JOURNAL OF VIROLOGY, Dec. 1985, p. 830–838
0022-538X/85/120830-09$02.00/0

Vol. 56, No. 3

Transcriptional and Translational Mapping and Nucleotide Sequence Analysis of a Vaccinia Virus Gene Encoding the Precursor of the Major Core Polypeptide 4b

JOHANNES ROSEL AND BERNARD MOSS*

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20205

Received 14 June 1985/Accepted 19 August 1985

We prepared antiserum that reacted with a major core polypeptide of approximately 62,000 daltons (62K polypeptide), designated 4b, and its 74K precursor, designated P4b. A cell-free translation product of vaccinia virus late mRNA that comigrated with P4b was specifically immunoprecipitated. The late mRNA encoding P4b hybridized to restriction fragments derived from the left end of the *Hin*dIII A fragment and to a lesser extent from the right side of the *Hin*dIII D fragment. A polypeptide that comigrated with P4a, the precursor of another major core polypeptide, was synthesized by mRNA that hybridized to DNA segments upstream of the P4b gene. Complete nucleotide sequence analysis of the P4b gene revealed an open reading frame, entirely within the *Hin*dIII A fragment, that was sufficient to encode a 644-amino-acid polypeptide of 73K. The 5′ end of the P4b mRNA was located at or just above the translational initiation site.

Vaccinia virus, the best-studied member of the poxvirus family, has a 185-kilobase-pair (185-kb), linear, double-stranded DNA genome which codes for about 200 polypeptides. Transcriptional and translational mapping studies indicate that early and late genes are distributed along the length of the DNA molecule (1), except for about 2.5 kb at each end which consists largely of tandem repeats (38). The genes that have been examined have continuous coding sequences (4, 16, 33, 35, 36), and there is no evidence of splicing. Transcriptional regulatory signals, which are recognized by the vaccinia virus RNA polymerase and associated factors, are located immediately upstream of the RNA start sites of both early and late genes (4, 8, 22, 36) and may be 31 base pairs or less in length (9). Although transcriptional termination usually occurs just distal to the coding segments of early genes, late transcripts typically pass over several genes downstream and appear to lack discrete 3′ ends (11, 23, 36).

Thus far, only a few genes of known or suspected biological function have been mapped. These include a polypeptide with homology to epidermal growth factor (6, 7, 33), thymidine kinase (15, 34), DNA polymerase (18), one subunit of RNA guanylyltransferase (27), and several RNA polymerase subunits (E. V. Jones, C. Puckett, and B. Moss, communicated at the 1984 annual meeting of the American Society for Virology), all of which are early genes. Some of the major virion polypeptides, which are expressed late in infection, also have been mapped (36, 39, 40). Of the above early and late genes, the nucleotide sequences of only the thymidine kinase (16, 35), the growth factor (33), the 28-kilodalton precursor (28K precursor) of a core polypeptide (36), and the 11K structural polypeptide (4) have been reported.

In the present study we have mapped and sequenced the gene for a major core polypeptide of approximately 62K that was previously designated 4b (19, 29). This protein is of particular interest because its abundance suggests efficient gene expression and because processing of the 74K precursor polypeptide is linked to virus assembly (19, 29). While our sequencing was in progress, Wittek and co-workers (40) reported the mapping of 4b and of another core polypeptide designated 4a.

(A preliminary account of this work was presented at the 1984 Annual Meeting of the American Society for Virology.)

MATERIALS AND METHODS

Virus and cells. Vaccinia virus (strain WR) was grown in HeLa cell suspension cultures maintained in Eagle medium containing 5% horse serum.

In vitro labeling of proteins. HeLa cells were infected with 30 PFU of purified virus per cell and at 6 h were pulse-labeled for 20 min with 125 μCi of [^{35}S]methionine. Cells were harvested immediately after the pulse or after a 2-h chase in complete medium.

Preparation of antisera. Purified vaccinia virions were incubated with 0.25 M NaCl, 0.1 M Tris hydrochloride (pH 8.5), 0.01 M dithiothreitol, and 0.2% sodium deoxycholate for 30 min at 0°C. After centrifugation in an SW41 rotor at 35,000 rpm for 30 min at 4°C, the protein pellet was dissolved in gel loading buffer (0.0625 M Tris hydrochloride [pH 6.8], 4% sodium dodecyl sulfate, 2% mercaptoethanol, 10% glycine, 0.003% bromophenol blue) and heated at 100°C for 3 min. Proteins from about 10 mg of dissociated virions were purified by electrophoresis on a 10% polyacrylamide gel. The gel was lightly stained with Coomassie blue, the unseparated P4a-P4b protein doublet was excised, Dounce homogenized, and eluted overnight into 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA. Rabbits were immunized by subcutaneous injection of the purified protein in complete Freund adjuvant. The animals were boosted 4 weeks later by three weekly injections of purified protein in incomplete Freund adjuvant. Each rabbit received a total of 450 μg of protein over this time period.

Hybridization selection of RNA. Total cytoplasmic RNA was prepared from cells 4 h after infection in the presence of cycloheximide or 6 h after infection in the absence of the

* Corresponding author.

inhibitor. RNA was purified by the method of Glisin et al. (12) as previously described (10). A 200-mg sample of total cytoplasmic RNA was hybridized to restriction fragments from cloned vaccinia virus DNA immobilized on nitrocellulose filters (1). Filters were then washed, and the specifically bound RNA was eluted and ethanol precipitated with tRNA carrier (28).

In vitro translation. The selected RNA was translated in a micrococcal nuclease-treated reticulocyte lysate prepared as described by Jackson and Hunt (17).

Immunoprecipitation. A 15-μl sample of the in vitro translation products was incubated at 4°C with 10 μl of preimmune serum for 2 h and then with 10 μl of protein A-Sepharose for an additional 2 h. Protein A-Sepharose was removed by centrifugation, and the supernatant was incubated at 4°C with 20 to 30 μl of antiserum for 4 h and with 20 μl of protein A Sepharose for an additional 12 h. The protein A-Sepharose was washed seven times with a triple-detergent buffer (1% Triton X-100, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 0.05 M Tris hydrochloride [pH 7.5], 0.15 M NaCl) and twice with 2 M urea–0.4 M LiCl–10 mM Tris hydrochloride (pH 8.0). Finally, the proteins were solubilized in gel loading buffer by heating for 3 min at 100°C and analyzed on a 10% polyacrylamide gel. [^{35}S]methionine-labeled vaccinia virus or cytoplasmic cell extracts were prepared for polyacrylamide gel analysis in the same manner.

Mapping of the 5′ end of mRNA by nuclease digestion. End-labeled DNA fragments were hybridized to late RNA and nuclease treated as previously described (36) with 580 U of S1 nuclease or 240 U of mung bean nuclease at 25°C for 1 h.

DNA purification. Plasmid DNA was purified as described by Birnboim and Doly (5). DNA fragments were isolated by electroblotting onto DEAE-paper (34), phenol extraction (2), or absorption to glass beads (41). Other techniques were performed by the methods of Maniatis et al. (24).

DNA sequencing. DNA was sequenced as described by Maxam and Gilbert (25) or by Sanger et al. (32) with mp18 or mp19 derivatives of phage m13 (26).

Computer analysis. Computer analysis was performed on an IBM 370 or DEC 10 computer with programs devised by Queen and Korn (31) and Kyte and Doolittle (21).

Materials. Restriction endonucleases were obtained from Bethesda Research Laboratories, New England Biolabs, or Boehringer Mannheim Biochemicals. T4 polynucleotide kinase, S1 nuclease, and mung bean nuclease were from Pharmacia P-L Biochemicals; T4 DNA ligase and large fragment of DNA polymerase I were from Bethesda Research Laboratories; calf intestinal phosphatase was from Boehringer Mannheim Biochemicals.

RESULTS

Preparation of antiserum. The object of this study was to map polypeptides 4a and 4b by cell-free translation of mRNA that was selected by hybridization to vaccinia virus DNA fragments. These core polypeptides are formed by proteolytic processing of higher-molecular-weight precursors designated P4a and P4b, respectively (19, 29). Since the processing step is complex and associated with virion assembly, we anticipated that the precursors would be the final product of in vitro translation. Nevertheless, it seemed likely that polyclonal antiserum directed against 4a and 4b would also react with the precursors. Therefore, we planned to identify P4a and P4b by immunoprecipitation of cell-free translation products of late vaccinia virus mRNA by antiserum raised against the smaller polypeptides 4a and 4b as well as by their characteristic electrophoretic mobilities on polyacrylamide gels.

Previous enzyme isolation studies (30) indicated that both 4a and 4b remained in a readily sedimentable form after treatment of purified vaccinia virions with sodium deoxycholate and dithiothreitol. Since this procedure solubilizes more proteins than Nonidet P-40 and mercaptoethanol, which are usually used to prepare cores, it provided a significant initial enrichment step. The protein pellet was dissociated with sodium dodecyl sulfate and mercaptoethanol and applied to a polyacrylamide gel. After electrophoresis, the 4a-4b doublet was visualized by lightly staining with Coomassie blue, excised, and eluted from the gel. The polypeptide mixture was then used to immunize rabbits.

The specificity of the antiserum was evaluated by immunoprecipitation of proteins from cytoplasmic extracts of vaccinia virus-infected cells that were labeled with [^{35}S]methionine. Initially we examined the total proteins that were pulse-labeled at 6 h after infection (Fig. 1A). After a 2-h chase, the amounts of labeled P4a and P4b were diminished, whereas other polypeptides including ones that migrated just above P4a and P4b actually increased slightly in intensity. The appearance of the prominent 4a-4b doublet during the chase was striking (Fig. 1A). Antiserum prepared against the mixture of 4a and 4b polypeptides immunoprecipitated P4b, but unexpectedly did not bind to P4a. After a 2-h chase, polypeptide(s) migrating with the 4a-4b doublet also were immunoprecipitated (Fig. 1A). The specific reactions of the

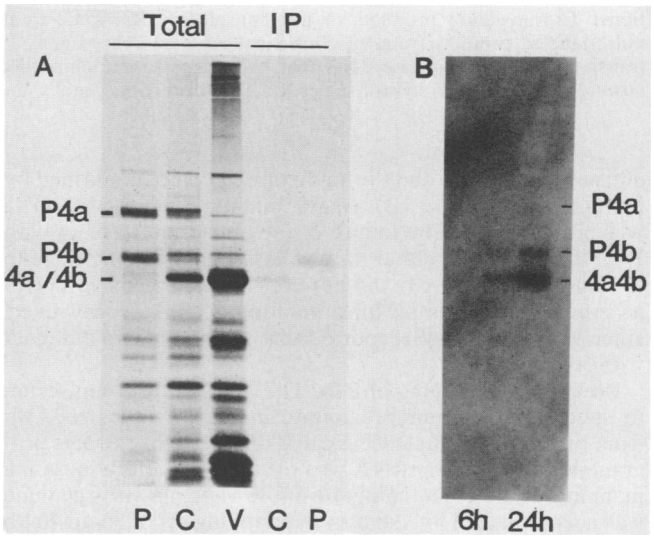

FIG. 1. Specificity of antiserum for P4b. (A) Immunoprecipitation of P4b and 4b from vaccinia virus-infected cells with rabbit antiserum. Cells were labeled for 20 min with [^{35}S]methionine at 6 h after infection and harvested either immediately (lane P) or after a 2-h chase in complete medium (lane C). Lane V shows [^{35}S]methionine-labeled structural proteins of vaccinia virus. Total and IP indicate proteins analyzed directly and after immunoprecipitation, respectively. An autoradiograph is shown. (B) Immunoblot analysis of proteins from vaccinia virus-infected cells. Cell extracts were obtained 6 or 24 h after infection. Proteins were separated on a sodium dodecyl sulfate-polyacrylamide gel and transferred to nitrocellulose. Antiserum was added, and ^{125}I-protein A was used to detect immobilized antibody-antigen complexes. An autoradiograph is shown.

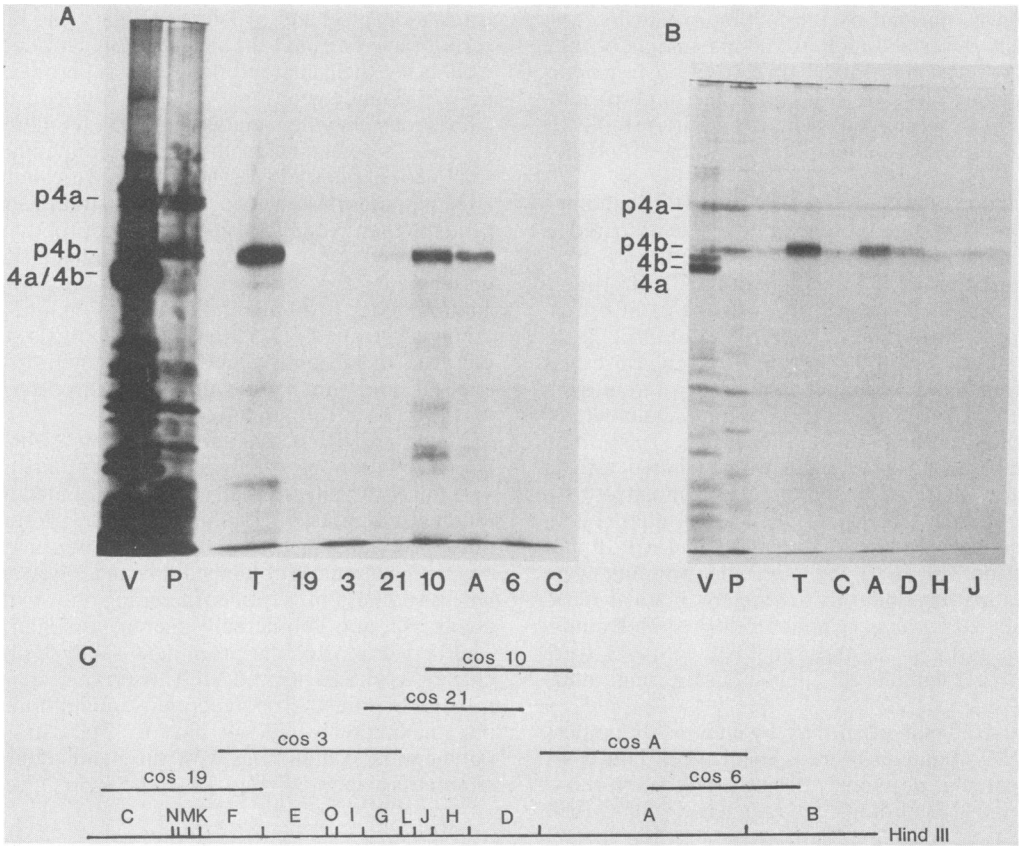

FIG. 2. Immunoprecipitation of the in vitro translation products of mRNA selected by hybridization to vaccinia virus DNA. Abbreviations: T, translation products of total unselected RNA; C, translation products with no RNA added; V, viral structural proteins; P, pulse-labeled proteins from vaccinia virus-infected cells. Lanes 19, 3, 21, 10, A, and 6 (panel A) and A, D, H, and J (panel B) contain translation products of late RNA that hybridized to immobilized cosmids or plasmids. *Hin*dIII restriction maps of vaccinia virus and the cosmid clones used to hybrid select RNA are shown in panel C.

antiserum with P4b and the 4a-4b doublet were confirmed by immunoblotting (Fig. 1B). Again, binding of antibody to P4a was not detected. The failure of the antiserum to react with P4a suggested that either 4a was less immunogenic than 4b or that antibodies to 4a did not cross-react extensively with its precursor. The latter interpretation is unlikely, however, since Wittek et al. (40) reported that antibody to 4a did react with its precursor.

Preliminary mapping of P4b. The ability of our antiserum to specifically immunoprecipitate in vitro synthesized P4b from micrococcal nuclease-treated reticulocyte extracts programmed with viral mRNA was established before initiating mapping studies. Screening the entire vaccinia virus genome was accomplished by using an overlapping set of 30- to 40-kb DNA fragments that were cloned in a cosmid vector (Fig. 2C). The cosmids were immobilized on nitrocellulose filters and used for hybridization selection of mRNA. Cosmids 10 and A, which contain overlapping vaccinia virus DNA, selected mRNA that codes for P4b (Fig. 2A). Trace amounts of P4b also were detected with mRNA that hybridized to cosmid 21.

Since cosmid A contains the *Hin*dIII A fragment of the vaccinia virus genome and the DNA segment in cosmid 10 overlaps *Hin*dIII-A, -D, -H, and -J (Fig. 2C), the P4b gene must map in the left portion of *Hin*dIII-A and possibly in the adjacent *Hin*dIII D fragment. To examine this further, late mRNA was also selected by hybridization to cloned *Hin*dIII D, H, and J fragments immobilized on nitrocellulose filters. Synthesis of P4b was directed by mRNA that hybridized to the *Hin*dIII D fragment, although to a lesser extent than by mRNA that hybridized to the *Hin*dIII A fragment (Fig. 2B). P4b was not detected in the translation products of mRNA that hybridized to *Hin*dIII-J or -H; the faint band just below P4b is found in all tracks including the control lacking added mRNA (Fig. 2B).

Fine mapping within the *Hin*dIII A fragment. The above data suggested that P4b mapped near the junction of *Hin*dIII fragments A and D. For finer mapping, we cloned the left 14-kb *Hin*dIII-*Kpn*I subfragment of *Hin*dIII-A (designated Aa) and four *Bam*HI subfragments of Aa (designated A_{24}, A_{25}, A_{26}, and A_{27}) in a plasmid vector (Fig. 3D). Each plasmid was immobilized on a nitrocellulose filter and used to select late mRNA species. Both total and immunoprecipitated translation products of the selected mRNAs were analyzed by polyacrylamide gel electrophoresis. As shown by examination of total and immunoprecipitated polypeptides (Fig. 3B), P4b was synthesized by total late mRNA and mRNA selected by hybridization to the leftmost 4.1-kb *Hin*dIII-*Bam*HI subfragment of Aa designated A_{24}. P4b also was detected by translation of mRNA that hybridized to Aa itself, although it is very faint in the reproduction of the immunoprecipitation analysis. P4b was not detected among the translation products of RNA selected by hybridization to the other *Bam*HI fragments; however, mRNA coding for a

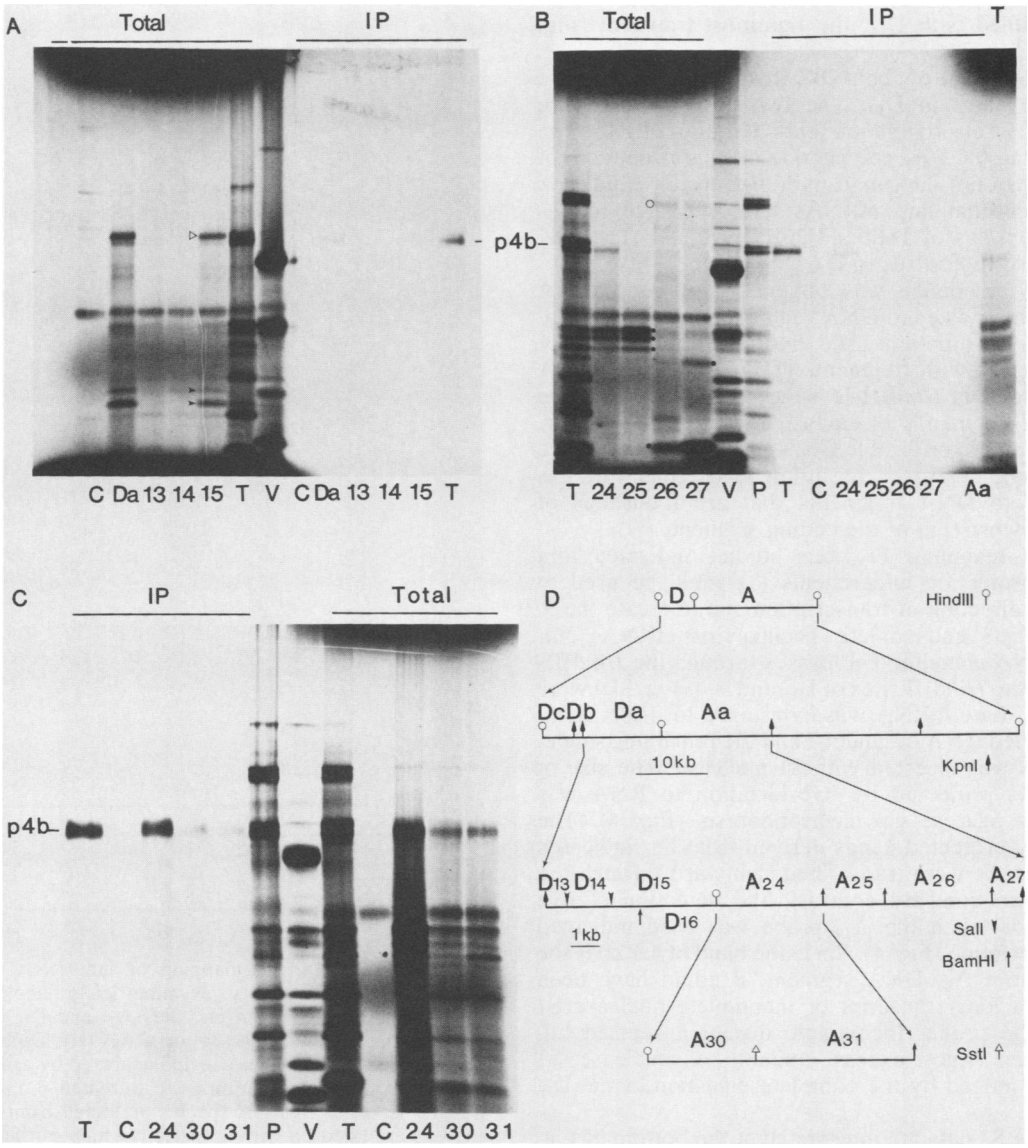

FIG. 3. Further analysis of the in vitro translation products of RNA selected by hybridization to vaccinia virus DNA. Restriction sites (↓, *Sal*I; ↑, *Bam*HI; ♀, *Hin*dIII; ↟, *Kpn*I; ♀, *Sst*I) within the *Hin*dIII D and A fragments as well as the code names for recombinant plasmids are indicated in panel D. Shown are autoradiographs of polyacrylamide gels containing cell-free translation products of late mRNA that was selected by hybridization to DNA fragments subcloned form *Hin*dIII-D (panel A) or *Hin*dIII-A (panels B and C). Code numbers of plasmids used to select mRNA are indicated below the lanes and correspond to those in panel D. Abbreviations: IP, immunoprecipitation; C, no RNA; T, total RNA; P, pulse-labeled proteins from vaccinia virus-infected cells; V, virion proteins.

polypeptide that comigrated with P4a was selected with subfragments A_{26} and A_{27} (Fig. 3B). Although we could not confirm the identity of the polypeptide by immunoprecipitation because of the specificity of our antiserm for P4b, the map location agrees with that recently reported for P4a by Wittek and co-workers (40).

A single *Sst*I site in the *Hin*dIII-*Bam*HI fragment A_{24} was used to subclone the *Hin*dIII-*Sst*I fragment A_{30} and the *Sst*I-*Bam*HI fragment A_{31}. Late mRNA encoding P4b was selected by hybridization to both of these plasmids (Fig. 3C).

During the course of mapping P4b we also determined the locations of additional polypeptides within the left 14 kb of the *Hin*dIII A fragment. Since these data may be useful to others, we note that mRNAs coding for late proteins of 38K, 36K, and 34K are selected by fragments A_{24} and A_{25} (Fig. 3B) and therefore map very close to P4b. Late mRNAs that encode polypeptides of 18K and 30K hybridized to fragments A_{26} and A_{27}, respectively (Fig. 3B). A major 14K translation product was made with RNA that hybridized to both of the latter DNA fragments (Fig. 3B). The identities of some of these proteins are currently under investigation.

Fine mapping within the *Hin*dIII D fragment. To complete the mapping studies, we subcloned the two large *Hin*dIII-*Kpn*I fragments of *Hin*dIII-D and used them for hybridization selection (Fig. 3D). Since P4b was synthesized only with the right subfragment Da, the latter was further subcloned as the *Kpn*I-*Sal*I D_{13}, *Sal*I-*Sal*I D_{14}, and *Sal*I-*Hin*dIII D_{15} subfragments. Positive selection of P4b mRNA

was only obtained with D_{15}, the rightmost fragment (Fig. 3A).

A major polypeptide of about 70K that migrates just above P4b (Fig. 3A) maps in D_{15}. The two polypeptides were distinguished by selective immunoprecipitation of P4b. The mRNA encoding the 70K polypeptide maps downstream of P4b since it does not appear to hybridize to the *Hin*dIII A fragment. Additional late mRNAs that hybridize to D_{15} encode polypeptides of 18K and 20K (Fig. 3A). The 20K polypeptide also hybridizes to D_{14} (Fig. 3A).

In summary, evidence was obtained that late mRNA encoding P4b hybridizes to DNA segments on both sides of the *Hin*dIII D-A junction. The hybridization was more efficient, however, with fragments derived from *Hin*dIII-A than with those from *Hin*dIII-D, suggesting that the coding region was predominantly or exclusively within the former. The long region of hybridizable DNA may be a consequence of the 3′-terminal heterogeneity of late transcripts: mRNAs can hybridize to DNA fragments that are thousands of nucleotides downstream of the coding segment.

Nuclease S1 mapping. Previous studies indicated that nuclease S1 protection experiments (3) could be used to determine the direction of transcription and to locate the 5 ends, but not the 3′ ends, of late vaccinia virus mRNAs (36). Since the mRNA encoding P4b passes through the *Hin*dIII-D-A junction, the *Hin*dIII sites of D_{16} and A_{24} (Fig. 3D) were 5′ end labeled. Late mRNA was hybridized to these asymmetrically labeled DNA fragments, and the remaining single-stranded DNA was digested with S1 nuclease. The size of DNA segments protected by hybridization to RNA was determined by agarose gel electrophoresis (Fig. 4). The absence of any protected bands derived from D_{16} indicated that no late RNAs were transcribed rightward through the *Hin*dIII-D-A junction. By contrast, the detection of five protected bands when the A_{24} probe was used indicated leftward transcription (Fig. 4). Since the band of 4.1 kb is the size of the entire A_{24} DNA segment, it could have been derived from a long transcript or incomplete nuclease S1 digestion of the probe. Incomplete digestion seemed unlikely, however, since excess amounts of nuclease S1 were used, as judged by the complete digestion of the D_{16} probe.

The nuclease S1 data are interpreted at the bottom of Fig. 4. Since mRNA that encodes polypeptide P4b was hybrid selected by the *Sst*I-*Bam*HI fragment D_{31} (Fig. 3D), the three RNA start sites downstream of the *Sst*I site (which lead to protection of 0.5-, 1.2-, and 1.4-kb DNA segments) must be downstream of the P4b gene translational start site. The mRNA encoding polypeptide P4b was not selected by hybridization to A_{25} (Fig. 3B), suggesting that it is encoded within the *Sst*I-*Bam*HI fragment. Therefore, the RNA that protected the 3.5-kb DNA segment was thought to be the message for P4b.

Sequencing. To clarify the gene organization, we determined the DNA sequence between the two *Xba*I sites in fragment A_{24} (Fig. 4). The *Xba*I fragment was cloned in M13 mp18 and mp19, and unidirectional deletion mutants were constructed by using time-controlled exonuclease III treatment as described by Henikoff et al. (14). The segment between successive deletions was kept to about 100 base pairs, and both strands were completely sequenced by using a universal primer. There is an ATG-initiated open reading frame of 1,932 base pairs between nucleotides 227 and 2158 (Fig. 5). If the first ATG were used to initiate translation, the predicted 644 amino acid polypeptide would be 73K. If the second ATG in the same open reading frame were used for initiation, a 558-amino-acid polypeptide of 63K would be formed. Initiation further downstream would give proteins considerably smaller than P4b.

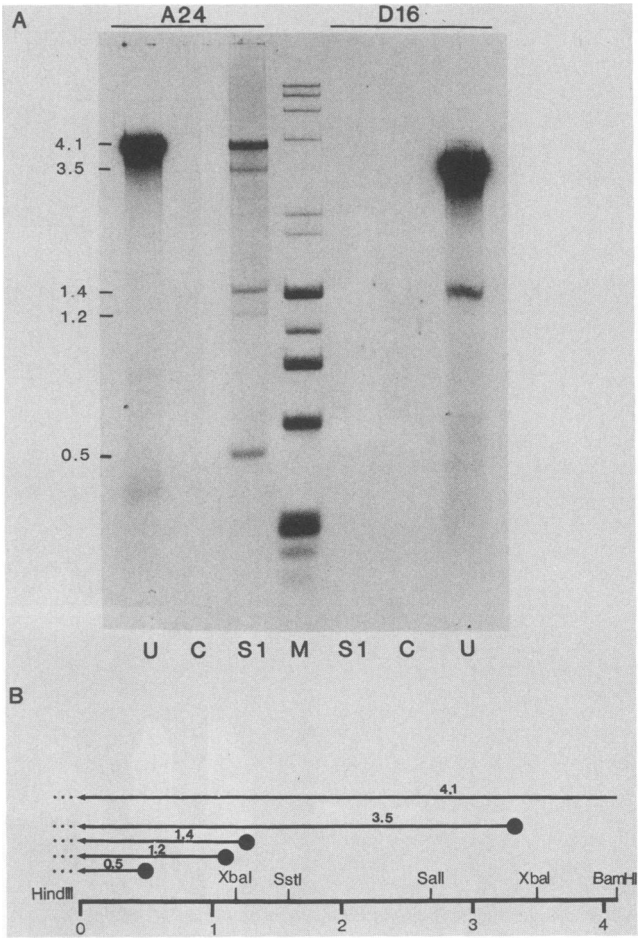

FIG. 4. Nuclease S1 mapping of late RNA transcripts. (A) *Bam*HI-*Hin*dIII fragments, asymmetrically labeled at the latter sites, were obtained from plasmids A_{24} and D_{16} (Fig. 3D). After hybridization to late RNA (S1) or tRNA (C), single-stranded DNA was digested with S1 nuclease and analyzed by agarose gel electrophoresis. Other abbreviations: U, untreated probe; M, molecular weight markers. (B) RNA species predicted from the S1 nuclease analysis are lined up with a restriction map of the *Hin*dIII-*Bam*HI fragment A_{24}.

Computer-aided analysis, performed on the amino acid sequence predicted from the long open reading frame, indicated regions of hydrophilicity (amino acids 65 to 70 and 465 to 470) and of hydrophobicity (around amino acids 240, 290, and 315) (Fig. 6).

Location of the 5′ end of the mRNA. Nuclease protection analysis was performed with a DNA fragment asymmetrically labeled at an *Sal*I site downstream of the second ATG. The major nuclease S1-protected band corresponded to an RNA start site at nucleotide position 227, which coincides with the A residue of the first ATG. A minor S1 nuclease-protected band was located about 14 nucleotides upstream of the ATG (Fig. 7). There was no evidence of an RNA start located downstream of the first ATG (data not shown). Parallel experiments with mung bean nuclease indicated minor and major RNA start sites 3 and 7 nucleotides, respectively, before the first ATG (Fig. 7). In interpreting

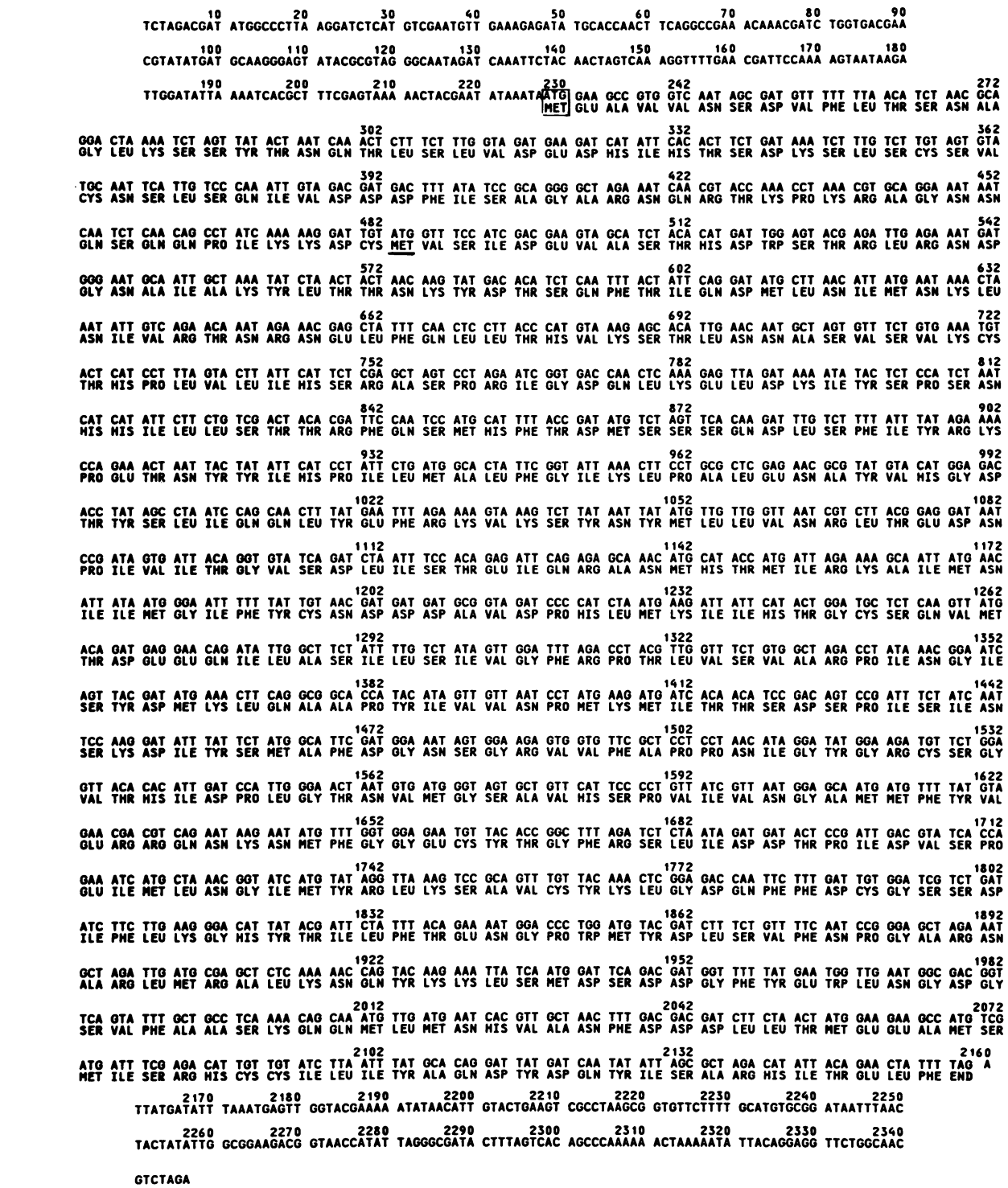

FIG. 5. Nucleotide and derived amino acid sequences of the *Xba*I fragment containing the p4b gene. The open reading frame starts at nucleotide 227 and ends at nucleotide 2161. A second in-frame ATG is located at position 485.

both the nuclease S1 and mung bean nuclease experiments, an additional correction of 1 to 1.5 bases has to be made to account for the elimination of the modified terminal nucleotide during chemical sequencing (13). The small discrepancy between the results obtained with the two nucleases may reflect differences in nibbling of the extremely A+T-rich sequence. Taken together, the data suggest that the 5′ end of the mRNA encoding P4b maps just upstream of the first ATG of the open reading frame.

DISCUSSION

In this communication we describe the mapping, nucleotide sequence, putative RNA start site, and predicted amino

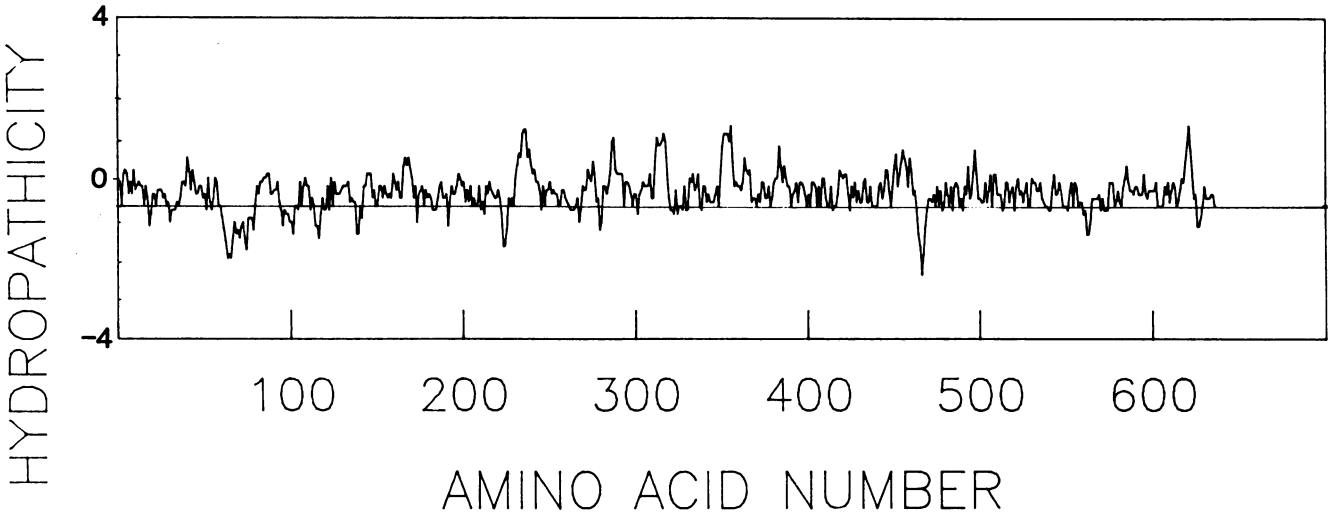

FIG. 6. Computer-aided analysis of the predicted amino acid sequence. The hydropathicity profile of the P4b gene product was determined by using a computer program described by Kyte and Doolittle (21). Positive and negative values $\times 10^{-1}$ represent hydrophobicity and hydrophilicity, respectively.

acid sequence of a gene encoding the precursor to polypeptide 4b, a major core polypeptide. To date, only two other vaccinia virus late genes have been similarly analyzed (4, 36).

Antibodies prepared against polypeptide 4b immunoprecipitated a larger polypeptide designated P4b. The latter was previously shown to be the precursor of 4b by pulse-chase experiments and tryptic peptide analysis (19, 29). P4b was synthesized in a micrococcal nuclease-treated reticulocyte cell-free extract that was programmed with cytoplasmic RNA obtained at late times after vaccinia virus infection. Evidence that the polypeptide was virus coded was obtained by hybridization of the mRNA to vaccinia virus DNA before translation. This hybridization selection, cell-free translation procedure was used to map the P4b gene within the left portion of the *Hin*dIII A fragment, in complete agreement with recent data of Wittek et al. (40).

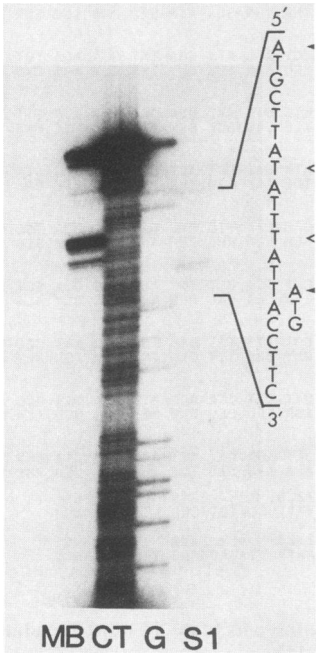

FIG. 7. Determination of the 5′ end of P4b mRNA by nuclease protection. Late RNA was hybridized to a restriction fragment that was 5′ end labeled at a *Sal*I site downstream of the RNA start site and treated with S1 or mung bean (MB) nuclease. Maxam-Gilbert sequence reactions C+T (CT) and G (G) were performed on the same labeled DNA fragment. The positions of S1 nuclease-protected (◀) and mung bean nuclease-protected (<) bands are shown next to the template strand. The first ATG of the open coding frame is indicated.

Hybridization selection experiments indicated that the P4b message extended through the *Hin*dIII-D-A junction into the *Hin*dIII D fragment. Since nuclease protection experiments suggested that at least three late RNA start sites occur downstream of the P4b coding sequence, it seems likely that the P4b message runs over other late genes. In addition, a late RNA starting upstream of the P4b gene may run over P4b. A similar failure of other vaccinia virus RNAs to terminate discretely was described previously and appears to be a general property of late transcripts (11, 23, 36).

Nucleotide sequence analysis indicated the presence of a long open reading frame beginning with ATG that can code for a 73K polypeptide that would have 644 amino acids. An in-phase TAA stop codon occurs just three nucleotides before the first ATG. If the second in-frame ATG represented the initiation codon, then the polypeptide would be 63K. Estimates of P4b have varied from 68K to 74K (29, 40), suggesting that the longer open reading frame is correct. The sequence around the first ATG also is similar to that of other late genes that we have examined in having the triplet TAA immediately preceding it (J. Rosel, J. P. Weir, and B. Moss, manuscript in preparation), instead of the eucaryotic consensus sequences proposed by Kozak (20).

The presence of overlapping late transcripts causes some uncertainty in mapping mRNAs. At present, we assume that the mRNA for a particular gene is the one that starts just upstream of the coding sequence. Our nuclease protection experiments indicate that the 5′ end of the mRNA mapping 3.5 kb to the right of the *Hin*dIII site in Fig. 4 is very close to the first ATG of the long open reading frame discussed

above. The next potential major RNA start site, which Wittek et al. (40) considered to correspond to the P4b message, maps about 900 nucleotides further upstream.

High-resolution nuclease S1 analysis under conditions intended to minimize nibbling (0.28 M NaCl; 25°C) revealed major and minor protected bands at the A of the first ATG and 14 nucleotides above it, respectively. With mung bean nuclease, which reportedly has a lower tendency to nibble than S1 nuclease (13), minor and major protected bands were 3 and 7 nucleotides, respectively, above the first ATG. Thus far, we have not succeeded in mapping the RNA start site by primer extension with oligonucleotides, possibly because of overlapping and complementary late transcripts.

Nucleotide sequences are now available for two genes that encode polypeptides that are proteolytically processed to form core proteins. By determining the terminal amino acid sequences of the mature proteins, it should be possible to locate the putative cleavage sites.

ACKNOWLEDGMENTS

We thank E. V. Jones for cosmids, N. Cooper for purified vaccinia virus, J. Weir for helpful advice, G. Rohrmann and K. Brechling for critically reading the manuscript, and M. Rust for editorial assistance.

J.R. was supported by a Deutsche Forschungs Gemeinschaft grant.

LITERATURE CITED

1. **Belle Isle, H., S. Venkatesan, and B. Moss.** 1981. Cell-free translation of early and late mRNAs selected by hybridization to cloned DNA fragments from the left 14 million to 72 million daltons of the vaccinia virus genome. Virology **112**:306–317.
2. **Benson, S. A.** 1984. A rapid procedure for isolation of DNA fragments from agarose gels. BioTechniques **2**:66–68.
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. **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.
5. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
6. **Blomquist, M. C., L. T. Hunt, and W. C. Barker.** 1984. Vaccinia virus 19 kilodalton protein: relationship to several mammalian proteins, including two growth factors. Proc. Natl. Acad. Sci. USA **81**:7363–7367.
7. **Brown, J. P., D. R. Twardzik, H. Marquardt, and G. J. Todaro.** 1985. Vaccinia virus encodes a polypeptide homologous to epidermal growth factor and transforming growth factor. Nature (London) **313**:491–492.
8. **Cochran, M. A., M. Mackett, and B. Moss.** 1985. Eucaryotic transient expression system dependent on transcription factors and regulatory DNA sequences of vaccinia virus. Proc. Natl. Acad. Sci. USA **82**:19–23.
9. **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.
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. **Glisin, V., R. Crkvenjakow, and C. Buys.** 1974. Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry **13**:2633–2637.
13. **Green, M. R., and R. G. Roeder.** 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. Cell **22**:231–242.
14. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoint for DNA sequencing. Gene **28**:351–359.
15. **Hruby, D. E., and L. A. Ball.** 1982. Mapping and identification of vaccinia virus thymidine kinase gene. J. Virol. **43**:403–409.
16. **Hruby, D. E., R. A. Maki, D. B. Miller, and L. A. Ball.** 1983. Fine structure analysis and nucleotide sequences of the vaccinia virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **80**:3411–3415.
17. **Jackson, R. J., and T. Hunt.** 1983. Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eucaryotic messenger RNA. Methods Enzymol. **96**:50–74.
18. **Jones, E. V., and B. Moss.** 1984. Mapping of the vaccinia virus DNA polymerase gene by marker rescue and cell-free translation of selected RNA. J. Virol. **49**:72–77.
19. **Katz, E., and B. Moss.** 1970. Formation of a vaccinia virus structural polypeptide from a higher molecular weight precursor: inhibition by rifampicin. Proc. Natl. Acad. Sci. USA **66**:677–684.
20. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
21. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **159**:105–132.
22. **Mackett, M., G. L. Smith, and B. Moss.** 1984. A general method for the production and selection of infectious vaccinia virus recombinants expressing foreign genes. J. Virol. **49**:857–864.
23. **Mahr, A., and B. E. Roberts.** 1984. Arrangement of late RNAs transcribed from a 7.1-kilobase *Eco*RI vaccinia DNA fragment. J. Virol. **49**:510–520.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.
25. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
26. **Messing, J., and J. Viera.** 1982. A new pair of M13 vectors for selecting either strand of double digest restriction fragments. Gene **19**:269–276.
27. **Morgan, J. R., L. K. Cohen, and B. Roberts.** 1984. Identification of the DNA sequence encoding the large subunit of the mRNA-capping enzyme of vaccinia virus. J. Virol. **52**:206–214.
28. **Morgan, J. R., and B. E. Roberts.** 1984. Organization of RNA transcripts from a vaccinia virus early gene cluster. J. Virol. **51**:283–297.
29. **Moss, B., and E. N. Rosenblum.** 1973. Protein cleavage and poxvirus morphogenesis: tryptic peptide analysis of core precursors accumulated by blocking assembly with rifampicin. J. Mol. Biol. **81**:267–269.
30. **Paoletti, E., H. Rosemond-Hornbeak, and B. Moss.** 1974. Two nucleic acid-dependent nucleoside triphosphate phosphohydrolases from vaccinia virus: purification and characterization. J. Biol. Chem. **249**:3273–3280.
31. **Queen, C., and L. J. Korn.** 1980. Computer analysis of nucleic acids and proteins. Methods Enzymol. **65**:595–609.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
33. **Venkatesan, S., A. Gershowitz, and B. Moss.** 1982. Complete nucleotide sequences of two adjacent early vaccinia virus genes located within the inverted terminal repetition. J. Virol. **44**:637–646.
34. **Weir, J. P., G. Bajszar, and B. Moss.** 1982. Mapping of the vaccinia virus thymidine kinase gene by marker rescue and by cell-free translation of selected mRNA. Proc. Natl. Acad. Sci. USA **79**:1210–1214.
35. **Weir, J. P., and B. Moss.** 1983. Nucleotide sequence of the vaccinia virus thymidine kinase gene and the nature of spontaneous frameshift mutations. J. Virol. **46**:530–537.
36. **Weir, J. P., and B. Moss.** 1984. Regulation of expression and nucleotide sequence of a late vaccinia virus gene. J. Virol.

51:662–669.

37. **Winberg, G., and M.-L. Hammarskjold.** 1980. Isolation of DNA from agarose gels using DEAE paper. Application to restriction site mapping of adenovirus type 16 DNA. Nucleic Acids Res. **8**:253–264.
38. **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.
39. **Wittek, R., M. Hänggi, and G. Hiller.** 1984. Mapping of a gene coding for a major late structural polypeptide on the vaccinia virus genome. J. Virol. **49**:371–378.
40. **Wittek, R., B. Richner, and G. Hiller.** 1984. Mapping of the genes coding for the two major vaccinia virus core polypeptides. Nucleic Acids Res. **12**:4835–4848.
41. **Yang, R. C.-A., J. Lis, and R. Wu.** 1979. Elution of DNA from agarose gels after electrophoreses. Methods Enzymol. **68**: 176–182.