JOURNAL OF VIROLOGY, Jan. 1989, p. 226–232
0022-538X/89/010226-07$02.00/0

Vol. 63, No. 1

Capped Poly(A) Leaders of Variable Lengths at the 5′ Ends of Vaccinia Virus Late mRNAs

BYUNG-YOON AHN AND BERNARD MOSS*

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

Received 15 June 1988/Accepted 15 September 1988

Evidence for capped poly(A) leaders of variable lengths located immediately upstream of the translation initiation codon was obtained by direct analyses of a major late mRNA species. A decapping-recapping method was used to specifically substitute a radioactively labeled phosphate for an unlabeled one within the cap structure. RNase H-susceptible sites were made by hybridizing synthetic oligodeoxyribonucleotides to the mRNA encoding a late major structural protein of 11 kilodaltons. Sequences of the type $m^7G(5')pppAmp(Ap)_nUpG \cdots$, where *n* varies from a few to more than 40 nucleotides, were deduced by analysis of the length and sequence of RNase H, RNase T_1, and RNase U2 digestion products.

Vaccinia virus, a member of the poxvirus family, transcribes and replicates its 187,000-base-pair linear DNA genome in the cytoplasm of infected cells (22). Gene expression is temporally separated into pre- and postreplicative phases. Considerable information exists regarding the structure and synthesis of early mRNAs. These transcripts are of discrete length (10) with a methylated cap structure at the 5′ end (38) and a poly(A) tail at the 3′ end (15). There is no evidence for endo- or exonucleolytic processing of early mRNA: the initiating nucleotide forms part of the mature cap structure (37). The 3′ end appears to form by a termination mechanism (30, 34, 35), and splicing has not been observed. DNA sequences forming the early promoter (8, 41; A. J. Davison and B. Moss, manuscript in preparation) and termination signal (45) have been identified.

In comparison with the information cited above, much less is known about late transcription. Kinetic studies have suggested that at least two classes of mRNAs may be synthesized after DNA replication (24, 28). Accumulated evidence indicated that late mRNAs are heterogeneous in length (10, 11, 17), can form complementary duplex structures (7, 9), and are polyadenylated (25) and capped (5). The length heterogeneity, as well as the high sequence complexity of late RNA (6, 16, 26, 27), evidently results in part from the failure of late transcripts to stop at early termination signals (45). Unlike early mRNAs, which have cap sequences $m^7G(5')pppN_1pN_2 \cdots$, where N_1 is 2′-O-methylated adenosine or guanosine and N_2 is any of the four ribonucleosides in 2′-O-methylated or unmethylated form, late mRNAs have predominantly adenosine residues in methylated or unmethylated form in both the N_1 and N_2 positions (5). The subsequent nuclease S1 mapping of the 5′ ends of late mRNAs within the three A residues of the conserved TAAAT sequence just upstream of the open reading frames (31, 39) was consistent with the sequence of the cap structures. Nevertheless, subsequent reports indicated the presence of a novel modification in which additional A residues are present at the 5′ end of late mRNAs (4, 33). A complex structure consisting of a poly(A) segment interposed between the 3′ and 5′ ends of two unrelated RNAs was deduced by cDNA cloning and electron microscopy (4). However, a simpler structure consisting of a short 5′-terminal poly(A) leader was consistent with the analysis of reverse transcripts and cap-specific antibody binding (33). Preliminary functional analyses of late promoters indicated that they can be even shorter than early promoters (2, 3, 40) and that the TAAAT sequences must be conserved (3, 14; A. J. Davison and B. Moss, manuscript in preparation). Although the soluble transcription system derived from virus particles is specific for early promoters (30), a transcription system derived from cells infected for 6 h or longer will transcribe both early and late vaccinia virus genes (29, 32, 44). In these infected-cell extracts, RNAs transcribed from late genes contain a poly(A) leader (32, 44), whereas RNAs transcribed from early genes do not (44).

The apparent conflict in the interpretation of data regarding the structure of late mRNA from infected cells makes it difficult to proceed with mechanistic studies. Since previous conclusions were based principally on indirect analyses involving reverse transcriptase and S1 nuclease digestion of poly(A) · poly(dT) hybrids (4, 33), we considered that an alternative and more direct analysis of the 5′ ends of late mRNAs was crucial.

MATERIALS AND METHODS

Materials. Aminophenylthioester cellulose paper was obtained from Schleicher & Schuell, Inc., and was converted to diazo form as instructed by the manufacturer. Oligodeoxyribonucleotides were synthesized by the Applied Biosystem DNA Synthesizer model 380B. Ribonucleases, RNase inhibitor, and avian myeloblastosis virus reverse transcriptase were from Bethesda Research Laboratories, Inc., or Boehringer Mannheim Biochemicals. T7 RNA polymerase was from Promega Biotec. Marker nucleotides for cap structure and other nucleotides were from Pharmacia, Inc. [α-^{32}p]GTP (3,000 Ci/mmol) was from Amersham Corp. Capping enzyme (19), purified from vaccinia virions, was provided by Stuart Shuman.

Preparation and labeling of RNA. HeLa S3 cells were infected with 30 PFU of vaccinia virus per cell for 6 h. Cytoplasmic RNA was isolated by CsCl centrifugation as described previously (10). RNA encoding the 11-kilodalton (kDa) structural polypeptide was isolated by hybridization of the cytoplasmic RNA to complementary single-stranded DNA immobilized to cellulose paper. Single-stranded DNA was made by introducing a 3.7-kilobase-pair *Kpn*I-*Sst*I fragment of vaccinia virus DNA into M13mp18. RNA was specifically labeled at the cap structure by a decapping-

* Corresponding author.

recapping method. The terminal GMP of the cap structures was removed by periodate oxidation and β-elimination as described previously (23). Briefly, 2 to 5 μg of RNA was incubated in the dark at 0°C for 30 min in a 0.2-ml solution of 0.9 mM sodium periodate and 0.15 M sodium acetate (pH 5.3). The RNA was precipitated twice with 2.5 M ammonium acetate and ethanol. Subsequently, RNA was suspended in a 0.2-ml solution containing 0.3 M aniline and 0.01 M acetic acid adjusted to pH 5 with concentrated HCl. After 3 h at room temperature, the RNA was precipitated three times with ammonium acetate and ethanol. Conditions for the enzymatic guanylation and methylation of RNA were essentially the same as described previously (13, 20). For 2 to 5 μg of decapped RNA, the reaction mixture (0.1 ml) contained 50 mM Tris hydrochloride (pH 7.6), 2 mM dithiothreitol, 2.5 mM $MgCl_2$, 50 μM *S*-adenosylmethionine, 2 mM [α-^{32}P]GTP (3,000 Ci/mmol), and 5 to 10 U of capping enzyme. After incubation at 37°C for 30 min, unincorporated nucleotides were removed by passing the reaction mixture through a 5-ml column of Sephadex G-50 (Pharmacia) and subjecting it to two precipitations with ammonium acetate and ethanol.

Enzymatic digestions. For RNase H cleavage, cap-labeled RNA (total of 10^5 cpm) was annealed at 65°C for 15 min with 0.1 nmol of appropriate oligodeoxyribonucleotide in a 20-μl reaction mixture containing 400 mM Tris hydrochloride (pH 7.9), 40 mM $MgCl_2$, 4 mM dithiothreitol, 1 μg of yeast tRNA, and 1 U of RNase inhibitor. The temperature of the reaction was brought to 25°C for 30 min before 1 to 2 U of RNase H was added, and the RNA was incubated for 30 min at 37°C. The RNA was precipitated with ethanol and dried in vacuo. The samples were suspended in 10 μl of dye mixture containing 90% formamide, 50 mM EDTA, and 0.05% bromophenol blue. The samples were boiled to 95°C for 2 min before being applied to a polyacrylamide gel. Labeled RNA was digested with RNase T_1, nuclease P1, and snake venom phosphodiesterase as recommended by suppliers. For partial digestion with RNase U2, 0.1 U of enzyme was used in a 10-μl reaction mixture containing 25 mM sodium citrate (pH 3.5), 7 M urea, 1 mM EDTA, and 1 μg of yeast tRNA. The reaction was incubated at 55°C for 15 min.

Primer extension and dideoxy sequencing. 5′-End-labeled oligodeoxyribonucleotide was incubated with RNA at 65°C for 10 min in a 10-μl reaction mixture containing 34 mM Tris hydrochloride (pH 8.3), 50 mM NaCl, and 6 mM $MgCl_2$. The mixture was allowed to cool slowly to 42°C, and 20 U of reverse transcriptase, 1 U of RNase inhibitor, and 1 μl of 100 mM dithiothreitol were added. The reaction mixture was aliquoted into five parts of 2.5 μl each, and 2.5 μl of nucleotide mixture was added to each aliquot to a final concentration of 250 μM. The ratios of deoxy- to dideoxynucleotides were 4, 2, 2, and 2 for C, T, A, and G, respectively. For the primer extension, dideoxynucleotides were omitted. The reaction was incubated for 30 min at 42°C and was chased with 1 μl of 2.5 mM deoxynucleoside triphosphate mixture for another 30 min. To stop the reaction, 5 μl of a solution containing 90% formamide, 50 mM EDTA and 0.05% bromophenol blue was added. After being boiled for 2 min at 95°C, 2 to 3 μl of the mixture was applied to an 8 to 12% polyacrylamide gel containing 7 M urea. Electrophoresis was carried out for 2 h at 60 to 70 W, and an autoradiograph was made.

Two-dimensional gel electrophoresis. The two-dimensional gel system was a minor modification of that of deWachter and Fiers (12). The first-dimension gel (33 cm by 40 cm by 1.6 mm) was composed of 14% acrylamide–0.7% N,N′-methylene bisacrylamide containing 7 M urea. The final pH of the solution was adjusted to 3.5 by adding a saturated solution of citric acid. Electrophoresis was carried out at 350 V for 20 h in 0.025 M citric acid buffer (pH 3.5) until xylene cyanol moved halfway down the gel. A strip of gel (3 cm by 32 cm) was excised and soaked for 5 min in 0.45 M Tris-borate (pH 8.0) buffer containing 7 M urea. The strip was placed at the bottom of a second gel apparatus (33 cm by 40 cm by 1.6 mm), and the second gel solution (24% acrylamide, 1.2% bisacrylamide, 7 M urea, 90 mM Tris-borate, 2.5 mM EDTA [pH 8.0]) was added. Ascending electrophoresis was performed at 600 V for 18 h until bromophenol blue migrated up to 3 to 5 cm from the top of the gel. An autoradiograph was prepared by incubating the wet gel at −70°C with an intensifying screen for 24 to 48 h.

RESULTS

RNase H digestion of cap-labeled late mRNA. The first step in our analysis was to isolate the RNA encoding the major 11-kDa structural protein (43) by hybridization of cytoplasmic RNA to complementary single-stranded DNA immobilized on cellulose filters. The isolated RNA was then labeled by a decapping-recapping procedure that is specific for natural 5′ ends. Decapping was accomplished by periodate oxidation and β-elimination, and the resulting triphosphate-ended molecules were recapped by using [α-^{32}P]GTP and purified vaccinia virus RNA guanylyltransferase (21). The latter enzyme transfers the GMP moiety of GTP only to the 5′ ends of RNA containing a di- or triphosphate (18). Because of this specificity, RNA ends derived by strand scission are not labeled. When *S*-adenosylmethionine is present, a methyl group is transferred to the added GMP to form an m^7 G cap.

In a typical experiment, starting with 10 mg of cytoplasmic RNA obtained from approximately 1.2×10^9 infected cells, approximately 20 μg of RNA was recovered after DNA hybridization and 5 to 10 pmol of [α-^{32}P]GMP (3,000 Ci/mmol) was incorporated. When the periodate oxidation and β-elimination procedures were omitted, transfer of GMP to the RNA was reduced by 99%, indicating that decapping was required.

The next step was to cut the mRNA encoding the 11-kDa protein at a specific site and then analyze the cap-labeled 5′ ends by polyacrylamide gel electrophoresis. This was accomplished by hybridization of the RNA to an oligodeoxyribonucleotide corresponding to sequences within the gene for the 11-kDa protein and digestion with RNase H. The latter endoribonuclease is specific for RNA-DNA hybrids and thus allowed us to select the site of cleavage. Autoradiographs of polyacrylamide gels revealed diffuse bands that correlated in mobility with the distance of the hybridized oligonucleotide from the beginning of the gene for the 11-kDa protein (Fig. 1). However, the estimated length of the cap-labeled RNA bands exceeded that predicted for transcripts initiating within the three A residues (Fig. 1C, triangles) just upstream of the ATG on the basis of S1 nuclease mapping analysis (4, 33). The predicted sizes when using hybrid probes I and II would be 9 and 54 nucleotides, respectively, whereas the observed sizes were up to 35 nucleotides longer. These results were confirmed by using a higher-resolution gel in which a smear could be seen to be composed of finely spaced bands (data not shown).

The possibility that the heterogeneity in length was due in part to incomplete RNase H digestion was assessed by analysis of the cleavage products of a uniformly ^{32}P-labeled transcript with a discrete 5′ end synthesized in vitro by using

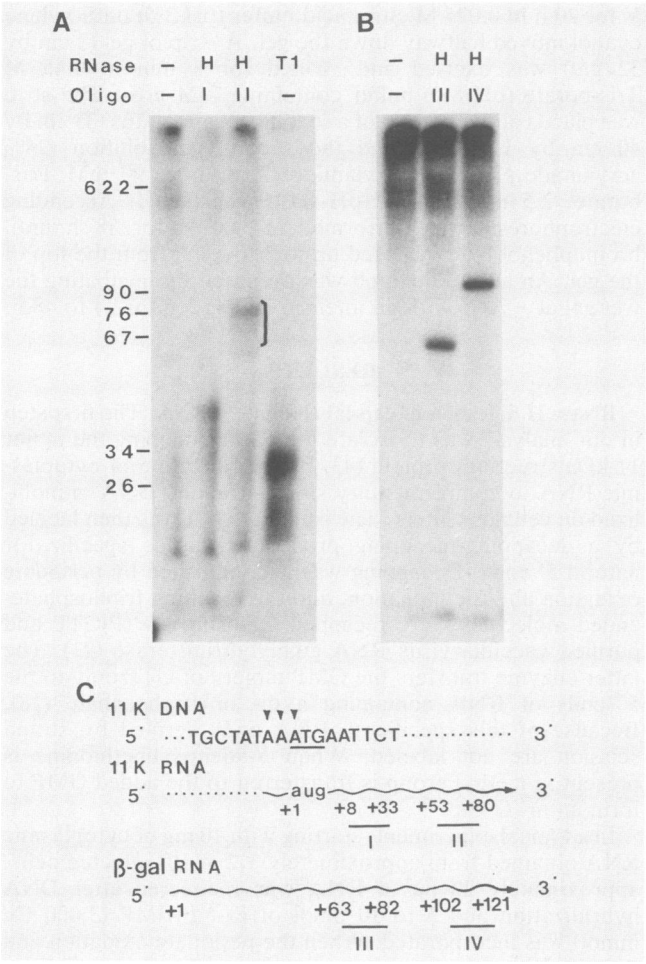

FIG. 1. Polyacrylamide gel electrophoretic analysis of RNase H and RNase T_1 digestion products of vaccinia virus cap-labeled late mRNA. (A) Cytoplasmic RNA from vaccinia virus-infected cells was hybridized to immobilized single-stranded DNA complementary to the mRNA encoding the 11-kDa protein. The selected RNA was cap labeled and then digested with either RNase H or RNase T_1. Oligodeoxyribonucleotides I and II used in RNase H digestion are shown in panel C. The digestion products were resolved by 4% polyacrylamide–7 M urea gel electrophoresis, and an autoradiograph was prepared. Numbers on the left refer to the lengths in nucleotides of DNA markers. (B) RNA encoding β-galactosidase was synthesized in vitro from 1 μg of linear DNA template containing the *lacZ* gene and 20-base-pair bacteriophage T7 promoter by using bacteriophage T7 RNA polymerase in the presence of [α-^{32}P]UTP. The RNA was hybridized to oligodeoxyribonucleotide III or IV (shown in panel C) and digested with RNase H. The digestion products were resolved by polyacrylamide gel electrophoresis, and an autoradiograph was prepared. (C) The DNA sequence containing the late conserved sequence of TAAAT and translation initiation codon (underlined) of the 11-kDa protein gene is shown. The S1 nuclease-resistant 5′ ends of the transcript are shown (▼). Below the DNA is a representation of the transcript made in vivo; the +1 position starts the open reading frame; complementary oligodeoxyribonucleotides I and II are indicated by bars. At the bottom, RNA encoding β-galactosidase synthesized in vitro by T7 RNA polymerase is represented; the +1 position refers to the first nucleotide of the RNA; complementary oligodeoxyribonucleotides III and IV are indicated by bars.

bacteriophage T7 RNA polymerase. Hybridization was carried out with two different oligodeoxyribonucleotides, and in each case RNase H digestion gave a sharp band of the predicted length. It therefore seems unlikely that either the additional length or the size heterogeneity of the vaccinia virus transcripts was an artifact of the digestion procedure.

RNase T_1 digestion provided another way of estimating the lengths of the cap-labeled 5′ ends of the RNAs. Since RNase T_1 cuts specifically 3′ to G residues, initiation within the TAAATG sequences would yield a capped 4-mer or 5-mer. Instead, a smear was obtained with increased intensity at a length of about 34 nucleotides (Fig. 1).

Sequences of the oligoribonucleotides produced by RNase H digestion. We interpreted the results of the RNase H and RNase T_1 cleavage experiments to indicate that up to 30 or 40 additional nucleotides are interposed between the labeled cap and the start of the RNA predicted by nuclease S1 analysis. To further determine their structure, the 5′-terminal fragments generated by RNase H cleavage were electroeluted from the polyacrylamide gel (Fig. 1A, lane 3, bracket). Primer extension and dideoxynucleotide sequencing revealed a poly(A) track starting immediately above the AUGAA and extending for about 30 nucleotides (Fig. 2A). The intensity of the primer extension track suggested that molecules with about 30 A residues were most abundant. Similar, although technically better, results were obtained by primer extension and sequencing of intact mRNA (Fig. 2B).

Analysis of 5′-terminal cap structure. The 5′-terminal cap structures of total late RNA or late RNA selected by hybridization to DNA encoding the 11-kDa protein were isolated by nuclease P1 digestion and DEAE-cellulose thin-layer electrophoresis. The majority of the material migrated with the m^7GpppAm marker, and less than 10% migrated in the position of m^7GpppGm (Fig. 3). This result is consistent with a previous analysis of cap structures of total late RNA labeled in vivo with ^{32}P (5).

The in vitro-labeled caps migrated faster than m^7GpppA and coincident with m^7GpppAm. Since the purified capping enzyme used for labeling lacks the 2′-*O*-methyltransferase, the 2′-*O*-methyladenosine residue must have been formed in vivo and was part of the original cap structure. Control experiments demonstrated that m^7GpppA (not m^7GpppAm) is produced when in vitro-synthesized 5′ triphosphate poly(A) is capped and methylated by purified capping enzyme. Snake venom phosphodiesterase digestion, which liberates m^7Gp, was used to demonstrate the completeness of methylation in the presence of *S*-adenosylmethionine.

Analysis of cap-labeled RNase T_1 oligoribonucleotides. The combination of cap analysis, RNase H digestion, primer extension, and dideoxynucleotide sequencing indicated that the mRNA encoding the 11-kDa protein has the structure m^7G(5′)pppAmp(Ap)$_n$UpGpApA · · ·, in which *n* is quite variable. To assess the variability more directly, the hybridization-selected cap-labeled RNA was digested with RNase T_1, which cleaves 3′ to G residues, and the resultant oligoribonucleotides were analyzed by two-dimensional polyacrylamide gel electrophoresis. Autoradiographs revealed a distinctive pattern, suggesting the presence of families of oligoribonucleotides with similar base composition but different lengths (Fig. 4). Previously, Whitkop et al. (42) obtained a very complex two-dimensional pattern of 5′-terminal oligonucleotides upon RNase T_1 digestion of cap-labeled total cytoplasmic RNA isolated at 6 h after infection; they attributed this finding to the presence of a mixture of early and late transcripts. In our work, we used specific

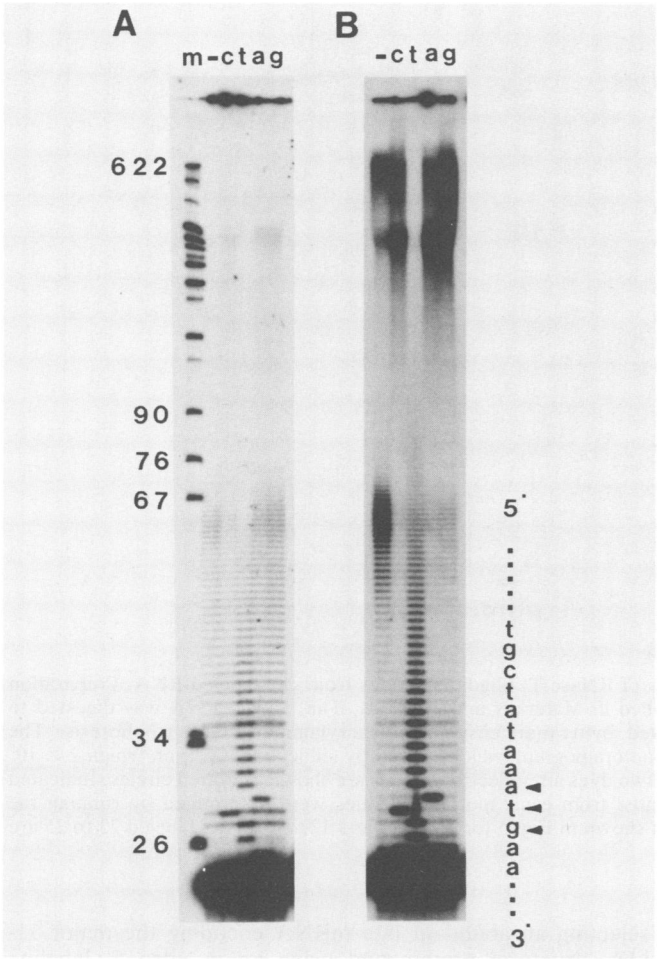

FIG. 2. Sequence of 5′-terminal RNA segments produced by RNase H digestion. (A) Oligodeoxyribonucleotide I (Fig. 1) was 5′ end labeled and used to prime reverse transcription, with RNA encoding the 11-kDa protein as a template. The latter was derived by hybridization of oligodeoxyribonucleotide II to the full-length RNA, RNase H digestion, and polyacrylamide gel electrophoresis. The diffuse band indicated by the bracket in Fig. 1A was electroeluted. Primer extension was carried out in the absence (lane −) or presence of the indicated dideoxyribonucleotide (C, T, A, or G). Numbers on the left refer to the length in nucleotides of DNA markers. (B) The same oligodeoxyribonucleotide used in panel A was used to prime reverse transcription of an intact RNA template. Lanes are marked as in panel A. Shown on the right is the DNA sequence of part of the gene encoding the 11-kDa protein. The translation initiation codon is indicated by arrowheads.

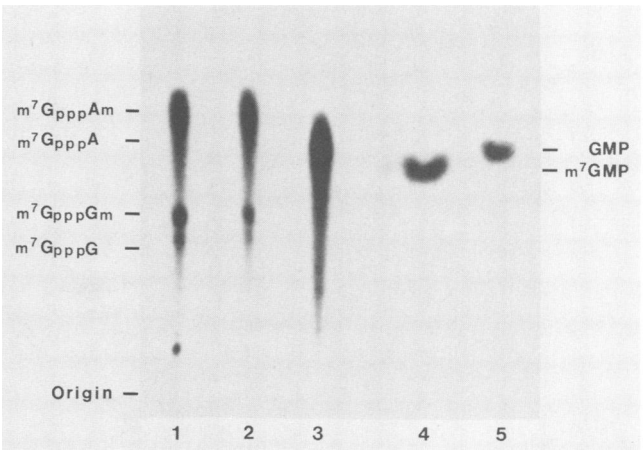

FIG. 3. Analysis of in vitro-labeled cap structures. RNA samples were digested with nuclease P1 (lanes 1, 2, and 3) or snake venom phosphodiesterase (lanes 4 and 5). The digestion products were separated by DEAE-cellulose thin-layer electrophoresis. The electrophoresis took place in 7% formic acid at 90 W for 2 h, and an autoradiograph was prepared. Unless indicated, capping was carried out with [α-^{32}P]GTP and unlabeled *S*-adenosylmethionine. The marker nucleotides were mixed with samples before electrophoresis and then visualized by UV illumination. Lanes: 1, cap-labeled total late RNA; 2 and 4, cap-labeled late RNA selected by hybridization to single-stranded DNA encoding the 11-kDa protein; 3, cap-labeled 5′-triphosphate poly(A) RNA (the latter was synthesized from denatured calf thymus DNA in vitro with *Escherichia coli* RNA polymerase in the presence of 2 mM Mn^{2+} as described previously [36]); 5, same as lane 4, except cap labeled in the absence of *S*-adenosylmethionine.

RNA that hybridized to the complementary strand of the late gene encoding the 11-kDa protein. Analysis of selected oligoribonucleotides of the major track (Fig. 4, spots 13, 14, and 15) by nuclease P1 digestion and DEAE-cellulose thin-layer electrophoresis indicated the presence of the m^7GpppAm cap, whereas oligoribonucleotides from the right adjacent track (spots 23, 24, and 25) had m^7GpppGm caps (Fig. 5A). Other minor tracks have not been analyzed but might be due to the presence of additionally methylated second nucleotides, as was reported previously (5).

Several individual RNase T_1 oligonucleotides of the predominant family were eluted from the polyacrylamide gel and subjected to partial digestion with RNase U2, which cuts specifically 3′ to adenosine residues. A ladder was produced that is consistent with a poly(A) structure (Fig. 5B). Since the initial digestion, prior to two-dimensional polyacrylamide gel electrophoresis, was with RNase T_1, the 3′-terminal nucleotide must be a G. Nucleotide sequencing of the 5′-terminal RNA segments produced by RNase H digestion had indicated that the next nucleotide is a U residue. In accord with this, a relative gap is present in the RNase U2 ladder just below the full-length RNase T_1 products (Fig. 5B).

From the number of partial U2 digestion products obtained, as well as their migration rates relative to oligoribonucleotides of known length (Fig. 5B), we deduced that spots 11 and 12 have seven and eight consecutive A residues between the cap and the UG sequence, respectively. By extrapolation, the spots below 11 and 12 in Fig. 4 have 9 to at least 30 A residues. The m^7G(5′)pppGm capped oligoribonucleotides in the parallel track appear to be of a similar size range.

DISCUSSION

A limitation of previous analyses of the 5′-terminal structure of vaccinia virus late mRNA was the primary reliance on indirect methods including nuclease S1 analysis of DNA-RNA hybrids and reverse transcriptase extension of oligodeoxyribonucleotide primers. Although these methods have proven reliable for conventional RNA structures, alternative and more direct analyses are needed for us to be confident of novel structures. The major technical difficulty with direct analysis is the relatively small amount of specific mRNA obtainable from infected cells. A second difficulty, which holds true for both direct and indirect methods, is that of distinguishing natural 5′ ends from degradation products. We have attempted to solve these problems by hybrid-

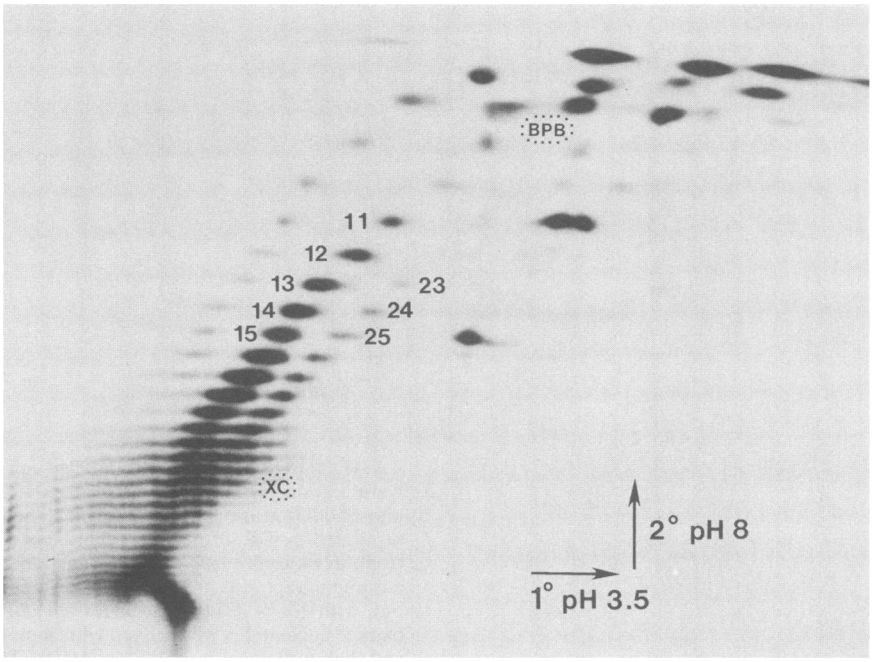

FIG. 4. Two-dimensional polyacrylamide gel electrophoretic analysis of RNase T_1 oligonucleotides from cap-labeled RNA. Preparation (and cap labeling) of RNA encoding the 11-kDa protein was as described in Materials and Methods. The labeled RNA was digested to completion with RNase T_1, and the digestion products were resolved by two-dimensional polyacrylamide gel electrophoresis. The two-dimensional gel system is described in Materials and Methods. An autoradiograph was prepared by subjecting a gel, on which 2×10^5 to 5×10^5 cpm was separated, to a 24- to 48-h exposure. The positions of two dyes after electrophoresis are shown as dotted circles. Indicated as numbers from 11 to 15 are the lengths of oligoribonucleotides isolated from each spot. The values were determined by running the oligonucleotides on a separate gel next to an RNA sequencing ladder as shown in Fig. 5 for spots 11 and 12. The spots labeled 23 to 25 are presumed to be 13 to 15 nucleotides in length.

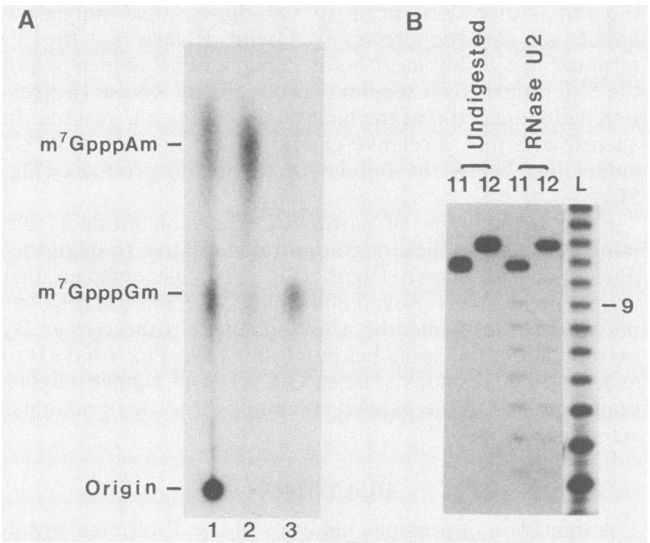

FIG. 5. Analysis of isolated RNase T_1 oligoribonucleotides. The indicated oligoribonucleotides were electroeluted from the gel shown in Fig. 4. After two precipitations with 2.5 M ammonium acetate and ethanol, oligoribonucleotides were subjected to further analysis. (A) Cap structure analysis. Oligoribonucleotides isolated from spots 13, 14, and 15 were combined and digested with nuclease P1. The digestion products were separated by DEAE-cellulose thin-layer electrophoresis as described in the legend to Fig. 3, and an autoradiograph was made (lane 2). Likewise, oligoribonucleotides were isolated from spots 23, 24, and 25 and were subjected to the same analysis (lane 3). As a control, cap-labeled RNA encoding the 11-kDa protein was digested with nuclease P1 and separated under the same conditions (lane 1). Indicated at the left are dinucleotide cap markers. (B) Partial RNase U2 digestion. Oligoribonucleotides isolated from spots 11 and 12 were subjected to a partial digestion with 0.1 U of RNase U2 in the presence of 1 μg of yeast tRNA. The digestion products were separated on a 12% polyacrylamide gel containing 7 M urea. An autoradiograph was prepared by exposure of a dry gel at −70°C for 5 days with an intensifying screen. As a control, cap-labeled 5′-triphosphate poly(A) (made with *E. coli* RNA polymerase) was digested with RNase U2 under the same conditions (lane L). Indicated on the right is a molecular size marker (9 nucleotides).

selecting an abundant late mRNA encoding the major 11-kDa structural protein and using an in vitro 5′ labeling procedure which relies on the specificity of the vaccinia virus capping enzyme for di- or triphosphate-ended molecules. Since we knew from earlier studies (5) that vaccinia virus late mRNAs are capped (estimate of one cap per 1,200 nucleotides), our first step was the removal of the m^7G. The procedure (periodate oxidation and β-elimination) takes advantage of the *cis*-diol structure of the cap. Vaccinia virus capping enzyme in the presence of [α-^{32}P]GTP and *S*-adenosylmethionine added an m^7Gp with a labeled phosphate residue. We know that we were relabeling cap structures and not preexisting di- or triphosphate-ended RNAs for two reasons. First, if periodate oxidation and β-elimination were omitted, the labeling decreased by more than 99%. Second, analysis of the labeled cap structures indicated that $m^7GpppAm$ was formed. Since the purified capping enzyme had no detectable 2′-*O*-methyltransferase, the adenosine must have been methylated in vivo. Furthermore, since the vaccinia virus 2′-*O*-methyltransferase requires a cap struc-

ture for activity (1), it follows that the labeled ends were $m^7G(5')pppAm$ originally.

Next, we wished to determine whether the labeled cap was located on leaders attached to the mRNA encoding the 11-kDa protein and, if so, to determine the length and sequence of the leader. To accomplish this, we needed a way of cutting at a precise site within the mRNA. Our approach was to make RNase H-susceptible sites by hybridizing complementary oligodeoxyribonucleotides to the mRNA. Polyacrylamide gel electrophoresis revealed that leaders of heterogeneous length up to 30 to 40 nucleotides were present. Control experiments, in which RNAs with defined 5′ ends were used, ensured that neither the length estimates nor the heterogeneity was due to artifacts of RNase H digestion. Furthermore, RNase T_1 digestion of the cap-labeled RNA gave a family of oligonucleotides of the type $m^7G(5')pppAmp(Ap)_nUpGp$, in which n varies from a few to more than 30 nucleotides.

The present results are consistent with previous data indicating that late mRNAs contain caps with predominantly two A nucleotides following the m^7G (5) and that cap-specific antibody binds to the poly(A) leaders (33). Our analysis suggests considerable variability in the length of the capped poly(A) leaders, which may be of importance in determining the mechanism of their formation. Schwer and Stunnenberg (32) reported that the length of the poly(A) leader made in vitro depended on the concentration of ATP. Although a mechanism involving reiterative transcription of T residues within the conserved ATTTAC sequence in the template strand is favored, critical experiments to rule out a mechanism(s) involving capped primers or ligation have not been carried out.

We wish to emphasize that our data do not rule out the existence of additional forms of late mRNA consisting of a poly(A) segment interposed between two unrelated RNAs (4). Indeed, we did see very long molecules, whether analysis was carried out by RNase H digestion or primer extension by reverse transcriptase. Although most of the experiments presented here pertained to RNA selected by hybridization to the 11-kDa protein gene, essentially similar results were obtained with total vaccinia virus late RNA, indicating that the capped poly(A) leader structure must be a function of many late RNAs.

ACKNOWLEDGMENTS

We thank Nicholas Stamatos for carrying out preliminary RNase H digestion experiments, Stuart Shuman for providing samples of purified vaccinia virus capping enzyme, and Norman Cooper for performing cell cultures.

LITERATURE CITED

1. **Barbosa, E., and B. Moss.** 1978. mRNA(nucleoside-2′-)-methyltransferase from vaccinia virus. Purification and physical properties. J. Biol. Chem. **253**:7692–7697.
2. **Bertholet, C., R. Drillien, and R. Wittek.** 1985. One hundred base pairs of 5′ flanking sequence of a vaccinia virus late gene are sufficient to temporally regulate late transcription. Proc. Natl. Acad. Sci. USA **82**:2096–2100.
3. **Bertholet, C., P. Stocco, E. Van Meir, and R. Wittek.** 1986. Functional analysis of the 5′ flanking sequence of a vaccinia virus late gene. EMBO J. **5**:1951–1957.
4. **Bertholet, C., E. Van Meir, B. Haggeler-Bordier, and R. Wittek.** 1987. Vaccinia virus produces late mRNAs by discontinuous synthesis. Cell **50**:153–162.
5. **Boone, R. F., and B. Moss.** 1977. Methylated 5′-terminal sequences of vaccinia virus mRNA species made in vivo at early and late times after infection. Virology **79**:67–80.
6. **Boone, R. F., and B. Moss.** 1978. Sequence complexity and relative abundance of vaccinia virus mRNAs synthesized in vivo and in vitro. J. Virol. **26**:554–569.
7. **Boone, R. F., R. P. Parr, and B. Moss.** 1979. Intermolecular duplexes formed from polyadenylated vaccinia virus RNA. J. Virol. **30**:365–374.
8. **Cochran, M. A., C. Puckett, and B. Moss.** 1985. In vitro mutagenesis of the promoter region for a vaccinia virus gene: evidence for tandem early and late regulatory signals. J. Virol. **54**:30–37.
9. **Colby, C., C. Jurale, and J. R. Kates.** 1971. Mechanism of synthesis of vaccinia virus double-stranded ribonucleic acid in vivo and in vitro. J. Virol. **7**:71–76.
10. **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.
11. **Cooper, J. A., R. Wittek, and B. Moss.** 1981. Extension of the transcriptional and translational map of the left end of the vaccinia virus genome to 21 kilobase pairs. J. Virol. **39**:733–745.
12. **deWachter, R., and W. Fiers.** 1972. Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. Anal. Biochem. **49**:184–197.
13. **Ensinger, M. J., S. A. Martin, E. Paoletti, and B. Moss.** 1975. Modification of the 5′-terminus of mRNA by soluble guanylyl and methyltransferases from vaccinia virus. Proc. Natl. Acad. Sci. USA **72**:2525–2529.
14. **Hanggi, M., W. Bannwarth, and H. G. Stunnenburg.** 1986. Conserved TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. EMBO J. **5**:1071–1076.
15. **Kates, J., and J. Beeson.** 1970. Ribonucleic acid synthesis in vaccinia virus. II. Synthesis of polyriboadenylic acid. J. Mol. Biol. **50**:19–23.
16. **Kaverin, N. V., N. L. Varich, V. V. Surgay, and V. I. Chernos.** 1975. A quantitative estimation of poxvirus genome fraction transcribed as early and late mRNA. Virology **65**:112–119.
17. **Mahr, A., and B. E. Roberts.** 1984. Arrangement of late mRNAs transcribed from a 7.1-kilobase *Eco*RI vaccinia virus DNA fragment. J. Virol. **49**:510–520.
18. **Martin, S. A., and B. Moss.** 1975. Modification of RNA by mRNA guanylyltransferase and mRNA (guanine-7-)methyltransferase form vaccinia virions. J. Biol. Chem. **250**:9330–9335.
19. **Martin, S. A., E. Paoletti, and B. Moss.** 1975. Purification of mRNA guanylyltransferase and mRNA (guanine.7) methyltransferase from vaccinia virions. J. Biol. Chem. **250**:9322–9329.
20. **Moss, B.** 1977. Utilization of the guanylyltransferase and methyltransferases of vaccinia virus to modify and identify the 5′-terminals of heterologous RNA species. Biochem. Biophys. Res. Commun. **74**:374–383.
21. **Moss, B.** 1981. 5′-End labeling of RNA with capping and methylating enzymes. Gene amplification and analysis of nucleic acids, p. 254–266. *In* J. G. Chirikjian and T. S. Papas (ed.), Structural analysis of nucleic acids, vol. 2. Elsevier/North Holland Publishing Co., New York.
22. **Moss, B.** 1985. Replication of poxviruses, p. 685–703. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, J. Melnick, B. Roizman, and R. Shope (ed.), Virology. Raven Press, New York.
23. **Moss, B., and F. Koczot.** 1976. Sequence of methylated nucleotides at the 5′ terminus of adenovirus-specific RNA. J. Virol. **17**:385–392.
24. **Moss, B., and N. P. Salzman.** 1968. Sequential protein synthesis following vaccinia virus infection. J. Virol. **2**:1016–1027.
25. **Nevins, J. R., and W. K. Joklik.** 1975. Poly(A) sequences of vaccinia virus messenger RNA: nature, mode of addition and function during translation in vitro and in vivo. Virology **63**:1–14.
26. **Oda, K.-I., and W. R. Joklik.** 1967. Hybridization and sedimentation studies on "early" and "late" vaccinia virus messenger RNA. J. Mol. Biol. **27**:395–419.
27. **Paoletti, E., and L. Grady.** 1977. Transcriptional complexity of vaccinia virus in vivo and in vitro. J. Virol. **23**:608–615.
28. **Pennington, T. H.** 1974. Vaccinia virus polypeptide synthesis: sequential appearance and stability of pre- and post-replicative

polypeptides. J. Gen. Virol. **25**:433–444.
29. **Puckett, C., and B. Moss.** 1983. Selective transcription of vaccinia virus genes in template dependent soluble extracts of infected cells. Cell **35**:441–448.
30. **Rohrmann, G., L. Yuen, and B. Moss.** 1986. Transcription of vaccinia virus early genes by enzymes isolated from vaccinia virions terminates downstream of a regulatory sequence. Cell **46**:1029–1035.
31. **Rosel, J., P. 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 *Hin*dIII H genome fragment. J. Virol. **60**:436–449.
32. **Schwer, B., and H. G. Stunnenberg.** 1988. Vaccinia virus late transcripts generated in vitro have a poly(A) head. EMBO J. **7**: 1183–1190.
33. **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.
34. **Shuman, S., S. S. Broyles, and B. Moss.** 1987. Purification and characterization of a transcription termination factor from vaccinia virions. J. Biol. Chem. **262**:12372–12380.
35. **Shuman, S., and B. Moss.** 1988. Factor-dependent transcription termination by vaccinia virus RNA polymerase. J. Biol. Chem. **263**:6220–6225.
36. **Venkatesan, S., and B. Moss.** 1980. Donor and acceptor specificities of HeLa cell mRNA guanyltransferase. J. Biol. Chem. **255**:2835–2842.
37. **Venkatesan, S., and B. Moss.** 1981. In vitro transcription of the inverted terminal repetition of the vaccinia virus genome: correspondence of initiation and cap sites. J. Virol. **37**:738–747.
38. **Wei, C. M., and B. Moss.** 1975. Methylated nucleotides block 5′-terminus of vaccinia virus messenger RNA. Proc. Natl. Acad. Sci. USA **72**:318–322.
39. **Weir, J. P., and B. Moss.** 1984. Regulation of expression and nucleotide sequence of a late vaccinia virus gene. J. Virol. **51**: 662–669.
40. **Weir, J. P., and B. Moss.** 1987. Determination of the transcriptional regulatory region of a vaccinia virus late gene. J. Virol. **61**:75–80.
41. **Weir, J. P., and B. Moss.** 1987. Determination of the promoter region of an early vaccinia virus gene encoding thymidine kinase. Virology **158**:206–210.
42. **Whitkop, C., B. R. Lipinskas, S. Mercer, D. Panicali, and E. Paoletti.** 1982. Analysis of vaccinia virus transcriptional complexity in vitro and in vivo: characterization of RNase T_1-resistant 5′-terminal oligonucleotides. J. Virol. **42**:734–741.
43. **Wittek, R., M. Hanggi, and G. Hiller.** 1984. Mapping of a gene coding for a major late structural protein on the vaccinia virus genome. J. Virol. **49**:371–378.
44. **Wright, C. F., and B. Moss.** 1987. In vitro synthesis of vaccinia virus late mRNA containing a 5′ poly(A) leader sequence. Proc. Natl. Acad. Sci. USA. **84**:8883–8887.
45. **Yuen, L., and B. Moss.** 1987. Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. Proc. Natl. Acad. Sci. USA **84**:6417–6421.