# Identification of the Vaccinia Virus Gene Encoding an 18-Kilodalton Subunit of RNA Polymerase and Demonstration of a 5' Poly(A) Leader on Its Early Transcript

BYUNG-YOON AHN, ELAINE V. JONES, † AND BERNARD MOSS\*

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 5 January 1990/Accepted 22 March 1990

The DNA-dependent RNA polymerase of vaccinia virus contains 8 to 10 virus-encoded polypeptides. We have mapped the gene encoding an 18-kilodalton RNA polymerase subunit to D7R, the seventh open reading frame of the HindIII D genomic subfragment. Localization of this gene was achieved by using antibody to the purified RNA polymerase for immunoprecipitation of the in vitro translation products of in vivo-synthesized early mRNA selected by hybridization to cloned DNA fragments. The identification was confirmed by translation of D7R transcripts made in vitro with bacteriophage T7 RNA polymerase. The phenotypes of two previously isolated conditionally lethal temperature-sensitive mutants that map to D7R (J. Seto, L. M. Celenza, R. C. Condit, and E. G. Niles, Virology 160:110-119, 1987) are consistent with an essential role of this subunit in late transcription. This polymerase gene, designated rpo18, predicts a polypeptide of 161 amino acids with a molecular mass of 17,892. The rpo18 gene is transcribed early in infection, even though the 5'-TAAATG-3' motif, which is conserved among many genes of the late class, is present near the RNA start site. Characterization of the 5' end of the early transcript by several different methods, including cDNA cloning, revealed a poly(A) leader with up to 14 adenylate residues, whereas only 3 are present in the corresponding location of the DNA template. Similar but somewhat longer poly(A) leaders have previously been observed in mRNAs of late genes. We noted a TAAATG motif near the initiation site of several other early genes, including the viral DNA polymerase, and carried out additional experiments to demonstrate that their early transcripts also have 5' poly(A) leaders. Thus, formation of the poly(A) leader is not exclusively a late function but apparently depends on sequences around the transcription initiation site.

Vaccinia virus is a large double-stranded DNA virus that replicates in the cytoplasm of the host cell. The 185-kilobase-pair (kbp) genome encodes most, if not all, of the enzymes required for DNA replication and transcription, including a DNA polymerase, a DNA-dependent RNA polymerase, a poly(A) polymerase, and RNA capping and methylating enzymes (24). The expression of vaccinia virus genes is temporally regulated. Transcription of early genes is mediated by enzymes and factors packaged within the virion and begins immediately after infection, whereas viral DNA replication must precede the expression of late genes. The mechanisms underlying the transcriptional switch are not well understood, although different cis- and trans-acting factors are clearly involved (B. Moss, Annu. Rev. Biochem., in press).

The DNA-dependent RNA polymerase of vaccinia virus is a multisubunit enzyme (2, 25) composed of two large subunits of 147 and 132 kilodaltons (kDa) and six to eight small subunits of 17 to 34 kDa, all of which appear to be virus encoded (18). Whether all of the polypeptides represent unique gene products, however, has not been reported. The genes for the two large and one of the small 22-kDa RNA polymerase subunits of vaccinia virus or the closely related cowpox virus have been sequenced (4, 30). The predicted amino acid sequence of the large subunits showed significant similarity to the corresponding procaryotic RNA polymerase subunits and even greater similarity to the eucaryotic ones.

Poxviruses offer an excellent opportunity to combine biochemical and genetic approaches to the study of transcription. As a prerequisite, however, it is necessary to identify the genes encoding all of the enzymes and factors needed for RNA synthesis. In this report, we have located the gene (rpo18) for the 18-kDa subunit of vaccinia virus RNA polymerase. In addition, we found that the early transcripts of rpo18 and two other genes contain short 5' poly(A) leaders previously associated with late mRNAs.

# MATERIALS AND METHODS

Virus and cells. The Western Reserve strain of vaccinia virus was propagated in HeLa S3 Spinner cells in minimal essential Spinner medium (Quality Biologicals Inc.) supplemented with 5% horse serum.

Plasmid constructions. KpnI and SalI restriction endonuclease fragments of the vaccinia virus genomic HindIII D DNA segment were inserted into pUC19. Plasmid DNA prepared by alkaline lysis of bacteria was used for hybridization with viral RNA. For construction of the template used to synthesize rpo18 RNA in vitro, part of the HindIII D DNA fragment (nucleotides 8,751 to 9,320, numbering from the left HindIII site [27]) was amplified by 20 cycles of polymerase chain reaction (PCR). Amplification was performed in 100 μl of the standard reaction mixture containing 50 mM KCl, 10 mM Tris hydrochloride (pH 8.3), 1.5 mM MgCl<sub>2</sub>, 0.2 mM deoxynucleoside triphosphates, 1 μM of each primer, 0.01% (wt/vol) of gelatin, 0.1 to 1 μg of plasmid DNA, and 2 U of Taq polymerase. Reactions were for 2 min at 94°C, 2 min at 55°C, and 5 min at 74°C for denaturation,

In contrast, no protein homologous to the vaccinia virus 22-kDa polymerase subunit has been identified.

<sup>\*</sup> Corresponding author.

<sup>†</sup> Present address: Department of Molecular Genetics, Smith Kline Beecham Animal Health, Swedeland, PA 19479.

annealing, and polymerization, respectively. Amplified DNA was inserted into the Bluescript plasmid vector containing the promoter of bacteriophage T7 to construct plasmid pT7rpo18.

Selection of viral RNA, in vitro translation, and immunoprecipitation. Preparation of RNA by CsCl centrifugation and DNA hybridization was as described previously (7, 17). Early RNA was obtained from HeLa cells at 4 h after infection with 10 PFU per cell of vaccinia virus in the presence of 100 µg of cycloheximide per ml. Late RNA was isolated 6 h after infecting HeLa cells in the absence of cycloheximide. The in vitro transcript of the rpo18 gene was synthesized from linearized pT7rpo18 DNA by using T7 RNA polymerase and a cap  $m^7G(5')ppp(5')G$  primer (26). Hybrid-selected in vivo- or in vitro-synthesized RNAs were translated in the presence of [35S]methionine in micrococcal nuclease-treated rabbit reticulocyte lysates. The translated products were incubated with either preimmune serum or antiserum to vaccinia virus RNA polymerase (18), and the proteins that bound were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

RNA analysis. Nuclease S1 digestion and primer extension techniques were used to analyze the rpo18 transcript as previously described, with only minor modifications (1, 37). The complementary single-stranded DNA probe used in the nuclease S1 analyses was prepared by 30 cycles of asymmetric PCR of linearized plasmid DNA with a 5'-end-labeled primer (hereafter referred to as the c primer) complementary to the sequence downstream of the predicted transcriptional start site. The c primer for D7R was complementary to the sequence from nucleotides 8,823 to 8,845 (27), and the one for the DNA polymerase gene was complementary to the sequence from nucleotides 506 to 524 (11). T4 polynucleotide kinase and  $[\gamma^{-32}P]ATP$  were used to label the 5'-end of the c primers. The SpeI site located 770 base pairs upstream of the translational initiation codon of D7R and the EcoRI site 422 base pairs upstream of the DNA polymerase gene were used to linearize the plasmids containing the respective genes. The first step in the preparation of the poly(dT)-containing probe used to characterize the poly(A) leader was the amplification by 20 PCR cycles of a segment of the plasmid containing the target gene. The c primer and another primer containing 40 dA residues followed by the translational initiation codon and 20 nucleotides of the coding sequence were used for the PCR. The amplified DNA was subjected to 30 cycles of a second asymmetric PCR by using the 5'end-labeled c primer. The latter product, with a 3' poly(dT) tail, was directly loaded onto a 4% polyacrylamide gel containing 7 M urea, and the position of the probe was determined by a 5-min autoradiographic exposure at ambient temperature. The probe was electroeluted by applying a voltage of 150 V for 1 h to pieces of gel placed into a dialysis membrane. Typically, 1 µg of plasmid DNA and 20 pmol of end-labeled primer yielded a probe with approximately 1 ×  $10^8$  dpm/µg of DNA.

For the nuclease S1 analysis, 30 µg of early or late RNA was annealed at 42°C for 3 h with 1 pmol of probe in 30 µl of 40 mM PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)] (pH 6.5)–0.4 M NaCl-0.1 mM EDTA-80% formamide and then treated with 400 U of nuclease S1 at room temperature for 1 h in 300 µl of S1 buffer containing 30 mM sodium acetate (pH 4.5), 200 mM NaCl and 1 mM ZnSO<sub>4</sub>. Nuclease-resistant material was ethanol precipitated and analyzed by electrophoresis in a 5% polyacrylamide gel containing 7 M urea. For primer extension analysis, 30 µg of early RNA was annealed with the 5'-end-labeled c primer as described for

the S1 analysis. After ethanol precipitation, the dried material was suspended and incubated for 30 min at 42°C in the buffer containing avian myeloblastosis virus reverse transcriptase with or without dideoxynucleotides. The reaction mixture was loaded on the gel after addition of sequencing stop solution.

Conditions used for tailing of cDNA by the terminal deoxynucleotidyl transferase were as described previously (22).

Sources of materials. Restriction enzymes, T4 polynucleotide kinase, T4 DNA ligase, S1 nuclease, terminal deoxynucleotidyl transferase, and avian myeloblastosis virus reverse transcriptase were purchased from Boehringer Manheim Biochemicals. Sequenase was from United States Biochemicals Corp. The in vitro transcription system, including the T7 RNA polymerase and the rabbit reticulocyte lysate, were obtained from Promega Biotec. The Bluescript vector was from Stratagene. Taq polymerase and the thermal cycler used in PCR were products of The Perkin-Elmer Corp. Oligonucleotides were made by an Applied Biosystems DNA synthesizer 380B. Radioactive nucleotides were from Amersham Corp. Other reagents including dideoxynucleotides were purchased from Pharmacia, Inc.

### **RESULTS**

Identification of the rpo18 gene. We demonstrated previously that antiserum to the purified vaccinia virus RNA polymerase specifically immunoprecipitated the in vitro translation products of early viral mRNAs (18). By using genomic subfragments to isolate mRNAs by hybridization, the approximate locations of several RNA polymerase genes were determined. Thus, the mRNA encoding a subunit of about 19 kDa hybridized to the 16-kbp HindIII D genomic fragment. In the present study, further mapping of this gene was accomplished by hybridization of viral early mRNA to cloned KpnI and SalI restriction endonuclease segments of the HindIII D DNA fragment (Fig. 1B) and translation of the selected RNAs in micrococcal nuclease-treated reticulocyte extracts containing [35S]methionine. The antiserum immunoprecipitated several of the labeled small subunits of RNA polymerase from extracts programmed with total early RNA (Fig. 1A, lane T) but only one major translation product, estimated to be about 19 kDa, from a reticulocyte lysate programmed with RNA that had hybridized either to the entire HindIII D fragment (Fig. 1A, lane 4), the 10.8-kbp KpnI-HindIII subfragment (Fig. 1A, lane 3), or the 3.3-kbp SalI-SalI subfragment (Fig. 1A, lane 6). No translation products were immunoprecipitated from extracts programmed with RNAs that hybridized to other subfragments of HindIII D DNA (Fig. 1A, lanes 1, 2, 5, and 7), suggesting that only one RNA polymerase subunit is encoded in this region of the vaccinia virus genome.

The nucleotide sequence of the entire *HindIII* D fragment was reported by Niles et al. (27). Examination of the nucleotide sequence of the 3.3-kbp *SalI-SalI* segment indicated the presence of two major open reading frames (ORFs), D6R and D7R, with capacities to encode polypeptides of 68,362 and 17,892 daltons, respectively. The polypeptide encoded by D7R (from nucleotide 8,768 to 9,250) was considered to be a strong candidate for the RNA polymerase subunit, based on the similarity of its predicted molecular mass to the 19 kDa estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of the immunoprecipitated polypeptide. In order to confirm this prediction, the D7R ORF was cloned into a vector containing a bacte-

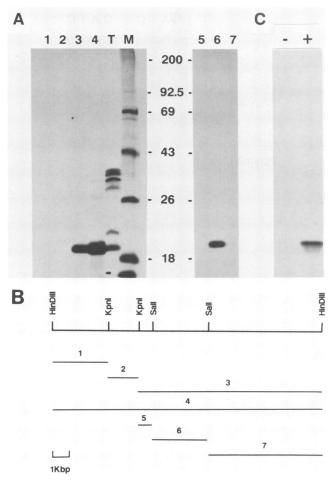


FIG. 1. Mapping of the rpo18 gene. (A) Total early RNA isolated from cells at 4 h after vaccinia virus infection in the presence of cycloheximide was hybridized to DNA fragments bound to a nitrocellulose membranes. Total or hybrid-selected mRNA was translated in a rabbit reticulocyte lysate in the presence of [35S]methionine, and the products were precipitated with antiserum to vaccinia virus RNA polymerase. An autoradiograph of the radioactively labeled protein, separated by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis, is shown. Lanes indicate the immunoprecipitated translation products of 10 µg of total early RNA (lane T) or of RNA that hybridized to the intact HindIII D fragment (lane 4) or various restriction enzyme subfragments (lanes 1 to 3 and 5 to 7), as defined in panel B. 14C-Labeled protein markers are shown in lane M. (B) A partial restriction map of the 16-kbp HindIII D fragment. The numbered bars represent DNA subclones used for RNA hybridization. (C) Translation and immunoprecipitation of the in vitro-synthesized rpo18 RNA. Plasmid pT7rpo18 containing the bacteriophage T7 promoter adjacent to the rpo18 gene was linearized with BamHI and transcribed by T7 RNA polymerase. One microgram of the rpo18 RNA was translated in reticulocyte extracts in the presence of [35S]methionine and precipitated with either preimmune serum (lane -) or antiserum to purified RNA polymerase (lane +).

riophage T7 promoter and transcribed in vitro by using T7 RNA polymerase. Translation of the capped in vitro RNA produced a polypeptide of 19 kDa which reacted with antiserum to the viral RNA polymerase but not with the preimmune serum (Fig. 1C). Based on these results, we concluded that the D7R ORF encodes a subunit of the vaccinia virus RNA polymerase and named this gene rpo18, according to its predicted molecular mass of 17,892.

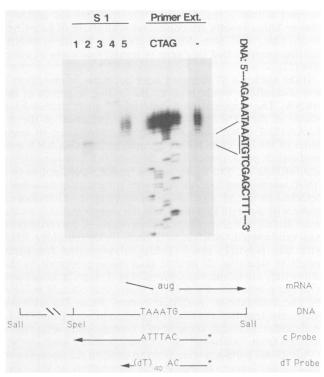


FIG. 2. Analyses of the 5' end of the rpo18 transcript. Total RNA was isolated from cells at 4 h after infection in the presence of cycloheximide or at 6 h in the absence of inhibitor and analyzed by nuclease S1 protection or primer extension. The two 5' 32P-labeled single-stranded DNA probes used for nuclease S1 protection have the same 5' ends and are indicated at the bottom of the figure. Also shown are the D7R mRNA and the 3.3-kbp SalI DNA fragment. The 850 nucleotide c probe contains a 78-nucleotide segment of the template strand of D7R and upstream sequences. The poly(dT) probe has the same 78 nucleotides of D7R, followed by 40 dT residues immediately 3' of the translation initiation codon. The preparation and precise location of the probes are described in Materials and Methods. Asterisks and arrowheads indicate labeled ends and polarity of strands, respectively. Lanes contain the following: untreated complementary probe (lane 1), complementary probe with early (lane 2) and late (lane 3) RNA, untreated (dT) probe (lane 4), and (dT) probe with early RNA (lane 5). In lanes 1 and 4, the 850-nucleotide and 118-nucleotide full-length probes, respectively, are not seen because the upper part of the autoradiograph has been cut off. The five lanes at the right contain extensions of the 5'-end-labeled primer in the presence of ddG (lane C), ddA (lane T), ddT (lane A), and ddC (lane G) and in the absence of any dideoxynucleotide (lane -). The 5' end of the primer is the same as that of the probes used for nuclease S1 protection. The DNA sequence around the translational initiation codon (underlined) of the rpo18 gene is shown on the right.

In a global computer search with the FASTA program (31), we could not find any sequence with significant similarity to rpo18 among either the translated nucleic acid sequences of GenBank (release 61) or the NBRF protein sequence library (release 21).

5' poly(A) leader in the transcript of rpo18. The 5' end of the rpo18 transcript was analyzed by the nuclease S1 procedure by using a 5'-end-labeled, single-stranded complementary DNA probe prepared by PCR. A major 5' end was detected with early RNA (Fig. 2, lane 2) but not with late RNA (Fig. 2, lane 3). By comparison with a sequence ladder in the same gel, from a primer extension experiment to be described shortly, the 5' end of the rpo18 transcript was

mapped within a few bases of the translational initiation codon. These results are consistent with those of Lee-Chen et al. (20), who also demonstrated by Northern (RNA) blotting analysis that this gene is transcribed early in infection (21).

If the nuclease S1 mapping were accurate, at most only a few nucleotides would precede the coding sequence of the mRNA. The potential of an extremely short untranslated leader of the D7R transcript and the presence of the 5'-TAAATG-3' DNA sequence which is conserved in many genes of the late class led us to consider that this RNA may contain a 5' poly(A) leader similar to those found naturally in late mRNAs (3, 35). Since such a leader would not have a DNA complement to hybridize with, it would not have been detected in the S1 analysis described above. Therefore, a probe with a poly(dT) sequence in the complementary strand immediately upstream of the coding region of the rpo18 gene was synthesized. Nuclease S1 analysis of the early RNA with this poly(dT) probe generated protected fragments longer in size (Fig. 2, lane 5) than those protected by the completely complementary probe (Fig. 2, lane 2), consistent with our prediction of a poly(A) leader immediately preceding the translational initiation codon. From the multiple bands, the length of the leader appeared to vary from 3 to 14 nucleotides, although some of this heterogeneity might have resulted from "nibbling" of A-T base pairs by nuclease S1.

Additional methods were used to confirm the presence of the poly(A) leader. Primer extension analysis of the early RNA yielded an array of bands similar in length to those protected from nuclease S1 by the dT probe and longer than those protected by the completely complementary probe (Fig. 2). Efforts were made to determine the sequence of the 5' ends by using dideoxynucleotides. The nucleotides complementary to AAATG could be read easily, but the sequence above this was ambiguous, probably due to the length hetrogeneity (Fig. 2). The latter problem was overcome by cDNA cloning.

A partial first-strand cDNA of the rpo18 transcript was synthesized by reverse transcriptase extension of the c primer, which is complementary to the sequence from 56 to 78 base pairs downstream of the translational initiation codon. A homopolymeric (dG) tail was added to the 3' end of the cDNA, and the latter was then amplified into doublestranded DNA by 25 cycles of PCR by using the c primer and a synthetic primer, 5'-GGAAGCTT(C)<sub>20</sub>-3'. The HindIII restriction site in the latter primer was designed for the subsequent cloning of the amplified DNA into a plasmid vector. Recombinant bacterial colonies were picked, and a sequence analysis of several cDNA clones revealed the presence of poly(dT), which varied between 6 and 10 nucleotides in length, followed by the oligo(dG) tail (Fig. 3). These results unambiguously identified the structure of a poly(A) leader present in the early transcript of the rpo18 gene.

Poly(A) leaders in other early transcripts. It was important to determine whether similar poly(A) sequences are present in other early mRNAs. Among those genes that were previously characterized as belonging to the early class, several have the TAAATG sequence at or near the experimentally determined RNA start site. These include the gene for the viral DNA polymerase (ORF E9L [11]), a candidate gene for a 20-kDa subunit of RNA polymerase (ORF A6R; Ahn et al., unpublished data), a gene encoding an 82-kDa polypeptide essential in viral DNA replication (ORF D5R [13, 33]), and an RNA capping enzyme subunit gene (ORF D12L [28]). Each of these genes, as well as rpo18, contains elements of the early promoter critical region sequence (8), indicated in

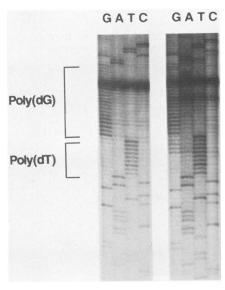


FIG. 3. Nucleotide sequence of cDNA clones containing the 5' end of the rpo18 transcript. Plasmids containing the amplified cDNA insert were alkali denatured and sequenced with Sequenase in the presence of  $[\alpha^{-35}S]dATP$ . Sequences of two positive clones are shown. The poly(dT) and poly(dG) sequences are indicated on the left

bold letters (Fig. 4), which have been used to line up the promoters. The TAAAT sequences occur 11 to 15 nucleotides downstream of the critical region, in accordance with the predicted locations of early RNA start sites (8). An analysis of the 5' ends of early E9L transcripts is shown in Fig. 5. By using a complementary probe, nuclease S1 analysis indicated major RNA start sites near the TAAAT sequence, consistent with the previous results of Earl et al. (11). As before, an array of longer protected fragments was obtained with a poly(dT) probe. The reverse transcriptase primer extension products were similar in length to the fragments protected by the poly(dT) probe and were consistent with a short 5' poly(A) leader of about 4 to 12 nucleotides. Very similar results also were obtained upon analysis of the 5' ends of the early A6R transcript (not shown). These data suggest that the 5' poly(A) leader is a feature shared by at least several early genes and therefore is not exclusively a function of late transcription.

- D7R CTGATGTAAAACTGATGAAATATATAAAGAAA<u>TAAATG</u>CGAGCTTT
  E9L GATTCATTAAATGAAAATATATTTCTAAATTCTA<u>TAAATG</u>ATGTT
  A6R AGGCTTTAAAATTGAATTGCGATTATAAGAT<u>TAAATG</u>CAGACACAG
  D5R TTAATGCTTTAGTGAAATTTTAACTTGTGTTC<u>TAAATG</u>ATGCGGCT
- D12L AACTAATAATAATGAAAACTATAGAGTTG<u>TAAATG</u>ATGAAAT

FIG. 4. Aligned nucleotide sequences of the promoters of several early vaccinia virus genes containing TAAATG motifs near the RNA start sites. The nucleotides constituting the early promoter critical region, as defined by Davison and Moss (8), are represented by bold letters. In each case, including E9L and A6R (Ahn, unpublished data), the ORF starts within the TAAATG.

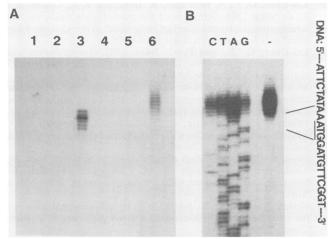


FIG. 5. Analyses of the 5'-end of the transcript for the vaccinia virus DNA polymerase gene. (A) RNA as well as complementary and dT probes were made and used for nuclease S1 analysis as described in the legend to Fig. 2. Lanes contain the following: untreated complementary probe (lane 1), complementary probe with 30 µg of yeast tRNA (lane 2) and early RNA (lane 3), untreated (dT) probe (lane 4), dT probe with tRNA (lane 5), and early RNA (lane 6). The 525- and 143-nucleotide full-length probes in lanes 1 and 4, respectively, are not seen because the upper part of the autoradiograph has been cut off. (B) Primer extension analysis of early RNA was performed as described in the legend to Fig. 2. The 5'end-labeled primer was annealed to early RNA and extended by reverse transcriptase in the presence of ddG (lane C), ddA (lane T), ddT (lane A), and ddC (lane G) and in the absence of any dideoxynucleotide (lane -). Shown on the right is the DNA sequence around the translational initiation codon (underlined).

## DISCUSSION

The vaccinia virus DNA-dependent RNA polymerase contains two large subunits and multiple small ones. Of the genes for the small subunits, only the one encoding a 22-kDa polypeptide had been mapped previously (4). Our identification of the D7R ORF as the gene encoding a subunit of about 19 kDa was based on the immunoprecipitation of the appropriate size translation products of in vivo and vitro transcripts with antibody to vaccinia virus RNA polymerase. This gene is essential, since two temperature-sensitive mutants have been mapped to the same ORF (36). Moreover, the phenotypes of these temperature-sensitive mutants are similar to those found for mutants of the 147- and 22-kDa RNA polymerase subunits (12, 15). At the nonpermissive temperature, normal levels of early gene expression and DNA replication occur but late gene expression is severely diminished (5, 6). Apparently, the RNA polymerase packaged into virus particles at the permissive temperature (and used for early gene expression) is not sensitive to later elevations of temperature, whereas the RNA polymerase synthesized at the nonpermissive temperature (and used for late gene expression) is temperature sensitive.

In contrast to the sequence similarity between the two large subunits of poxvirus RNA polymerase and corresponding subunits of procaryotic or eucaryotic origin, the predicted amino acid sequence of rpo18 shows no significant similarity to the sequence of any known genes. Since few eucaryotic RNA polymerase small-subunit sequences have been published (19, 23, 38), it is quite likely that the homolog of rpo18 has not been identified yet. Alternatively, some subunits of RNA polymerase may be distinctive in sequence

among diverse organisms. The two large subunits may be conserved to carry out functions shared by all RNA polymerases, such as chain elongation and substrate binding, whereas some of the small subunits may have more specialized roles.

Our RNA analyses confirmed those of Lee-Chen et al. (20) regarding the early transcription of the D7R gene. The rpo22 gene is also expressed early (4). In contrast, rpo132 and rpo147 have both early and late RNA start sites. Protein synthesis studies are needed, however, before considering the biological consequences of this apparent difference in gene expression. Nuclease S1 mapping placed the start site of the D7R transcript just upstream of the ORF and within a TAAAT sequence, reminiscent of the situation with late mRNAs (14, 32). Since late mRNAs contain a variablelength capped 5' poly(A) leader (1, 3, 29, 35), we examined the 5' end of the rpo18 mRNA for a similar structure. Two entirely different methods, nuclease S1 protection analysis with a probe containing poly(dT) attached to the complement of the ORF and primer extension with reverse transcriptase, indicated the presence of a short poly(A) leader of 4 to 14 nucleotides. The presence of this structure was confirmed by sequencing cDNA clones. Since the RNA used in the analyses was obtained at 4 h after infection in the presence of cycloheximide, the poly(A) leader is not derived from late transcripts expressed anomalously early in infection. The absence of a detectable nuclease S1 protected band with late RNA also discounts this possibility. A similar result was obtained with RNA made as early as 2 h after infection in the absence of cycloheximide, excluding the possibility of an artifactual effect of the drug. We found a 5' poly(A) leader on the early mRNAs of two additional early genes that contain a TAAATG sequence at the RNA start site, and Ink and Pickup (16) have found the poly(A) leader on another, suggesting that the structure is associated with the sequence rather than with the regulatory class of the promoter. This idea is further supported by mutagenesis of an early cowpox virus promoter (16).

Although the precise mechanism of poly(A) formation has not been determined, the data point to an RNA polymerase slippage mechanism requiring three consecutive A residues and influenced by neighboring sequences (9, 10, 16, 34). The lengths of 5' poly(A) leaders have been determined for only a few early and late mRNAs, but there does appear to be a size difference, with the latter being two to three times as long as the former. This difference may reflect the positioning of the RNA polymerase, with respect to the RNA initiation site, by early and late transcription factors. Mutagenesis studies to determine the role of the TAAAT sequence in 5' poly(A) formation has been difficult with late genes because this motif appears to be required for efficient promoter function. Despite the presence of the TAAATG sequence in some early genes, we suspect from promoter mutagenesis studies (8) that it is not essential for transcription. However, the poly(A) leader could be important for translation, since the predicted initiation codon for each of the early genes listed in Fig. 4 is contained within the TAAATG. Without the poly(A) addition to the 5' end, the untranslated leader would be virtually nonexistent. Studies to determine the mechanism of formation and function of poly(A) leaders on early mRNAs are in progress.

# ACKNOWLEDGMENTS

We thank Norman Cooper for providing cells and virus, Nelson Cole for preparing oligonucleotides, Patricia Earl and Paul Gershon for helping with computer analysis of nucleic acid sequences, and Michael Merchlinsky for helping in the preparation of the manuscript.

### LITERATURE CITED

- 1. Ahn, B.-Y., and B. Moss. 1989. Capped poly(A) leader of variable lengths at the 5' ends of vaccinia virus late mRNAs. J. Virol. 63:226-232.
- Baroudy, B. M., and B. Moss. 1980. Purification and characterization of a DNA-dependent RNA polymerase from vaccinia virions. J. Biol. Chem. 255:4372-4380.
- Bertholet, C., E. Van Meir, B. ten Heggeler-Bordier, and R. Wittek. 1987. Vaccinia virus produces late mRNAs by discontinuous synthesis. Cell 50:153-162.
- Broyles, S. S., and B. Moss. 1986. Homology between RNA polymerases of poxviruses, prokaryotes, and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. Proc. Natl. Acad. Sci. USA 83:3141-3145.
- Condit, R. C., and A. Motyczka. 1981. Isolation and preliminary characterization of temperature-sensitive mutants of vaccinia virus. Virology 113:224-241.
- Condit, R. C., A. Motyczka, and G. Spizz. 1983. Isolation, characterization, and physical mapping of temperature-sensitive mutants of vaccinia virus. Virology 128:224-241.
- Cooper, J. A., and B. Moss. 1979. In vitro translation of immediate early, early and late classes of RNA from vaccinia virus infected cells. Virology 96:368-380.
- Davison, A. J., and B. Moss. 1989. The structure of vaccinia virus early promoters. J. Mol. Biol. 210:771-769.
- 9. Davison, A. J., and B. Moss. 1989. The structure of vaccinia virus late promoters. J. Mol. Biol. 210:771-784.
- de Magistris, L., and H. G. Stunnenberg. 1988. cis-Acting sequences affecting the length of the poly(A) head of vaccinia virus late transcripts. Nucleic Acids Res. 16:3141-3156.
- Earl, P. L., E. V. Jones, and B. Moss. 1986. Homology between DNA polymerases of poxviruses, herpesviruses, and adenoviruses: nucleotide sequence of the vaccinia virus DNA polymerase gene. Proc. Natl. Acad. Sci. USA 83:3659-3663.
- Ensinger, M. J. 1987. Phenotypic characterization of temperature-sensitive mutants of vaccinia virus with mutations in a 125,000-M<sub>r</sub> subunit of the virion-associated DNA-dependent RNA polymerase. J. Virol. 61:1842-1850.
- Evans, E., and P. Traktman. 1987. Molecular genetic analysis of a vaccinia virus gene with an essential role in DNA replication. J. Virol. 61:3152-3162.
- Hänngi, M., W. Bannwarth, and H. G. Stunnenberg. 1986. Conserved TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. EMBO J. 5:1071-1076.
- Hooda-Dhingra, U., C. L. Thompson, and R. C. Condit. 1989.
   Detailed phenotypic characterization of five temperature-sensitive mutants in the 22- and 147- kilodalton subunits of vaccinia virus DNA-dependent RNA polymerase. J. Virol. 63:714-729.
- Ink, B. S., and D. J. Pickup. 1990. Vaccinia virus directs the synthesis of early mRNAs containing 5' poly(A) sequences. Proc. Natl. Acad. Sci. USA 87:1536-1540.
- Jones, E. V., and B. Moss. 1984. Mapping of the vaccinia virus DNA polymerase gene by marker rescue and cell-free translation of selected mRNA. J. Virol. 49:72-77.
- Jones, E. V., C. Puckett, and B. Moss. 1987. DNA-dependent RNA polymerase subunits encoded within the vaccinia virus genome. J. Virol. 61:1765-1771.
- Kolodziej, P., and R. Young. 1989. RNA polymerase II subunit RPB3 is an essential component of the mRNA transcription apparatus. Mol. Cell. Biol. 9:5387-5394.

- Lee-Chen, G. J., N. Bourgeois, K. Davidson, R. C. Condit, and E. G. Niles. 1988. Structure of the transcription initiation and termination sequences of seven early genes in the vaccinia virus HindIII D fragment. Virology 163:64-79.
- Lee-Chen, G. J., and E. G. Niles. 1988. Transcription and translation mapping of the 13 genes in the vaccinia virus HindIII D fragment. Virology 163:52-63.
- 22. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 217–219. Cold Spring Harbor Laboratory. Cold Spring Harbor, N.Y.
- Mann, C., J.-M. Buhler, I. Treich, and A. Sentenac. 1989.
   RPC40, a unique gene for a subunit shared between yeast RNA polymerase A and C. Cell 48:627-637.
- Moss, B. 1985. Replication of poxviruses, p. 685-703. In B. N. Fields, and D. M. Knipe, R. M. Chanock, J. Melnick, B. Roizman, and R. Shope (ed.), Virology. Raven Press, New York.
- Nevins, J. R., and W. K. Joklik. 1977. Isolation and properties of the vaccinia virus DNA-dependent RNA polymerase. J. Biol. Chem. 252:6930-6938.
- Nielsen, D. A., and D. J. Shapiro. 1986. Preparation of capped RNA transcripts using T7 RNA polymerase. Nucleic Acids Res. 14:5936.
- Niles, E. G., R. C. Condit, P. Caro, K. Davidson, L. Matusick, and J. Seto. 1986. Nucleotide sequence and genetic map of the 16-kb vaccinia virus HindIII D fragment. Virology 153:96-112.
- Niles, E. G., G.-J. Lee-Chen, S. Shuman, B. Moss, and S. S. Broyles. 1989. Vaccinia virus gene D12L encodes the small subunit of the viral mRNA capping enzyme. Virology 172: 513-522.
- Patel, D. D., and D. J. Pickup. 1987. Messenger RNAs of a strongly-expressed late gene of cowpox virus contains a 5'terminal poly(A) leader. EMBO J. 6:3787-3794.
- Patel, D. D., and D. J. Pickup. 1989. The second-largest subunit of the poxvirus RNA polymerase is similar to the corresponding subunits of procaryotic and eucaryotic RNA polymerases. J. Virol. 63:1076-1086.
- Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.
- 32. Rosel, J. L., P. L. Earl, J. P. Weir, and B. Moss. 1986. Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the *HindIII* H genome fragment. J. Virol. 60:436-439.
- Roseman, N., and D. E. Hruby. 1987. Nucleotide sequence and transcript organization of a region of the vaccinia virus genome which encodes a constitutively expressed gene required for DNA replication. J. Virol. 61:1398-1406.
- Schwer, B., and H. G. Stunnenberg. 1988. Vaccinia virus late transcripts generated in vitro have a poly(A) head. EMBO J. 7:1183-1190.
- 35. Schwer, B., P. Visca, J. C. Vos, and H. G. Stunnenberg. 1987. Discontinuous transcription or RNA processing of vaccinia virus late messengers results in a 5' poly(A) leader. Cell 50:163-169.
- Seto, J., L. M. Celenza, R. C. Condit, and E. G. Niles. 1987.
   Genetic map of the vaccinia virus HindIII D fragment. Virology 160:110-119.
- Weir, J. P., and B. Moss. 1984. Regulation of expression and nucleotide sequence of a late vaccinia virus gene. J. Virol. 51:662-669.
- 38. Woychik, N. A., and R. A. Young. 1989. RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. Mol. Cell. Biol. 9:2854–2859.