect on the amount of abortive cycling. Similar interactions might play a role during initiation and subsequent slippage occurring during vaccinia late transcription. The strain energy might in this model be relieved by a displacement of the primary pppApApA transcript. An interaction of the poly(A) leader and a factor(s) of the vaccinia initiation complex could trigger the transition into productive transcription. Mutations within sequences involved in polymerase-promoter interactions might lead to an alteration in the conformation of the complex thereby affecting the specificity of poly(A) head formation. A detailed analysis of the mechanism of poly(A) head formation and putative interactions of the poly(A) head with components of the transcription complex is needed to verify this hypothesis.

ACKNOWLEDGEMENT

We thank Drs. Munns (St Louis) and Lanzer (Heidelberg) for their kind gift of anti-$m^7G$ and anti-dhfr antibodies, respectively. We thank Jacky Schmitt for technical assistance and our colleagues for discussions and critical reading of the manuscript.

*To whom correspondence and reprint requests should be addressed

REFERENCES

1. Moss,B. (1985) In Virology, ed. B.N. Fields pp 685-703. Raven Press, New York.
2. Hänggi,M., Bannwarth,W. and Stunnenberg,H. G. (1986) EMBO J **5**, 1071-1076.
3. Rosel,J.L.,Earl,P.L. Weir,J.P. and Moss,B. (1986) J. Virol. **60**, 436-449.
4. Bertholet,C., Van Meir,E., Heggeler-Bordier,B., and Wittek,R. (1987) Cell **50**, 153-162.
5. Schwer,B., Visca,P., Vos,J. C. and Stunnenberg,H. G. (1987) Cell **50**, 163-169.
6. Patel,D.D. and Pickup,D.J. (1987) EMBO J. **6**, 3787-3794.
7. Schwer,B. and Stunnenberg,H.G. (1988) EMBO J. April vol., in press.
8. Kassavetis,G. A., Zentner,P. G., and Geiduschek,E. P. (1986) J. Biol. Chem **261**, 14256-14265.
9. Bertholet, C., Stocco,P., Van Meir,E. and Wittek,R. (1986) Proc. Natl. Acad. Sci. USA **82**, 2096-2100.
10. Drillien,R. and Spehner,D. (1983) Virology **131**, 385-393.
11. Kieny,M.P., Lathe,R., Drillien,R., Spehner,D., Skory,S., Schmitt,D., Wiktor,T., Koprowski,H. and Lecocq,J.P. (1984) Nature **312**, 163-166.
12. Chang,A.C.Y., Nunberg,J.H., Kaufman,R.J., Erlich,H.A., Schimke,R.T. and Cohen,S.N. (1978) Nature **275**, 617-624.
13. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.
14. Munns,T.W., Liszewski,M.K., Tellam,J.T., Sims,H.F. and Rhoads,R.E. (1982). Biochemistry **21**, 2922-2928.
15. Laemmli,U.K. (1970) Nature **227**, 680-687.
16. Mackett,M., Smith,G.L. and Moss,B. (1982) Proc. Natl. Acad. Sci. USA. **79**, 7415-7419.
17. Panicali,D. and Paoletti,E. (1982).Proc. Natl. Acad. Sci. USA **79**, 4927-4931.
18. Carpuosis,A.J. and Gralla,J.D. (1985) J. Mol. Biol. **198**, 165-177.
19. Straney,D.C. and Crothers,D.M. (1987) J. Mol. Biol. **193**, 267-278.