Transcriptional and Translational Analysis of a Strongly Expressed Early Region of the Vaccinia Virus Genome

FRED GOLINI* AND JOSEPH R. KATES

Department of Basic and Clinical Research, Division of Cellular Biology, Scripps Clinic and Research Foundation, La Jolla, California 93027

Received 8 July 1983/Accepted 19 October 1983

A transcriptional and translational map was obtained for a 7-kilobase pair *Eco*RI fragment of vaccinia virus DNA containing 4% of the viral genome. This particular region contains a cluster of early genes which are transcribed in viral cores in vitro and in infected cells before uncoating of the viral DNA. Expression of this region was characterized by using vaccinia virus DNA sequences cloned in phage and plasmid vectors. Polypeptides encoded on the 7-kilobase pair fragment were identified by cell-free translation of viral mRNA selected by hybridization to immobilized DNA fragments. Early RNA programed the synthesis of six proteins having apparent molecular weights of 63,000, 38,000, 37,000, 25,000, 15,000, and 13,000. The same result was obtained with RNA synthesized in vitro. A new species of 40 kilodaltons was synthesized in response to late RNA. Of the six "early" polypeptides, only the 13-kilodalton component was synthesized from late mRNA. RNA derived from the 7-kilobase pair region was analyzed by a variety of methods including hybridization blot, in vitro recapping, and S1 nuclease techniques. A surprisingly complex pattern of early transcription was found, indicating the existence of families of overlapping RNA species which share common 5'-proximal sequences. In addition, larger RNAs were identified spanning two or more adjacent genes. These RNAs appear to possess a common 5' terminus with transcripts derived from the first gene and are coterminal at the 3' end with RNAs from the downstream gene. Late RNA encoding the 40-kilodalton protein was shown to be heterogeneous in size. A single 5' terminus but no unique 3' terminus was identified for this class of transcripts. RNA species synthesized by cores in vitro were of similar size to authentic transcripts isolated from infected cells at early times.

The poxviruses, as a consequence of the cytoplasmic location of their replication cycle, provide a unique model system for the study of gene expression which is independent of host cell nuclear function. Transcription of early genes occurs within subviral core particles utilizing a transcription apparatus contained within the virion and activated upon penetration into the cell, resulting in the extrusion of capped, polyadenylated mRNAs. In view of the ability of permeabilized vaccinia virus cores to produce RNA in vitro with characteristics similar to early RNA from infected cells, core particles provide a suitable starting point for analysis of the components required for accurate synthesis and post-transcriptional modification of early RNA.

Progress in understanding the details of mRNA synthesis in vaccinia virus has been hindered by the relatively large size, 180 kilobase pairs (kbp), and complexity of the poxvirus genome. In recent studies, Moss and colleagues have utilized recombinant DNA methodologies to obtain a preliminary transcriptional map of several cloned DNA restriction fragments comprising >10% of the vaccinia virus genome (1, 7, 8, 22, 23). Three of these fragments were derived from the left end of the genome, a region containing one of the inverted terminal repetitions and a proximal "nonessential" region which is deleted from the genome of variant strains. mRNAs encoding early and late proteins were identified. Although these RNAs were capped and polyadenylated as are cellular RNAs which must function in the cytoplasmic milieu, they differed from most nuclear transcripts in that they were not spliced; the initiating nucleotide of the primary transcript was retained at the 5' end of mature RNAs contiguous to the body of the coding sequence (20).

In this communication we have focused on the characterization of transcripts and polypeptides whose synthesis is directed by a 7-kbp *Eco*RI fragment of vaccinia virus DNA which showed particularly strong hybridization with "immediate early" RNA obtained from infected cells or transcribed by permeabilized viral cores. It was presumed that this DNA fragment encodes a number of relatively abundant early mRNAs and is a likely candidate to contain one or more strong early promoters. This region is therefore suitable for our longer-term goal of studying viral gene transcription in a soluble in vitro system derived from purified vaccinia virions. It was also anticipated that the availability of higher concentrations of RNA derived from this fragment would facilitate construction of a more detailed transcription map.

MATERIALS AND METHODS

Radioisotopes and reagents. Media and serum for cell culture were purchased from Irvine Scientific. Restriction enzymes, terminal transferase, and T4 polynucleotide kinase were obtained from Bethesda Research Laboratories. [³⁵S]methionine and [γ -³²P]ATP were supplied by Amer-sham Corp. [α -³²P]GTP and [α -³²P]dCTP came from New England Nuclear.

Vaccinia virus DNA. Vaccinia virus WR strain was propagated in HeLa cells maintained in suspension in Eagle minimum essential medium for spinner cells essentially as described previously (11), except that crude stocks of virus were produced by low-multiplicity passage in HeLa monolayers rather than L cells. The method of virus purification was also modified by substitution of equilibrium density banding in 20 to 50% (wt/vol) potassium tartrate gradients containing 10 mM Tris-hydrochloride (pH 8.0) in place of the sucrose gradient centrifugation used previously. DNA was

* Corresponding author.

prepared from purified vaccinia virus suspended at 2×10^{11} particles per ml in 10 mM Tris-hydrochloride (pH 8.0)–100 mM NaCl–5 mM EDTA–1% β -mercaptoethanol–1% sodium dodecyl sulfate (SDS)–25% sucrose. Proteinase K (Boehringer Mannheim) was added to 250 μ g/ml, and the suspension was incubated for 1 h at 37°C before successive extractions with equal volumes of buffer-saturated phenol-chloroform (1:1) (twice) and diethyl ether (four times). DNA was precipitated from the aqueous phase with ethanol and dissolved in 10 mM Tris-hydrochloride (pH 8.0)–1 mM EDTA.

Molecular cloning. An *Eco*RI limit digest of vaccinia virus DNA was ligated to alkaline phosphatase-treated pBR322 DNA linearized at the *Eco*RI site, and the resulting chimeric plasmids were used to transform competent *Escherichia coli* (5). Ampicillin- and tetracycline-resistant colonies were grown on nitrocellulose membranes and screened with a hybridization probe of 32 P-labeled vaccinia virus core RNA (10). One recombinant plasmid containing a 7-kbp vaccinia virus insert, pV7, was further characterized. Hybridization of nick-translated pV7 DNA to a blot of *Hind*III-restricted vaccinia virus DNA identified the genomic location as the F fragment. A DNA restriction map was constructed by a combination of methods, including multiple enzyme digests, secondary digestion of purified restriction fragments, and partial digestion of end-labeled molecules. To facilitate transcriptional mapping studies, the *Hpa*II subfragments of the 7-kbp sequence were recovered from agarose gels by electroelution (18) and cloned in *Clal*- or *Clal*/*Eco*RI-treated pBR322. Ampicillin-resistant, tetracycline-sensitive recombinants were selected by plating transformants on fusaric acid-containing medium (4).

A genomic library of vaccinia virus sequences in the phage vector λ Charon 4 was constructed by the method of Maniatis et al. (12). Phage DNA arms were ligated to 10 to 20-kbp vaccinia virus DNA fragments produced by partial *Eco*RI digestion and purified by sucrose gradient fractionation. After in vitro packaging recombinant molecules were amplified in *E. coli* K802. The separated strands of recombinant clone B26 were obtained by CsCl density centrifugation of denatured DNA in the presence of polyguanylic-uridylic acid (Sigma Chemical Co.) according to published procedures (9).

Vaccinia virus RNA. HeLa cells grown in spinner culture were collected by centrifugation and resuspended at 10^7 cells per ml in warm adsorption medium (1 part Puck saline, 1 part Eagle minimum essential medium for spinner cells containing 5% heat-inactivated horse serum). Purified vaccinia virus was added at a multiplicity of 1,000 particles per cell, and the suspension was stirred slowly for 30 min at 37°C before dilution with prewarmed Eagle minimum essential medium for spinner cells plus 5% heated horse serum to 10^6 cells per ml. Immediate early or early RNA was prepared from cells infected for 4 h in the presence of cycloheximide (CHI; 100 μ g/ml) or cytosine arabinoside (CAR; 40 μ g/ml). Late RNA was prepared from cells harvested after 6 h in the absence of drug treatment. Infected cells were washed once with ice-cold phosphate-buffered saline and suspended at 5×10^7 cells per ml in hypotonic buffer: 20 mM Tris-hydrochloride (pH 7.6), 3 mM $MgCl_2$, 10 mM NaCl, and 10 mM vanadyl nucleotides (Bethesda Research Laboratories). To 1 volume of cells one-third volume of lysing buffer (hypotonic buffer containing 5% [wt/vol] sucrose and 1.2% Triton X-100) was added; after gentle mixing cells were held at 4°C for 15 min on ice. Nuclei and cell debris were removed by centrifugation at $15,000 \times g$ for 10 min. EDTA and SDS were added to

the supernatant to 10 mM and 1%, respectively, followed by extraction with phenol-chloroform (1:1) at least three times or until no additional material appeared at the interphase between organic and aqueous layers. RNA was precipitated from the aqueous phase by the addition of NaOAc (pH 5.5) to 0.3 M and 2 volumes of ethanol. The polyadenylated fraction was isolated by chromatography of total RNA on oligodeoxythymidylate-cellulose (P-L Biochemicals).

RNA was synthesized in vitro by purified viral core particles as follows: 2×10^{11} virus particles were incubated at 37°C for 30 min in 1 ml of EP buffer (50 mM Tris-hydrochloride, pH 8.5, 5 mM $MgCl_2$, 10 mM dithiothreitol, 0.5% Triton X-100). This suspension was layered over a 5-ml cushion of 36% (wt/vol) sucrose in EP buffer and spun for 60 min at 12,000 rpm in a Sorvall HB4 swinging-bucket rotor ($23,500 \times g$) at 4°C. The pellet of washed viral cores was suspended in 4 ml of EP buffer supplemented with 2.5 mM ATP, 1.25 mM each CTP, UTP, and GTP, and 10 μ M S-adenosylmethionine and incubated at 37°C. Radiolabeled RNA was synthesized in similar reactions containing 50 μ M [α - 32 P]GTP (typically 10^3 to 10^4 cpm/pmol). After 30 min, incubations were chilled and viral cores were removed by centrifugation ($23,000 \times g$ for 15 min). From the supernatant fraction extruded RNA was extracted with phenol-chloroform and chromatographed on oligodeoxythymidylate-cellulose as described above.

RNA hybridization to immobilized DNA filters. Circles 6 mm in diameter were punched from a sheet of aminophenylthioether paper prepared by the procedure of Seed and diazotized as described (17). A solution containing 10 μ g of sonicated denatured DNA in 20 mM NaOAc (pH 4.0) was applied, and coupling was allowed to proceed overnight at 4°C. Unbound DNA was removed by washing filters with 0.4 M NaOH at 37°C. Approximately 200 μ g of polyadenylated RNA was hybridized to each filter in 200 μ l of 50% formamide–0.4 M NaCl–0.04 M sodium PIPES {piperazine-N,N'-bis(2-ethanesulfonic acid)}, pH 6.4–0.005 M EDTA–0.2% SDS–1 mg of yeast tRNA (Boehringer Mannheim) per ml–50 μ g of polyadenylic acid (Sigma) per ml at 37°C for 6 to 8 h. Nonhybridized RNA was removed by four successive washes of 5-min duration in 50% formamide–0.1 M NaCl–0.01 M sodium PIPES (pH 6.4)–0.005 M EDTA–0.1% SDS at 25°C followed by two 10-min washes in 0.2 M NaCl–0.02 M sodium PIPES (pH 6.4)–0.005 M EDTA–0.1% SDS at 55°C. Specifically retained RNA was eluted by incubation of filters in 0.2 ml of 99% formamide containing 1 mM EDTA at 55°C for 5 min. RNA was precipitated with ethanol after the addition of 5 μ g of yeast tRNA as carrier and sodium acetate (pH 5.5) to 0.3 M.

Cell-free protein synthesis. Micrococcal nuclease-treated rabbit reticulocyte lysates were prepared by the method of Pelham and Jackson (16). Polyadenylated RNA (0.1 μ g) was translated in 20- μ l incubations containing 20 μ Ci of [35 S]methionine for 1 h at 30°C. Samples (2 μ l) were analyzed by electrophoresis in 12% polyacrylamide gels containing SDS. The gels were subjected to fluorography as described by Bonner and Lasky (6). 14 C-labeled protein molecular-weight standards were purchased from Bethesda Research Laboratories.

Electrophoretic analysis of nucleic acids. DNA or RNA was denatured as described by McMaster and Carmichael (14) by incubation in 20- μ l reaction mixtures containing 50% dimethyl sulfoxide, 1 M glyoxal, and 10 mM sodium phosphate, pH 7.0. After 30 min at 50°C reactions were cooled to room temperature, and 5 μ l of a dye mixture containing 0.4% bromophenol blue and 0.4% xylene cyanol in dimethyl

sulfoxide was added to samples. Electrophoretic separation was accomplished in 1.5 or 2% horizontal agarose gels in 10 mM sodium phosphate (pH 7.0) with constant recirculation of the electrode buffer to maintain correct pH. Gels of radiolabeled RNAs were dried and subjected to direct autoradiography. Unlabeled RNAs were detected after transfer to nitrocellulose membranes by the method of Thomas (19). Blots were prehybridized for at least 2 h at 41°C in sealed bags in 50% formamide–5× SSPE, (1× SSPE = 0.18 M NaCl, 10 mM sodium phosphate, pH 7.0, 1 mM EDTA)–0.1% SDS–0.2 mg of denatured sonicated calf thymus DNA per ml–5× Denhardt solution. Hybridization was conducted in the same buffer containing 10% dextran sulfate, and DNA probes were labeled by nick translation. After 16 h of hybridization at 41°C, blots were washed two times in 2× SSPE–0.1% SDS at room temperature and twice in 0.1× SSPE–0.1% SDS at 37°C for 15 min per wash. Labeled bands were detected by autoradiography at –70°C with intensifying screens.

Nuclease S1 mapping. DNA restriction fragments were terminally labeled in the 5' position with [γ - 32]ATP and T4 polynucleotide kinase and in the 3' position with [α - 32]dCTP and terminal transferase as described by Maxam and Gilbert (13). Polyadenylated RNA was annealed with end-labeled DNA restriction fragments in 20- μ l volumes containing 80% formamide–0.8 M NaCl–0.08 M sodium PIPES (pH 6.4)–0.01 M EDTA at 41°C for 12 h. Hybridizations were terminated by addition of 200 μ l of digestion buffer (0.25 M NaCl, 0.03 M NaOAc, pH 4.5, 0.001 M ZnCl₂, 20 μ g of denatured calf thymus DNA per ml)

containing 5,000 U of S1 nuclease (Boehringer Mannheim) per ml. Digestion of single-stranded tails was allowed to proceed for 1 h at 41°C. Nuclease-resistant heteroduplexes were precipitated with ethanol and analyzed by gel electrophoresis after glyoxal denaturation as described above.

RESULTS

Digestion of vaccinia virus DNA with endonuclease *Eco*RI produces approximately 50 fragments suitable for cloning in plasmid vectors. Recombinant plasmids containing the vaccinia virus terminal DNA fragments were not expected as a consequence of the block to ligation posed by the terminal loop cross-linking the two DNA strands (22). A limit digest of vaccinia virus DNA was cloned in pBR322 and recombinants were screened with [32 P]RNA synthesized by permeabilized viral cores in vitro as a hybridization probe. One particular recombinant plasmid containing a 7-kb vaccinia virus insert, pV7, was selected for further study on the basis of its particularly strong hybridization, reflecting the high level of transcription of this region in viral cores. The 7-kbp *Eco*RI fragment has been located in the genome entirely within the 12.8-kbp *Hind*III-F fragment at approximately 20% of the genome length in from the left end (Fig. 1) by Southern blot analysis.

Expression of genes located on the 7-kbp region in infected cells at early and late times and in permeabilized viral cores was compared by analysis of RNA prepared under a variety of conditions. Infection of cells treated with CHI or CAR has been used to distinguish two classes of early genes. In the presence of CHI, an inhibitor of protein synthesis,

FIG. 1. Genome location and endonuclease restriction maps of 7-kbp *Eco*RI fragment. Upper portion represents a restriction map of the left end of the vaccinia virus genome including sites for *Hind*III (▼), lettered segments, and *Eco*RI (▲). The positions of the 7-kbp *Eco*RI fragment cloned in pV7 and the 10.1-kbp insert of recombinant phage clone B26 are shown. In the lower portion sites for several endonuclease restriction enzymes within the 7-kbp region are given.

infected cells accumulate large amounts of virus-specific RNA resulting from the prolonged transcription of immediate early genes within intracytoplasmic viral core particles. A small class of "delayed early" genes has also been described whose expression requires prior viral protein synthesis. The nucleoside analog CAR prevents the onset of DNA synthesis but allows the expression of all early genes. Polyadenylated RNA was purified from infected cells treated with CHI, CAR, or at late times in the absence of any drugs. After denaturation, RNA was resolved by agarose gel electrophoresis and transferred to a nitrocellulose membrane. Those transcripts derived from the 7-kbp region were identified by virtue of their hybridization to 32 P-labeled pV7 DNA (Fig. 2). A surprising complexity was found with early RNA, including several major bands of 1 to 2 kilobases (kb) and at least three larger species of 3 or 4 kb. No qualitative difference between immediate early (CHI) and early (CAR) classes of RNA was detected (lanes 2 and 3). The same bands were also present in core RNA, but a significant background was seen (lane 1). In contrast, a different result was obtained with late RNA (lane 4). Although many of the

FIG. 2. Electrophoretic resolution of RNAs derived from the 7-kbp region. Unlabeled polyadenylated RNAs were denatured by treatment with glyoxal and resolved by electrophoresis in a 1.5% agarose gel. After transfer to a nitrocellulose membrane, the RNA blot was hybridized to pV7 DNA labeled by nick translation. An autoradiograph is shown. For reference the position and size (in kilobases) of denatured restriction fragments produced by *Eco*RI and *Hind*III digestion of phage λ DNA are indicated. Lane 1, Core RNA; lane 2, immediate early (CHI) RNA; lane 3, early (CAR) RNA; lane 4, late RNA.

FIG. 3. Cell-free translation of immediate early RNAs encoded on the 7-kbp region. CHI RNA was hybridized to filters containing immobilized pV7 DNA. Specifically retained RNA was eluted and translated in a rabbit reticulocyte lysate supplemented with [35 S]methionine. Polypeptide products were analyzed on a 12% SDS-polyacrylamide gel and visualized by fluorography. Apparent molecular weights were assigned by comparison to the migration of radiolabeled reference proteins. In vitro protein synthesis reactions were programmed with no mRNA, left lane; polyadenylated RNA from infected cells, center lane; or pV7-specific RNA, right lane.

same RNAs seen at early times were present, their relative abundance was shifted in favor of the larger bands. In addition, a heterogeneous smear of higher molecular weight was seen, indicating the synthesis of new RNA species at late times.

mRNAs encoded on the 7-kbp region were purified from each RNA preparation by hybridization to pV7 DNA immobilized on aminophenylthioether filters. After a stringent washing procedure, specifically retained transcripts were eluted and assayed by cell-free translation in the mRNA-dependent rabbit reticulocyte lysate. The results obtained with CHI RNA are shown in Fig. 3. In contrast to the large number of polypeptides synthesized in response to total polyadenylated vaccinia virus RNA (lane 2), pV7-specific RNA directed the synthesis of a unique subset of six proteins with apparent molecular weights of 63,000, 38,000, 37,000, 25,000, 15,000, and 13,000 (lane 3). The same six polypeptide products were also obtained from in vitro translation of CAR and core RNAs selected by pV7 DNA filters (data not

shown). In contrast, two polypeptides of 40 and 13 kilodaltons (kd) were obtained by in vitro translation of late RNA derived from the 7-kbp region (lane 2, Fig. 4). The smaller protein comigrated with the 13-kd band present at early times and may reflect the continued expression of this early gene throughout the infection. No other early proteins were detected. A band of 48 kd resulting from the labeling of an endogenous reticulocyte protein can be seen in each lane including the minus RNA control, lane 1. The 40-kd polypeptide is a new species whose synthesis was programmed exclusively by late RNA. These results identified six immediate early and a single late gene located on the 7-kbp *EcoRI* DNA fragment. No genes of the delayed early class were found in this region.

To more precisely locate the six early genes within the 7-kbp region, a series of DNA filters were prepared, each containing one of the seven cloned subfragments obtained by *HpaII* digestion. An analysis of the translation products of CHI RNA selectively hybridizing to each DNA filter is shown in Fig. 5. In some instances RNA derived from adjacent *HpaII* DNA fragments directed the synthesis of polypeptides of identical size. This indicated that the gene in question spanned the restriction site and that portions of the mRNA were derived from both fragments. Thus, it can be

FIG. 4. Cell-free translation of late RNAs encoded on the 7-kbp region. Hybridization of late RNA to pV7 DNA filters and translational analysis were as described in the legend to Fig. 3. Protein synthesis reactions were programmed with no RNA, left lane; pV7-specific late RNA, center lane; and polyadenylated late RNA, right lane.

FIG. 5. Translational mapping of immediate early genes. CHI RNA was hybridized to an ordered set of immobilized DNA filters and translated in a reticulocyte lysate, and the resulting polypeptides were analyzed as in the legend to Fig. 3. DNA filters contained either the entire 7-kbp fragment or the cloned *HpaII* subfragments of the region as indicated above each lane. No RNA was added to the in vitro translation reaction displayed in the first lane.

seen that RNA selected by both the B and F fragments program the synthesis of the 38-kd protein. RNA hybridizing to three adjacent fragments, F, D, and C, directed the synthesis of the 63-kd protein, but curiously, the D fragment also selected RNA encoding a second protein of 37 kd, suggesting that the two genes overlap. No identifiable mRNA activity was retained by filters with E or G DNA fragments. In addition to the endogenous 48-kd band seen in the minus RNA control (lane 1), another source of background apparent in other lanes of this long autoradiographic exposure results from the translation of traces of RNA derived from other regions of the vaccinia virus genome which were incompletely removed by the stringent washing procedure.

Transcripts for the 13-, 15-, and 25-kd polypeptides were all selected by hybridization to the 2-kbp-long *HpaII*-A fragment. To resolve this cluster of genes, three DNA filters were prepared with subfragments of the *HpaII*-A region obtained by digestion with the restriction enzyme *Clal* and purified by preparative agarose gel electrophoresis. RNA eluted from all three DNA filters was translated to produce the 25-kd protein (Fig. 6). It may be calculated that approximately 750 base pairs is required to encode a polypeptide of 25 kd, suggesting the transcription of extensive noncoding sequences in this region. DNA filter I selected mRNA

FIG. 6. Ordering of genes specifying 13-, 15-, and 25-kd polypeptides on the *HpaII*-A fragment by translational analysis. The *HpaII*-A fragment of pV7 DNA was digested with *ClaI*, and the resulting subfragments were recovered from agarose gel slices by electroelution. Three immobilized DNA filters were prepared from these subfragments as indicated at the bottom. Immediate early RNA which hybridized to each filter was subjected to translational analysis as described in the legend to Fig. 3. No RNA was added to the translation reaction displayed in the first lane.

encoding all three proteins, whereas filter II selected RNA specifying only the 15- and 25-kd proteins and filter III selected only the 25-kd protein. On the basis of these data the linear order of the three genes was determined to be 13, 15, and 25 kd from left to right.

A generalized transcriptional map of the 7-kbp region was obtained by using the ordered set of *HpaII* subfragments to probe strips cut from a blot of CHI RNA resolved by gel electrophoresis (Fig. 7). A major family of overlapping RNAs of 1.14, 1.42, and 1.75 kb was identified with both the *HpaII*-B and -F probes. A second major band of 1.95 kb could be seen with F, D, and C probes. The *HpaII*-D probe also identified a minor RNA of 1.2 kb. Based on a consideration of the translational map described above, it is apparent that the 1.2-kb RNA encodes the 37-kd protein, the 1.95-kb RNA encodes the 63-kd protein, and the RNA family derived from the B/F region is translated to produce the 38-kd polypeptide. Three minor RNA bands of 2.9, 3.2, and 3.5 kb could be seen in longer exposures in lanes probed with *HpaII*-B, -F, -D, or -C fragments. RNAs of several sizes

were identified with the *HpaII*-A fragment probe including major bands of 1.05, 1.25, and 1.77 kb and minor bands of 3.15 and 3.6 kb. In experiments in which RNAs were resolved on a 2% gel, additional species of 0.35 and 0.5 kb were seen. Occasionally a minor RNA of 1.9 kb was detected. The relative intensity of the higher-molecular-weight bands differed somewhat from one experiment to another. Whether these variations represent a real difference in RNA abundance or are due to the experimental technique is not known.

Direct autoradiography of RNA after in vitro cap labeling provides an alternative means of visualizing individual transcripts which is independent of RNA chain length. The naturally occurring 7-methylguanosine cap was removed from CHI RNA by sodium periodate oxidation and β -elimination. RNA was then recapped with [α - 32 P]GTP and vaccinia guanylyl transferase as described by Venkatesan and Moss (20). This procedure introduces 1 mol of radioactive phosphate per mol of RNA labeled. [32 P]RNA was fractionated by preparative hybridization to DNA filters containing cloned vaccinia virus sequences as above, and

FIG. 7. Hybridization blot analysis of immediate early RNAs derived from the 7-kbp region. CHI RNA was resolved on a 1.5% agarose gel after glyoxal denaturation. A nitrocellulose blot of the gel was cut into strips for hybridization to the cloned *HpaII* fragments of the 7-kbp region. [32 P]DNA probes labeled by nick translation are indicated above each strip shown in the autoradiograph. The size of RNA bands was determined by comparison to the position of denatured DNA restriction fragments produced by digestion of phage λ DNA with *EcoRI* and *HindIII*.

FIG. 8. Electrophoretic analysis of immediate early RNAs hybridizing to immobilized DNA filters. CHI RNA was cap labeled in vitro and hybridized to immobilized DNA filters. Specifically retained RNA was eluted, denatured with glyoxal, and resolved by electrophoresis in a 1.5% agarose gel. An autoradiograph of the gel is shown. For reference, the positions of denatured restriction fragments of *Eco*RI- and *Hind*III-digested phage λ DNA are indicated in the right margin. Filters contained the following DNA sequences as indicated above each lane: cloned *Hpa*II fragment B, D, C, or A; pV7; B26; or separated heavy (H) or light (L) strands of B26.

the selected RNAs were displayed by gel electrophoresis after denaturation (Fig. 8). RNA bands of the same size as those identified by the hybridization blot technique were found. It can be seen that the dominant transcripts derived from this region include the 1.95-kb RNA mapping to the *Hpa*II-F, -D, and -C fragments and the overlapping family of 1.14-, 1.42-, and 1.75-kb RNAs encoded on the B and F fragments. As detected by this method the larger RNAs of 3 to 4 kb appear to be relatively minor species. Transcripts of 1.05, 1.25, and 1.77 kb were found with the *Hpa*II-A fragment. In addition to these, RNAs of 350 and 500 nucleotides were also seen in similar experiments (not shown).

The polarity transcription was determined by using a recombinant phage vector to obtain the separated DNA strands of the 7-kbp region. Clone B26 was selected with a probe of pV7 DNA by plaque hybridization methods (2) from a genomic library of chimeric phage containing vaccinia virus sequences cloned in the λ Charon 4 vector. This clone was found to contain a 10.1-kbp vaccinia virus DNA insert including the 7-kbp *Eco*RI fragment and an adjacent 3.1-kbp sequence extending to the right into the *Hind*III-E fragment. The pattern of RNAs hybridizing to B26 DNA filters appears similar to those of RNA selected by pV7 filters (Fig. 8) but includes an additional band of 2.5 kb, presumably derived from the region to the right of the 7-kbp region. In experiments with filters containing the separated strands of B26

DNA, it can be seen that all RNAs detected are derived exclusively from one strand, the light strand. Additional hybridization studies (not shown) have established that transcription of sequences on the light strand corresponds to the leftward-reading direction in the vaccinia virus genome.

Independent confirmation of the polarity of transcription and a refinement of the RNA map resulting from hybridization blot methods have been obtained by S1 nuclease studies. A modification of the original method of Berk and Sharp (3) developed by Weaver and Weissman (21) was used. DNA probes labeled uniquely at the terminus of one strand were obtained by secondary restriction enzyme treatment of fragments which were end labeled after a primary enzyme digestion. Such probes unambiguously determine the polarity of transcription since heteroduplex formation will be detected only when the coding strand is labeled. The size of the labeled DNA sequence protected from S1 digestion corresponds to the distance from the RNA terminus to the restriction site in the heteroduplex. In initial experiments with S1 nuclease, hybridization conditions ensuring DNA excess were determined by titrating a fixed amount of DNA probe with various concentrations of CHI RNA (not shown). Under these conditions the intensity of multiple bands visualized with a single probe accurately reflected the relative abundance of overlapping RNA species.

Probes labeled at the *Sst*I site at map position 1.9 located the termini of the overlapping RNAs encoding the 38-kd protein. A single major 5' RNA fragment of 420 nucleotides (Fig. 9) and three 3' RNA fragments of 640, 930, and 1,260 nucleotides (Fig. 10) were identified, defining a family of three transcripts possessing a common 5' terminus mapping to 2.32 and extending leftward for 1.06, 1.35, and 1.68 kb (Fig. 11). These numbers are in good agreement with the values of 1.14, 1.42, and 1.75 kb resulting from the RNA blotting studies which include a polyadenylate tract of approximately 100 nucleotides not measured in the S1 technique.

Probes labeled at the *Pst*I site (position 2.85) detected mRNAs of 1.86 and 1.10 kb. Two 5'-terminal RNA fragments of 590 and 1,350 nucleotides and a single major 3'-terminal fragment of 510 nucleotides were seen. These bands are derived from two RNA species with 5' ends mapping to positions 3.44 and 4.20 and which are 3' coterminal at position 2.36. Thus, the entire nucleotide sequence of the smaller RNA is contained within the larger, which has its 5' terminus located 760 nucleotides upstream.

In addition to more precisely locating the ends of the major transcripts described above, S1 nuclease studies have identified RNAs species overlapping two or more genes. These larger RNAs appear to possess a common 5' terminus with transcripts derived from the first gene and are coterminal at the 3' end with RNAs derived from downstream genes. Thus, the 2,300-nucleotide RNA band seen with the 5' *Sst*I probe and the three larger bands seen with the 3' *Pst*I probe define a family of three transcripts extending leftward from 4.2, spanning the entire region encoding the 63-, 37-, and 38-kd polypeptides.

A similar pattern of overlapping transcription has been found for the three genes clustered on the *Hpa*II-A fragment. Bands of 1,600, 880, 200, and 560 nucleotides detected with the *Sal*I and *Hpa*I probes reflect transcripts whose 5' termini map to three positions at 5.72, 6.39, and 7.12. 3' RNA fragments terminating at closely spaced sites in two general locations have been identified. Three bands obtained with the *Cla*I probe reflect RNAs with 3' termini at 5.43, 5.40, and 5.29, and the doublet visualized with the *Hpa*I probe locates

FIG. 9. Nuclease S1 analysis of RNA 5' termini. Hybridization probes uniquely end labeled in one strand were prepared as follows: pV7 DNA was cleaved with *Sst*I, *Pst*I, *Sal*I, or *Hpa*I and 5' terminally labeled, using T4 polynucleotide kinase. After digestion with *Eco*RI, restriction fragments labeled at the left end and extending to the *Eco*RI site at the right side of the 7-kbp insert were recovered from agarose gels. Probes were hybridized in DNA excess to immediate early RNA under conditions promoting heteroduplex formation and overhanging ends were trimmed with S1 nuclease. For each analysis an identical hybridization was performed in which RNA was omitted. After glyoxal denaturation nuclease-resistant bands were resolved by electrophoresis in 1.5% agarose and visualized by autoradiography. For each probe the reaction in which RNA was omitted is shown in the left lane, and the reaction containing RNA is shown in the right lane. The size (in kilobases) of protected fragments was determined by comparison to the position of denatured restriction fragments of *Hind*III- and *Eco*RI-digested phage λ DNA. The location of the relevant restriction sites within the 7-kbp region and the map position of the S1-protected bands are schematically illustrated in the insert at the bottom.

RNAs 3' terminating at 5.98 and 5.93. These data permit the RNA species identified by hybridization blot studies to be positioned as shown in Fig. 11 and, in conjunction with the translational map obtained from the experiment shown in Fig. 6, suggest tentative assignment of the overlapping RNAs encoding the 13-, 15-, and 25-kd proteins. It can be seen that the 1,250- and 1,770-nucleotide transcripts have a

common 5' terminus at 7.12 and are complementary to sequences on all three DNA filters of Fig. 6; hence, these RNAs probably represent two forms of the messenger for the 25-kd protein. The 5' end of the 500- and 1,050-nucleotide transcripts originate at 6.4 and both contain sequences found on DNA filters I and II, correlating with the location of the mRNA activity for the 15-kd protein. The 350-

FIG. 10. Nuclease S1 analysis of RNA 3' termini. A strategy similar to that described in the legend to Fig. 9 was used to prepare 3' terminally labeled hybridization probes. pV7 DNA was cleaved with *Sst*I, *Pst*I, *Cla*I, or *Hpa*I and labeled in the 3' position, using terminal transferase. After digestion with *Eco*RI, restriction fragments labeled at the right end (*Sst*I, *Pst*I, and *Cla*I probes) and extending to the *Eco*RI site at the left end of the 7-kbp insert were recovered from agarose gels. (Due to the presence of several sites for *Hpa*I within the 7-kbp sequence, the 1.1-kbp *Hpa*I probe used was labeled at the 3' terminus of both strands.) Hybridization to immediate early RNA, S1 nuclease treatment, and denaturing agarose gel electrophoresis were described in the legend to Fig. 9. The location of relevant restriction sites and the map position of S1-resistant bands are schematically illustrated at the bottom.

nucleotide RNA starts at 5.7 and is 3' coterminal with both the 1,770- and 1,050-nucleotide transcripts. This 350-nucleotide transcript is derived from DNA sequences found only on filter I and is therefore presumed to encode the 13-kd protein. Other RNA species of 3.7, 3.1, 2.5, and 1.9 kb detected by hybridization blots are believed to represent RNAs containing 3'-proximal sequences derived from the *Hpa*II-A region and whose 5' termini map to a position to the right of the 7-kbp fragment. The 600-nucleotide band obtained with the 5'-labeled *Hpa*I probe corresponding to protection of the full-length DNA fragment suggests the

existence of such larger RNAs reading through into the *Hpa*II-A region.

RNA isolated from infected cells at late times includes many of the same transcripts found earlier, as shown by hybridization blot methods given above (Fig. 2). These early transcripts have also been detected by S1 nuclease analysis of late RNA. In addition, a new species encoding the 40-kd late protein was identified, transcribed in the same leftward direction as the six early genes. The 5' end of this RNA was located to position 5.43 as a result of the 450-nucleotide band obtained with a DNA probe 5' end labeled at the *Bgl*II site at

FIG. 11. Transcriptional and translational map of 7-kbp region. Beneath a map of the 7-kbp region and the lettered *HpaII* segments, horizontal arrows indicate the polarity and map location of RNA species. The length (in kilobases) is shown to the right of RNAs, and below each gene the size of the in vitro translation product is given in kilodaltons. Six genes are expressed at immediate early times, and the one encoding the 40-kd polypeptide is expressed only at late times. A dashed arrow indicates the 3'-terminal heterogeneity of late RNA.

4.98 (Fig. 12). In contrast, no distinct bands but only a broad smear was detected with 3'-labeled DNA probes (not shown). Thus, the size heterogeneity of transcripts derived from this late gene is apparently due to a variable length of 3'-proximal sequences.

Having characterized authentic transcripts derived from the 7-kbp region which accumulate in infected cells, we were in position to reevaluate the fidelity of transcription in vitro by permeabilized viral cores. Despite the high background obtained with core RNA, bands similar in size to the major species found in CHI RNA had been identified by hybridization blot analysis (see above). The close correspondence of core and CHI RNAs was further established by direct autoradiography of ^{32}P -labeled core RNA selected by immobilized DNA filter binding. Specifically retained transcripts were eluted and resolved by gel electrophoresis. The results of one experiment with *HpaII*-A DNA filters is shown in Fig. 13. Five RNA bands of 0.35, 0.50, 1.1, 1.4, and 1.8 kb were seen, in excellent agreement with the sizes of five CHI RNAs mapped to the *HpaII*-A region.

DISCUSSION

Until a map of the entire genome of vaccinia virus is available, local regions of the DNA will be selected for detailed transcriptional and translational characterization on the basis of possessing features of special interest. In view of our goal of obtaining a better understanding of poxvirus mRNA biogenesis, our approach was to choose a region of the genome strongly expressed at the level of RNA synthesis by the virion-associated transcription system. Although we have not presented any quantitative data, we have noted that transcripts of the genes located on the 7-kbp fragment are relatively abundant compared with other regions. RNA accumulation in infected cells is a complex variable affected by degradative as well as synthetic factors; representation in core RNA presumably reflects more directly the relative rate of transcription. By both criteria transcripts of this region were found in particular abundance, which facilitated the mapping of minor RNA species derived from these genes.

Six early genes were identified on the 7-kbp fragment

FIG. 12. Nuclease S1 5'-terminal mapping of late RNA encoding the 40-kd polypeptide. A 2.2-kbp *BglII*-*EcoRI* restriction fragment derived from the right portion of the 7-kbp region and 5' terminally labeled at the *BglII* site was hybridized to late RNA, and heteroduplex formation was analyzed as described in the legend to Fig. 9. Lanes (-) and (+) show the results of hybridizations in which RNA was omitted or included. The size of the RNA-dependent band was determined by reference to *EcoRI*- and *HindIII*-digested phage λ DNA (lane m).

FIG. 13. Electrophoretic analysis of in vitro RNA derived from the *HpaII*-A region. 32 P-labeled polyadenylated RNA synthesized by viral cores in vitro was hybridized to immobilized DNA filters containing the cloned *HpaII*-A fragment. Specifically retained RNA was denatured with glyoxal and resolved on a 1.5% agarose gel. An autoradiograph of the gel including the in vitro RNA (lane R) and size markers of *EcoRI*- and *HindIII*-digested phage λ DNA (lane m) is shown.

arranged in two clusters of three adjacent genes separated by a single late gene. All seven genes were encoded on the same DNA strand and transcribed toward the left end of the genome. In other studies a dominant direction of transcription has also been reported. At least 9 of 11 genes on the left end of the genome are transcribed leftward (8); however, RNA from the thymidine kinase and neighboring genes located on the *HindIII*-J fragment in the center of the genome are copied from the opposite, rightward-reading strand (1). It is evident that at least one switch in coding strands must occur somewhere between the 7-kbp *EcoRI* fragment characterized above and the *HindIII*-J region.

The striking feature of the transcriptional map reported here is the overlapping nature of early RNAs. In addition to a major RNA species 3' terminating shortly after the presumptive coding region of a particular gene, typically one or more minor species were also found containing 3'-proximal sequences derived from readthrough by RNA polymerase into downstream genes. For one gene specifying the 25-kd polypeptide only readthrough transcripts were identified. Another curious finding was the overlap of the two genes determining the 37- and 63-kd proteins. The 1.20-kb nucleotide sequence encoding the 37-kd polypeptide is also found

as the 3'-proximal portion of the 1.95-kb messenger for the 63-kd polypeptide. It is not known whether this sequence is translated in the same reading frame or a different one on the two RNAs. However, in both cases approximately the full coding capacity of the mRNA is utilized. In view of the evident lack of RNA splicing by poxviruses and the correspondence of RNA capping and initiation sites, this observation suggests that the 1.20-kb RNA initiates at a position which is internal to the coding region of the 63-kd polypeptide.

On the basis of sedimentation analysis and hybridization studies of total late mRNA, it had previously been inferred that some large late messengers contain sequences also present in the smaller early mRNA molecules (15). The transcripts we have identified derived from the late gene on the 7-kbp fragment do indeed read through into adjacent early genes. A unique 5' terminus has been located by S1 nuclease mapping, establishing that the size heterogeneity of these RNAs is entirely due to a variable length of 3'-proximal sequence.

Expression of early genes located on this DNA fragment was characterized throughout the course of infection. Transcripts synthesized by viral cores in vitro were indistinguishable from those produced by infected cells at early times, both in terms of physical criteria and by translational analysis in vitro. Although early messengers persist at low levels in infected cells at late times and were detected by S1 nuclease and hybridization blot methods, they apparently were largely biologically inactive. Only mRNA encoding the 13-kd polypeptide retained the ability to direct protein synthesis in a reticulocyte lysate when isolated at late times. The basis for this phenomenon is not understood but may suggest a role for translational regulation in the switching off of early gene expression.

ACKNOWLEDGMENTS

We thank Joel Gottesfeld for critical reading of the manuscript, M. Sarmiento-Batterson for the generous gift of rabbit reticulocyte lysate, and Ame Holden for excellent technical assistance.

This research was supported by Public Health Service research grant AI15941 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. Bajsar, G., R. Wittek, J. P. Weir, and B. Moss. 1983. Vaccinia virus thymidine kinase and neighboring genes: mRNAs and polypeptides of wild-type virus and putative nonsense mutants. *J. Virol.* **45**:62-72.
2. Benton, W. D., and R. W. Davis. 1977. Screening λ gt recombinant clones by hybridization to single plaques *in situ*. *Science* **196**:180-182.
3. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12**:721-732.
4. Bochner, B. R., H. Huang, G. L. Schieven, and B. F. Amers. 1980. Positive selection for loss of tetracycline resistance. *J. Bacteriol.* **143**:926-933.
5. Bolivar, F., and K. Backman. 1979. Plasmids of *E. coli* as cloning vectors. *Methods in Enzymol.* **68**:245-267.
6. Bonner, W. M., and R. A. Laskey. 1974. A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels. *Eur. J. Biochem.* **46**:83-88.
7. Cooper, J. A., R. Wittek, and B. Moss. 1981. Hybridization selection and cell-free translation of mRNA's encoded within the inverted terminal repetition of the vaccinia virus genome. *J. Virol.* **37**:284-294.
8. 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.
9. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics, p. 114–115. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 10. **Hanahan, D., and M. Meselson.** 1980. Plasmid screening at high colony density. *Gene* **10**:63–67.
 11. **Hruby, D. E., L. A. Guarino, and J. R. Kates.** 1979. Vaccinia virus replication. I. Requirement for the host-cell nucleus. *J. Virol.* **29**:705–715.
 12. **Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, and A. Efstratiadis.** 1978. The isolation of structural genes from libraries of eukaryotic DNA. *Cell* **15**:687–701.
 13. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–560.
 14. **McMaster, G. K., and G. C. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. U.S.A.* **74**:4835–4838.
 15. **Oda, K. I., and W. K. Joklik.** 1967. Hybridization and sedimentation studies on early and late vaccinia messenger RNA. *J. Mol. Biol.* **27**:395–419.
 16. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur. J. Biochem.* **67**:247–256.
 17. **Seed, B.** 1982. Diazotizable arylamine cellulose papers for the coupling and hybridization of nucleic acids. *Nucleic Acids Res.* **10**:1799–1810.
 18. **Smith, H. O.** 1980. Recovery of DNA from gels. *Methods Enzymol.* **65**:371–379.
 19. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. U.S.A.* **77**:5201–5205.
 20. **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.
 21. **Weaver, R. F., and C. Weissman.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S β -globin mRNA precursor and mature 10S β -globin mRNA have identical map coordinates. *Nucleic Acids Res.* **7**:1175–1193.
 22. **Wittek, R., E. Barbosa, J. A. Cooper, C. F. Garon, H. Chan, and B. Moss.** 1980. Inverted terminal repetition in vaccinia virus DNA encodes early mRNAs. *Nature (London)* **285**:21–25.
 23. **Wittek, R., J. A. Cooper, E. Barbosa, and B. Moss.** 1980. Expression of the vaccinia virus genome: analysis and mapping of mRNAs encoded within the inverted terminal repetition. *Cell* **21**:487–493.