JOURNAL OF VIROLOGY, Feb. 1987, p. 446–453
0022-538X/87/020446-08$02.00/0

Vol. 61, No. 2

Primary Structure and Transcription of the Genes Coding for the Two Virion Phosphoproteins pp65 and pp71 of Human Cytomegalovirus

BARBARA RÜGER,[1]* SABINE KLAGES,[1] BIRGITT WALLA,[1] JENS ALBRECHT,[1] BERNHARD FLECKENSTEIN,[1] PETER TOMLINSON,[2] AND BART BARRELL[2]

Institut für Klinische Virologie der Universität Erlangen-Nürnberg, D-8520 Erlangen, Federal Republic of Germany,[1] and Medical Research Council Laboratory of Molecular Biology, GB-Cambridge CB2 2QH, United Kingdom[2]

Received 23 June 1986/Accepted 28 October 1986

Human cytomegalovirus contains a phosphorylated matrix protein of 65,000 apparent molecular weight (65K phosphoprotein; pp65) and a related phosphoprotein of 71,0000 molecular weight (pp71). The 65K phosphoprotein is usually by far the most abundant structural component found in culture-grown purified virus particles. This study describes the precise mapping of the genes for both polypeptides, giving the entire nucleotide sequences and the exact positions of the respective transcripts. The 65K phosphoprotein is coded for by the 5′-terminal part of an abundant 4-kilobase (kb) mRNA. The 71K phosphoprotein corresponds to the single translational reading frame of a rare nonspliced 1.9-kb mRNA that is coterminal with the 4-kb transcript. The promoter for 4-kb mRNA appears to be unusual in structure; it does not contain a characteristic TATA sequence. The expression of antigenic epitopes from pp65 may allow improved serodiagnosis of human cytomegalovirus infections.

Human cytomegalovirus (HCMV) is an ubiquitous pathogenic herpesvirus causing a variety of disease conditions after prenatal infection or in patients with suppressed cellular immune functions. Problems are aggravated by the lack of specific chemotherapy (13), effective ways of passive immunization (1), or generally applicable vaccination schemes. To improve serodiagnostic procedures, the expression cloning of viral structural protein genes is required. HCMV codes for approximately 30 structural polypeptides; only a few of the respective viral genes have been mapped on the genome up to now. Cloning of late viral cDNA in the lambda gt11 system and screening of the expression library with monospecific polyvalent sera or monoclonal antibodies allowed identification of the coding sequences for the predominant envelope glycoprotein (gp58) (15); a second, smaller surface protein (p28), which is possibly glycosylated (M. Mach, H. Meyer, P. Landini, and B. Fleckenstein, unpublished data); and the highly immunogenic matrix phosphoprotein of about 150,000 apparent molecular weight (150K phosphoprotein) (pp150) (G. Jahn, M. Mach, B.-C. Scholl, B. Plachter, B. Traupe, B. Fleckenstein, T. Kouzarides, K. Weston, A. T. Bankier, and B. G. Barrell, manuscript submitted for publication).

Purified HCMV particles contain a very abundant matrix protein of about 64,000 to 69,000 apparent molecular weight (9, 11) which was shown to be modified by phosphorylation (26). This polypeptide had been designated in various reports as pp65 (21), 69,000-molecular-weight matrixlike protein (11), lower matrix protein (26), HCMV gp64 (4), or ICP27 (8). The structural protein pp65 forms about 95% of the protein mass in dense bodies, a type of defective virus particle usually found in HCMV-infected cell cultures. The viral gene coding for the matrix protein pp65 had been mapped by hybrid-selected in vitro translation (21) and by Southern blot hybridizations with synthetic oligonucleotides that were derived from a known amino acid sequence (22). The DNA coding for pp65 is contained within *Hin*dIII fragments L, c, and b of the HCMV AD 169 genome. A second protein of the inner envelope with about 71,000 molecular weight (pp71), sometimes also referred to as 74-kilodalton upper matrix protein (26), has also been shown to be phosphorylated (26; B. Nowak, Ph.D. thesis, Universität Erlangen-Nürnberg, Erlangen, Federal Republic of Germany, 1984); these two polypeptides, pp65 and pp71, appeared to be coded for by the same genomic region (21). The messenger for both proteins is transcribed from the same DNA strand, and a single abundant late 4-kilobase (kb) mRNA was correlated with the respective genomic region in the appropriate orientation. The relationship between the two phosphoproteins and their respective genomic organizations appeared to be of particular interest, as various transcriptional patterns or many ways of posttranscriptional and posttranslational processing could account for two phosphoproteins of widely different abundance from one distinct DNA locus. This study describes the nucleotide sequence of the entire coding domain for pp65 and pp71, identifies separate translational reading frames and transcripts for each protein, and analyzes their RNA structures.

MATERIALS AND METHODS

Bacteria and virus. The following *Escherichia coli* strains were used: JM83 [=F$^-$ *ara* (*lac-proAB*) *strA 80lacZ* M15 (19)], JM105 [=*thi strA endA sbc-15 hsdR4* (*lac-proAB*) F′ *traD36 proAB lacI^{q}Z* M15 (34)], HB101 [=*thi thr leu recA strA hsdR hsdM* (strain collection, Dept. of Microbiology, Universität Erlangen-Nürnberg)], and LE392 [=F$^-$ *supF supE hsdR galK trpR metB lacY tonA* (gift from Lynn Enquist)]. The following cloning vectors were used: pHC79 (7), pUC8 and pUC18 (32, 34), pSP64 and pSP65 (18), pGem1 and pGem2 (Promega-Biotec, CH-Geneva, Switzerland), and M13mp18 and M13mp19 (20). HCMV AD 169 was provided by U. Krech, CH-Sankt Gallen, Switzerland; the virus was grown on human foreskin fibroblasts by standard

* Corresponding author.

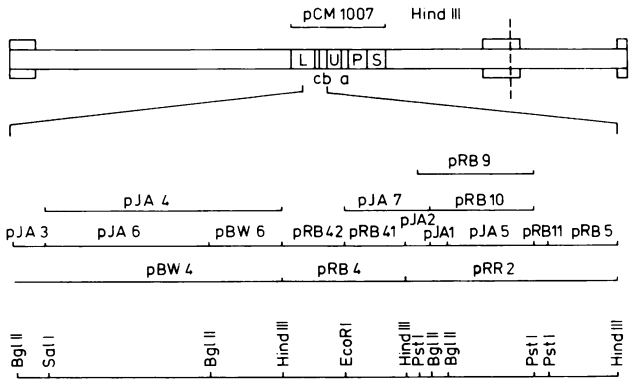

FIG. 1. Genomic localization of the bacteriophage and plasmid subclones from the coding region for phosphoproteins pp65 and pp71. The entire size of the DNA region corresponds to 4.1 kb. The plasmid vector pGem2 was used for clones pBW4, pJA4, pJA5, pJA6, pJA7, pRB9, pRB10, pRB11, pRB41, pRB42, and pRR2. pUC8, or pSP64 and pSP65, vector was used for pBW6, pRB4, and pRB5. The clones pJA1, pJA2, and pJA3 are M13mp18 derivatives.

methods. Virions were purified by centrifugation through a glycerol-tartrate gradient (31).

Plasmid and bacteriophage cloning. Fragments of viral DNA from cosmid clones were subcloned in plasmid vectors by standard procedures; cloning into M13 vectors was done according to a published protocol (19). M13mp18 and M13mp19 were propagated on *E. coli* JM105, pUC clones were propagated on strain JM83, pSP64, pSP65, pGem1 and pGem2 were propagated on strain LE392, and cosmid clones were propagated on *E. coli* HB101.

Isolation of DNA and sequencing. Plasmids and M13 replicative form DNA were purified by a cleared-lysate method (19). Single-stranded bacteriophage DNA for sequencing was also prepared according to a published protocol (19). The nucleotide sequence was determined by chemical degradation (17) or by the dideoxy chain termination method (27) by using bacteriophage M13 clones (32). Sequence fragments were aligned by using a computer program (29, 30).

RNA preparation. Fibroblast cultures were infected with HCMV AD 169; after about 7 days, the cells were removed with trypsin and lysed with 5.8 M guanidinium rhodanide containing 50 mM lithium citrate, 0.1 M β-mercaptoethanol, and 0.1% lithium laurylsulfate. The RNA was separated from DNA and proteins by centrifugation through a CsCl gradient (5.7 M CsCl, 0.1 M EDTA [pH 7.0]) in a Spinco SW41 rotor for 24 h at 28,000 rpm and 20°C. The RNA pellet was suspended in water (treated with diethyl pyrocarbonate) and precipitated with ethanol–0.4 M LiCl.

Northern blot hybridization. Gel electrophoresis of RNA after denaturing by glyoxal treatment was performed on 1% agarose–MOPS (3-N-morpholino propanesulfonic acid) gels containing 2.2 M formaldehyde. RNA was transferred to nitrocellulose filters by the standard method (12, 28). The RNA was hybridized with nick-repair-labeled plasmid DNA (25), labeled second strands of M13 clones, or cRNA generated from SP6 and T7 promoters of the corresponding clones. The hybridization was carried out in 5 × SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and in 50% (vol/vol) formamide at 46°C; those conditions are stringent, calculated based on an average guanine-plus-cytosine content of 65% for the respective fragments of HCMV DNA.

Hybrid-selected in vitro translation. Samples (100 μg each) of denatured cloned DNA were spotted on 10 nitrocellulose filters (5 by 5 mm in size) and vacuum dried for 2 h. After two washings in 6 × SSC, the filters were vacuum dried again and baked at 80°C for 4 h. Hybridization with 100 μg of total RNA from HCMV-infected human foreskin fibroblast cells was performed in 65% formamide–10 mM PIPES (piperazine-*N*,*N*-bis(2-ethanesulfonic acid) [pH 6.8])–400 mM NaCl–0.2% of polyvinyl sulfate at 58°C for 3 h. After hybridization, the filters were washed 10 times in 1 ml of 1× SSC with 0.5% sodium dodecyl sulfate and rinsed 3 times in 1 ml of 2 mM EDTA (pH 7.0) at 58°C. Bound RNA was eluted in 300 μl of distilled H_2O for 5 min at 100°C and ethanol precipitated with calf liver tRNA and 300 mM potassium acetate. The eluted RNA was translated in vitro by using rabbit reticulocyte lysate (New England Nuclear Corp., Boston, Mass.) according to the instructions of the supplier. In vitro translation products were fractionated in 15% sodium dodecyl sulfate-polyacrylamide gels by standard protocols.

Nuclease protection analyses. For S1 nuclease mapping experiments, 2 to 3 μg of end-labeled (16) DNA fragments was precipitated with 100 μg of total RNA from HCMV-infected human foreskin fibroblast cells, suspended in 20 μl of S1 hybridization buffer (80% formamide, 400 mM NaCl, 1 mM EDTA, 40 mM PIPES [pH 6.4]), treated for 15 min at 75°C to resolve secondary structures, and hybridized for 3 h at 58°C. Subsequently, 200 μl of S1 digestion buffer (30 mM sodium acetate [pH 7.0], 50 mM NaCl, 1 mM $ZnSO_4$, 50 mg of denatured calf thymus DNA per ml, 5% glycerol) and 1,000 U of S1 nuclease (Bethesda Research Laboratories, Inc., Rockville, Md.) were added, and S1 nuclease digestion was performed for 30 min at 37°C. After ethanol precipitation, protected fragments were sized on a 6% polyacrylamide sequencing gel. For RNase protection assays, DNA fragments were cloned into pSP64 and pSP65 or into pGem1 and pGem2 (Promega-Biotec). cRNA synthesis and RNase protection were performed according to the instructions of the supplier.

RESULTS

Subclones of the genomic region. The entire region of the HCMV genome which codes for the phosphoproteins pp65 and pp71 is transcribed into a late 4-kb mRNA (21); it is contained within *Hin*dIII fragments b and c and about 2 kb from the right part of *Hin*dIII-L (Fig. 1). The small genuine *Hin*dIII fragments were subcloned in the plasmid vector pGem2 or pUC8 starting from the cosmid clone pCM1007 (7). The rightmost one-third of *Hin*dIII-L was subcloned as a *Hin*dIII-*Bam*HI fragment of 2.4 kb into the vector pGem2, also starting from cosmid pCM1007 (7). Additional subclones were also constructed in plasmid and bacteriophage vectors (Fig. 1).

Nucleotide sequence. The entire 4.1-kb segment shown in Fig. 1 was sequenced by the dideoxy chain termination method (27) and, in part, by chemical degradation (17). The nucleotide sequence of the noncoding strand is shown in Fig. 2. Computer analysis revealed two long open translational reading frames in the direction of transcription (Fig. 3), whereas the opposite strand did not have sizable reading frames. The first coding frame did span between nucleotides 341 and 2106, and the second reading frame was located between nucleotides 2213 and 3992 (Fig. 2). Two TATA consensus sequences were found between the two translational reading frames (nucleotides 2058 to 2061 and 2117 to 2121) (Fig. 2). However, no canonical TATA sequence

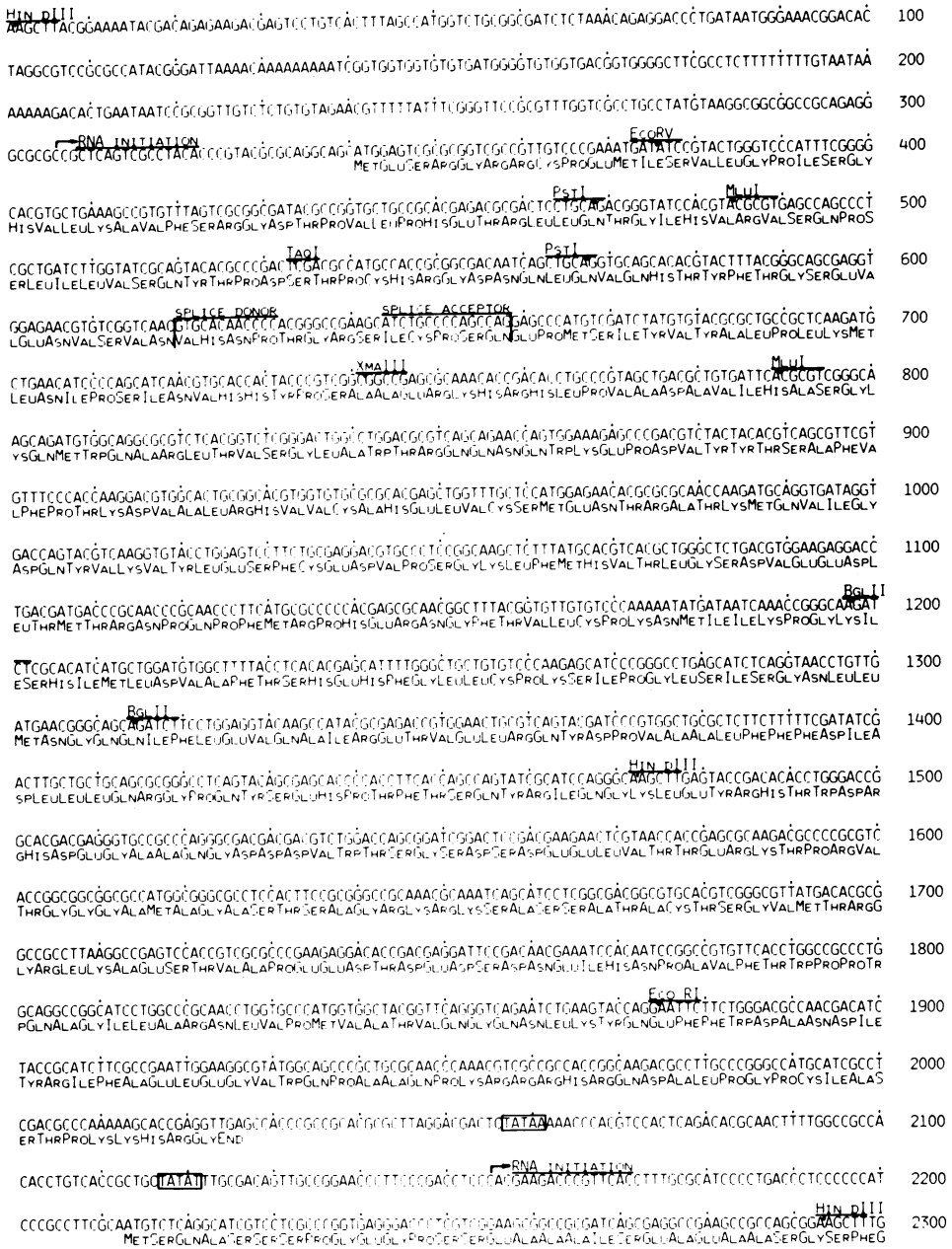

FIG. 2. Nucleotide sequence of the entire genomic region coding for the phosphoproteins pp65 and pp71. The derived amino acid sequences are indicated for the two long translational reading frames. TATA consensus sequences and the single polyadenylation signal, AATAAA, are boxed. All restriction sites used for subcloning (Fig. 1) are indicated.

became apparent upstream of the first open reading frame. A single polyadenylation signal, AATAAA, was found downstream from the reading frames, at nucleotides 3954 to 3959 (Fig. 2). The computer search did not reveal nucleotide or amino acid sequence homologies with any of the entries in the actual National Institutes of Health Genebank, the European Molecular Biology Laboratory gene library, or the National Biomedical Research Foundation protein data bank. No significant homologies became apparent with the entire Epstein-Barr virus genome; there were also no sequence similarities to the structural phosphoprotein of about 65,000 molecular weight that had been identified in herpes simplex virus as a *trans*-activator of immediate-early transcription (3).

Coding mRNAs. Previous Northern blot studies had shown that there is one abundant late 4-kb mRNA from the entire transcription unit (21). This genomic region comprises both open reading frames, and each of these translational frames could possibly code for either of the two phosphoproteins, pp65 and pp71. In order to correlate transcripts, reading frames, and proteins, a series of hybrid-selected in vitro translations were conducted by using small DNA subclones of the coding region (Fig. 1). When late mRNA was selected with the clones pBW4 and pBW6

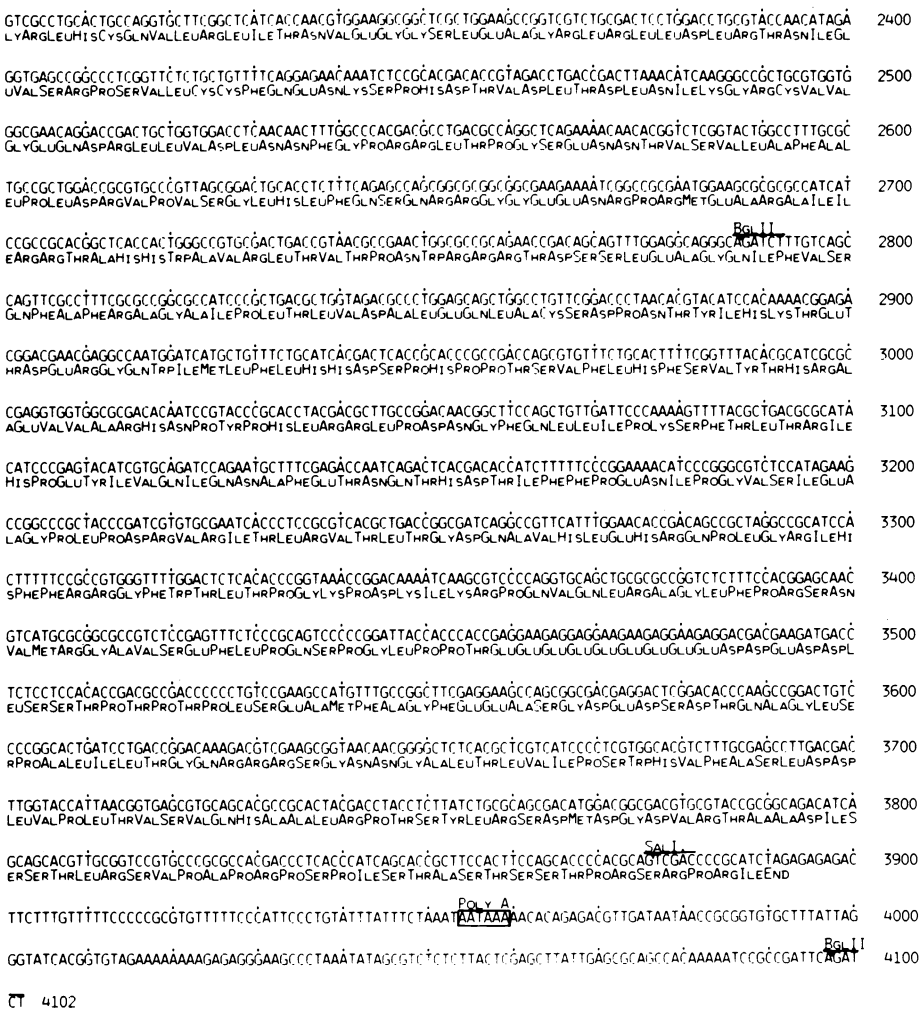

FIG. 2—*Continued.*

(subclones of *Hin*dIII-L) and translated in vitro with a rabbit reticulocyte lysate, two polypeptides of 65,000 and 71,000 molecular weight became apparent (Fig. 4A). However, the same assay, using plasmid pRR2 (*Hin*dIII-b), yielded the 65K polypeptide (Fig. 4A) and, if any, very low amounts of the 71K protein. Adjacent plasmid clones outside the sequenced 4.1-kb region (e.g., pRB7, pBW7, pBW8, pBW9, and pBW10) did not select mRNA coding for either of the two proteins (Fig. 4A). These experiments seemed to indicate first that a more dominant mRNA, presumably the 4-kb transcript, codes for the pp65, whereas pp71 is translated from a second, smaller overlapping mRNA of lower abundance. Subsequent hybrid-selected in vitro translations using additional small subclones showed that fragments from the 5′ part of the coding region selected mainly mRNA coding for the 65K protein, whereas the clones from the 3′ segment always selected the messages coding for both proteins (Fig. 4B). Very faint bands at the 71,000-molecular-weight position were sometimes seen after hybrid selection with 5′-oriented subclones; these bands may be due to selection background or, less likely, internal translation initiation from fragmented mRNA. In vitro translation from RNA selected by the clones pRB42 and pBW6 yielded another polypeptide of about 76,000 molecular weight; this may hint at additional late viral transcripts, though structural virion proteins of this size have not been identified so far. Altogether, these facts indicated that the first (5′) open reading frame codes for the 65-K protein and that the second (3′) frame codes for the 71K polypeptide. Thus, the 4-kb mRNA appears to code for pp65 and, in addition, contains all of the information for pp71; the smaller transcript (for pp71 only) seemed to be a 3′ subset of the 4-kb mRNA. This model was confirmed by additional Northern blot hybridizations using nick-repair-labeled DNA probes, radioactive second-strand probes from bacteriophage M13 clones, or cRNA from SP6 and T7 promoters. Labeled probes from the 5′ part of the 4.1-kb segment hybridized only with the 4-kb transcript; 3′ probes, such as pJA3, pJA4, and pRB4, recognized the dominant 4-kb mRNA and a minor 1.9-kb transcript. The characteristic Northern blot pattern obtained with the labeled *Bgl*II-*Sal*I fragment of 227 nucleotides in clone pJA3 is shown in Fig. 5.

Transcript mapping. The two transcripts of 4 and 1.9 kb were mapped more precisely by nuclease protection analyses. The 5′ end of the 4-kb transcript was placed at nucleotide 306 by S1 nuclease analysis (Fig. 6). The same nucleotide was identified as putative transcription initiation site when RNase protection experiments with SP6 transcripts

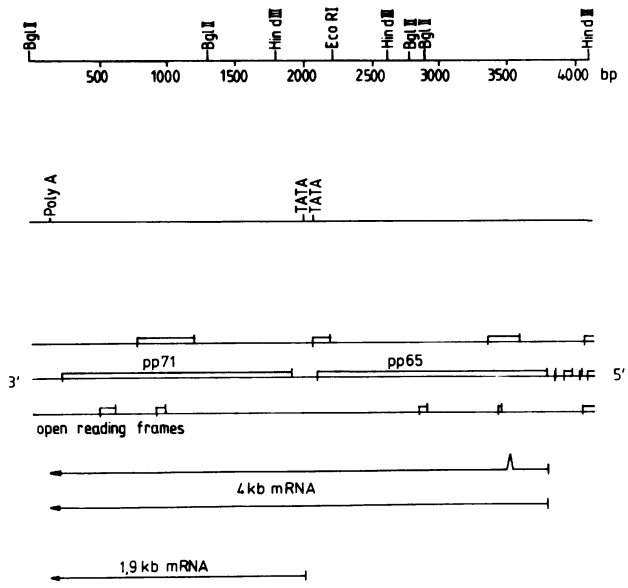

FIG. 3. Summary scheme of the organization of the genes coding for the phosphoproteins pp65 and pp71, including the relative orientations of the translational reading frames for the two polypeptides, the promoter consensus sequences, and the single polyadenylation site. Sizes, directions, and map positions of the coding RNAs are symbolized by arrows.

were conducted. Although it seemed remarkable that no TATA consensus sequence precedes the apparent cap site, it is very likely that it represents a promoter sequence. No DNA fragments further upstream hybridized with the 4-kb transcript in Northern blots when cloned *Hin*dIII fragments P, S, U, and a were used as radioactive probes. There was also no possible splice acceptor consensus sequence coinciding with the promoter sequence identified by nuclease protection assays. In addition, primer extension experiments using a synthetic oligonucleotide placed the transcription initiation site at the same promoter sequence.

Screening of the entire 4-kb transcription unit by RNase protection with cRNAs from SP6 and T7 promoters indicated that a single short intron was within the first open reading frame. S1 nuclease analyses placed the first spliced nucleotide at position 620 (data not shown). Residue 658 was identified as the last nucleotide of the intron (Fig. 2 and 7). This showed that 39 nucleotides of the first open reading frame in the 4-kb mRNA can be spliced out, leaving the frame in phase. However, sequencing of a cDNA clone covering 400 5'-oriented nucleotides of the 4-kb transcript showed that unspliced transcripts can occur in lytically infected cells. In vitro-synthesized RNA from the clones pRB41, pRB42, pJA2, and pJA was protected against degradation with RNases A and T1 if hybridized with RNA from HCMV-infected cells (data not shown), indicating the absence of additional introns. The 5' end of the 1.9-kb mRNA was found within the subclone pRB42 at nucleotide 2160

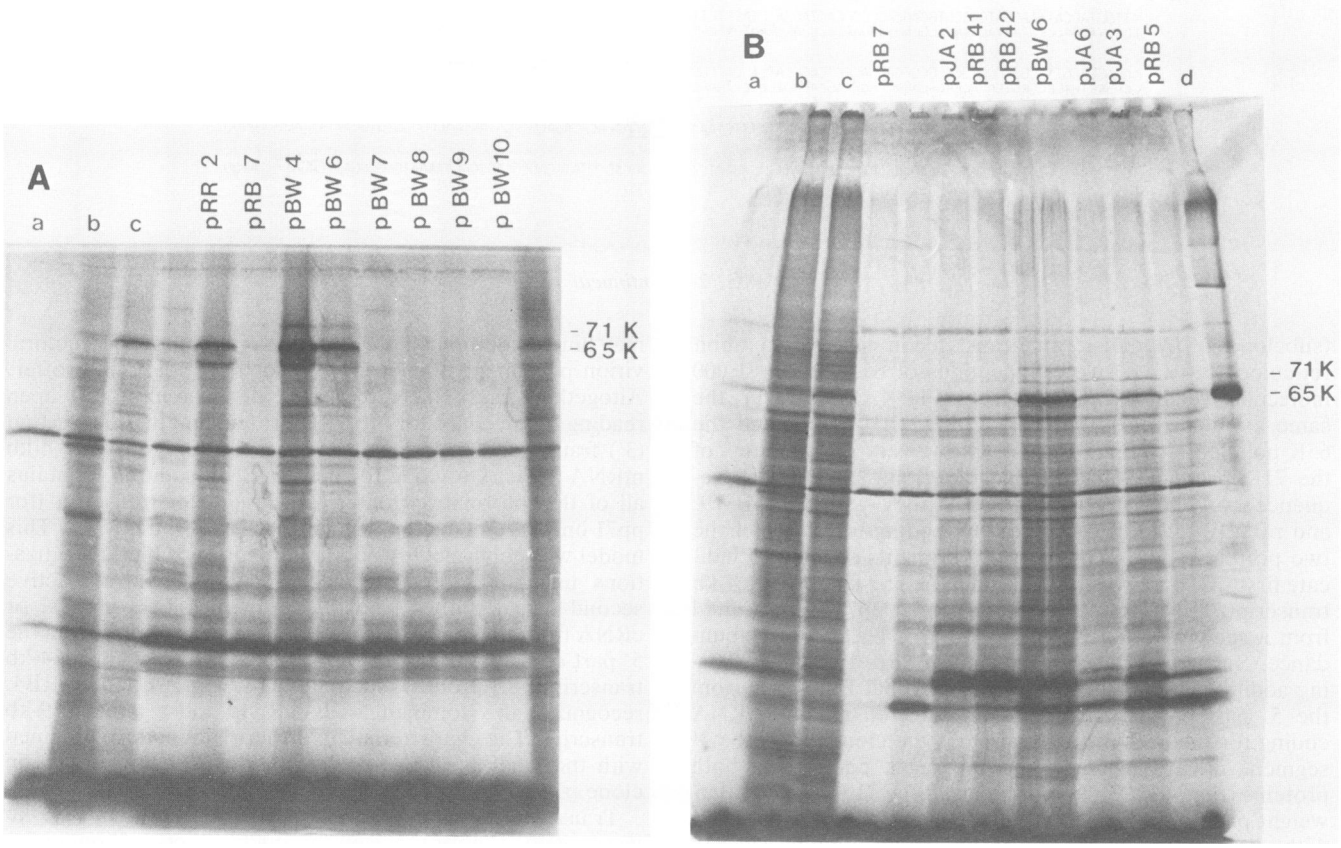

FIG. 4. Hybrid-selected in vitro translation of late viral mRNA by using rabbit reticulocyte lysates. (A) Hybrid-selection done with the plasmid clones pRR2 (*Hin*dIII b fragment), pRB7 (subclone of *Hin*dIII U), and pBW4, pBW6, pBW7, pBW8, pBW9, and pBW10 (subclones of *Hin*dIII L). Lanes a to c show control experiments. Lanes: a, translation products in the absence of exogenous RNA; b, in vitro translation of RNA from mock-infected fibroblasts; c, translation of late RNA of HCMV-infected cultures. (B) Hybrid-selection by a series of small subclones, as described in the legend to Fig. 1. Lanes: a to c, negative and positive controls, as in panel A; d, [^{35}S]methionine-labeled proteins of purified virions.

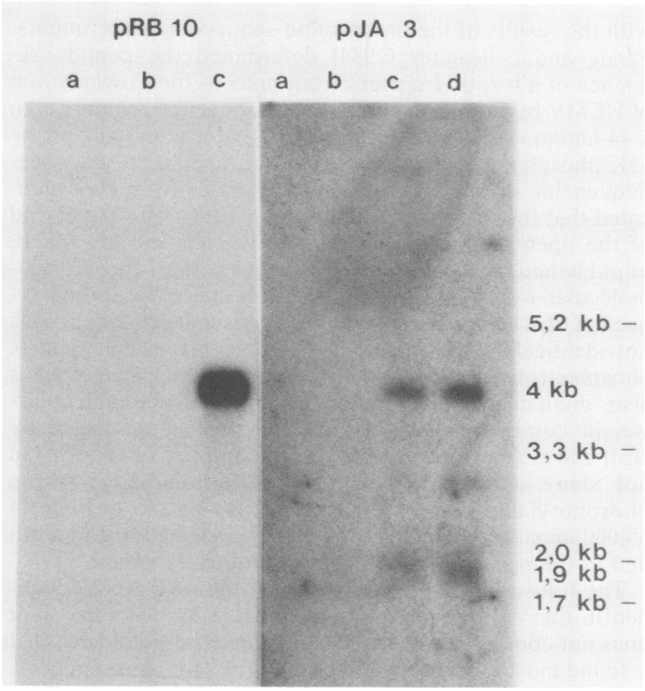

FIG. 5. Northern blot hybridizations of two small DNA subclones with late viral RNA. First panel, hybridization of the nick-repair-labeled plasmid clone pRB10 (Fig. 1) with RNA from mock-infected fibroblasts (lane a), with *E. coli* rRNA (lane b), and with RNA from fibroblasts late in HCMV infection. Second panel, hybridization of labeled second strand from the M13 subclone pJA3 (Fig. 1) with RNA from mock-infected fibroblasts (lane a), with *E. coli* rRNA (lane b) and with independently prepared batches of late HCMV RNA (lanes c and d). This clone regularly detects transcripts of 1.9 and 4.0 kb. Sometimes a very faint band became visible, representing RNA of about 3.0 kb.

(Fig. 2 and 7). This is 35 nucleotides downstream from the TATAT motif described above. Northern blots and hybrid-selected in vitro translations allowed us to conclude that both transcripts use the same single polyadenylation signal. Thus, disregarding poly(A) tails, the large abundant transcript for pp65 has close to 3,770 nucleotides, and the minor mRNA for pp71 has about 1,815 nucleotides. The precise molecular weight of phosphoprotein pp65 was found to be 61,520 (548 amino acids), and that of phosphoprotein pp71 was found to be 62,900 (561 amino acids). The topology of the two transcripts relative to open reading frames is schematically summarized in Fig. 3.

DISCUSSION

Purified HCMV particles contain a highly abundant phosphorylated protein of about 65,000 molecular weight which has been designated pp65 (21), or lower matrix protein (26). The protein is a constituent of the internal virion envelope, and it represents approximately 95% of the protein mass in dense body particles (26). Previous studies had identified the coding region for this phosphoprotein (21, 22) and for a related matrix phosphoprotein of 71,000 molecular weight (pp71, upper matrix protein) within *Hin*dIII fragments b, c, and L of HCMV AD 169. This paper describes the nucleotide sequence of the entire coding region, identifying separate translational reading frames for each of the two phosphoproteins by hybrid-selected in vitro translation. The lower matrix phosphoprotein (548 amino acids, 61,500 precise molecular weight) is coded for by an abundant 4-kb mRNA that can be formed through splicing; the short intervening sequence has 39 nucleotides. The upper matrix phosphoprotein (561 amino acids, 62,900 precise molecular weight) appeared to be coded for by an unspliced rare 1.9-kb mRNA. This model, assigning pp65 to 5′-oriented reading frame and pp71 to the 3′-terminal sequence, should receive additional support from current immunoprecipitation studies using antisera raised against procaryotic fusion proteins that are derived from each of the two open reading frames. The abundant 4-kb mRNA and the 1.9-kb transcripts appeared to be coterminal. This is consistent with the presence of a single polyadenylation signal, AATAAA, within the entire DNA sequence of nucleotides 3954 to 3959. Families of overlapping transcripts with coterminal 3′ ends may be common in herpesviruses (33).

The structure of the 4-kb mRNA implies that it is a bicistronic eucaryotic messenger. Yet it is difficult to decide whether the second reading frame of this mRNA is used to synthesize the pp71 precursor polypeptide in infected cells. Hybrid-selected in vitro translation of the 4-kb mRNA

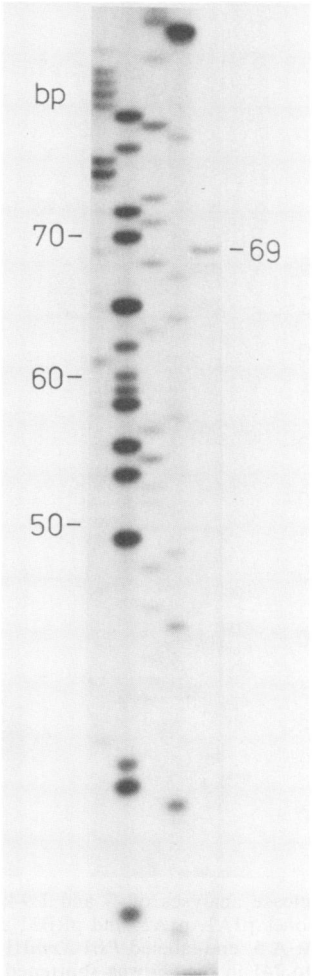

FIG. 6. S1 nuclease protection analysis of the transcription initiation site of the 4-kb mRNA. A *Hin*dIII-*Eco*RV fragment (nucleotides 1 to 375 in Fig. 2) was 5′ end-labeled with polynucleotide kinase and hybridized with HCMV late RNA before S1 nuclease digestion. The shortened product of 69 nucleotides was sized with a known DNA sequence ladder. bp, Base pairs.

resulted essentially in the synthesis of the 65K polypeptide. The single short intron identified by nuclease protection analyses in the 4-kb transcription unit is unusual, as it is located within the coding region of a herpesvirus structural protein. The intron is very small (39 nucleotides); however, it is still compatible with splicing rules (24). On the other hand, sequencing of a cDNA clone indicated that unspliced transcripts do occur in lytically infected cells. The relative proportion of spliced and unspliced transcripts in the pools of precursors and mature mRNA needs to be clarified in future studies.

The 65K phosphoprotein is coded for by the first open reading frame of the 4-kb mRNA. This is in good agreement with the results of the first protein-sequencing experiments. Pande and colleagues (1984) determined the peptide sequence of a tryptic fragment from pp65 of the Towne strain of HCMV by Edman degradation; its sequence is identical to a 44-amino-acid stretch in the carboxy-terminal part of the 65K phosphoprotein derived from our nucleotide sequence. Sequencing of the amino-terminal residues from pp65 indicated that the mature phosphoprotein starts with the 5′ end of the open reading frame (W. Lottspeich and M. Mach, unpublished data). This also proves that the 67,000-molecular-weight tegument phosphoprotein the coding sequence of which was reported by Davis and Huang (5, 6) is not identical to the major 65,000-molecular-weight matrix phosphoprotein described here. Our phosphoprotein pp65 is also distinct from the approximately 65,000-molecular-weight comigrating virion envelope protein described by Britt and Auger (2). The phosphoprotein pp65 of HCMV did not share any appreciable sequence homology with a phosphorylated herpes simplex virus matrix protein of 65,000 apparent molecular weight (23) that had been identified as the *trans*-activator of immediate-early genes (3).

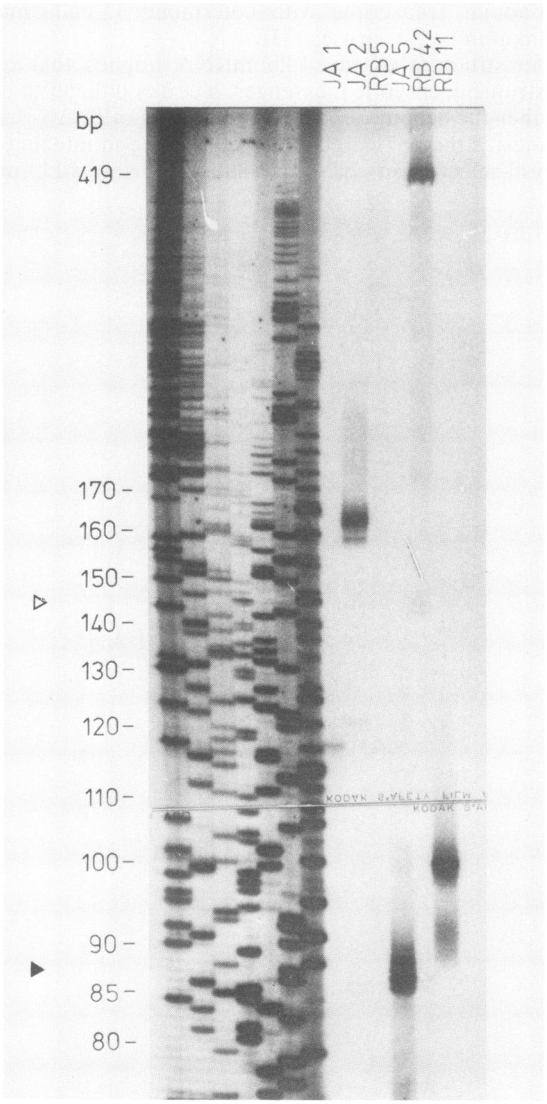

FIG. 7. S1 nuclease analyses of 4- and 1.9-kb mRNA. The 5′ end-labeled subclones pJA1, pJA2, and pRB11 are fully protected by late viral RNA. A 5′ end-labeled *Pst*I-*Xma*III fragment of pJA5 (nucleotides 568 to 741 in Fig. 2) was shortened to 87 nucleotides (splice acceptor). The plasmid subclone pRB42 (Fig. 1) was linearized with *Hin*dIII, 5′ end-labeled, and hybridized with late HCMV RNA before S1 nuclease digestion. Most of the fragment was entirely protected by the 4-kb mRNA; the radioactive probe that hybridized with the 1.9-kb mRNA was shortened to 143 nucleotides. bp, Base pairs.

The transcription initiation site for the 4-kb RNA identified in this study appears to be unusual in structure, as it does not contain the typical TATA consensus sequence that is found mostly in eucaryotic promoter upstream sequences. Nonetheless, there seemed to be little doubt that it represents in fact the functional promoter. The assumed promoter nucleotide sequence does not reveal a splice acceptor consensus motif, and none of the upstream restriction fragments did hybridize with 4-kb RNA in Northern blots. Also, hybrid-selected in vitro translations with cosmid clones covering the entire HCMV genome had not given any indication that the mRNA encoding pp65 has a leader derived from a DNA sequence outside of cosmid pCM1007 (21). Furthermore, transient expression assays with fusion genes that consisted of this promoter sequence up to nucleotide −310 and a reading frame for chloramphenicol acetyl transferase confirmed its nature as a late viral promoter (B. Rüger and R. Rüger, unpublished data). Some culture-grown laboratory strains of HCMV, such as AD 169 and Towne, express the phosphoprotein in excessive amounts; pp65 is apparently the major constituent of dense body particles. On the other hand, recent clinical isolates of the virus sometimes contain very small amounts of this protein (G. Jahn, B.-C. Scholl, B. Traupe, and B. Fleckenstein, manuscript submitted for publication), and sometimes antibody titers in sera from patients are low (14). An unusual type of transcriptional regulation for the 4-kb mRNA could explain the variable expression of pp65 late in HCMV replication. The absence of a TATA sequence near the transcription initiation site may be functionally related to a late promoter that escapes the normally stringent regulation under certain growth conditions.

ACKNOWLEDGMENTS

This work was supported by Wilhelm-Sander-Stiftung and, in part, by Deutsche Forschungsgemeinschaft.

We thank S. Staprans and T. Mettenleiter for technical advice.

LITERATURE CITED

1. **Bowden, R. A., M. Sayers, N. Flournoy, B. Newton, M. Banaji, D. Thomas, and J. D. Meyers.** 1986. Cytomegalovirus immune globulin and seronegative blood products to prevent primary cytomegalovirus infection after marrow transplantation. N. Engl. J. Med. **314**:1006–1010.
2. **Britt, W. J., and D. Auger.** 1985. Identification of a 65,000

dalton virion envelope protein of human cytomegalovirus. Virus Res. **4**:31–36.
3. **Bzik, D. J., and C. M. Preston.** 1986. Analysis of DNA sequences which regulate the transcription of herpes simplex virus immediate early gene 3: DNA sequences required for enhancer-like activity and response to trans-activation by a virion polypeptide. Nucleic Acids. Res. **14**:929–943.
4. **Clark, B. R., J. A. Zaia, L. Balce-Directo, and Y.-P. Ting.** 1984. Isolation and partial chemical characterization of a 64,000-dalton glycoprotein of human cytomegalovirus. J. Virol. **49**: 279–282.
5. **Davis, M. G., and E.-S. Huang.** 1985. Nucleotide sequence of a human cytomegalovirus DNA fragment encoding a 67-kilodalton phosphorylated viral protein. J. Virol. **56**:7–11.
6. **Davis, M. G., E.-C. Mar, Y.-M. Wu, and E.-S. Huang.** 1984. Mapping and expression of a human cytomegalovirus major viral protein. J. Virol. **52**:129–135.
7. **Fleckenstein, B., I. Müller, and J. Collins.** 1982. Cloning of the complete human cytomegalovirus genome in cosmids. Gene **18**:39–46.
8. **Geballe, A. P., F. S. Leach, and E. S. Mocarski.** 1986. Regulation of cytomegalovirus late gene expression: γ genes are controlled by posttranscriptional events. J. Virol. **57**:864–874.
9. **Gibson, W.** 1981. Structural and nonstructural proteins of strain Colburn cytomegalovirus. Virology **3**:516–537.
10. **Gibson, W.** 1983. Protein counterparts of human and simian cytomegaloviruses. Virology **128**:391–406.
11. **Irmiere, A., and W. Gibson.** 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. Virology **130**:118–133.
12. **Jahn, G., E. Knust, H. Schmolla, T. Sarre, J. A. Nelson, J. K. McDougall, and B. Fleckenstein.** 1984. Predominant immediate-early transcripts of human cytomegalovirus AD 169. J. Virol. **49**:363–370.
13. **Koretz, S. H., and the Collaborative DHPG Treatment Study Group.** 1986. Treatment of serious cytomegalovirus infections with 9-(1,3-dihydroxy-2-propoxymethyl)guanine in patients with AIDS and other immunodeficiencies. N. Engl. J. Med. **314**: 801–805.
14. **Landini, M. P., G. Mirolo, B. Baldassarri, and M. LaPlaca.** 1985. Human immune response to cytomegalovirus structural polypeptides studied by immunoblotting. J. Med. Virol. **17**: 303–311.
15. **Mach, M., U. Utz, and B. Fleckenstein.** 1986. Mapping of the major glycoprotein gene of human cytomegalovirus. J. Gen. Virol. **67**:1461–1467.
16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
18. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
19. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
20. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
21. **Nowak, B., A. Gmeiner, P. Sarnow, A. J. Levine, and B. Fleckenstein.** 1984. Physical mapping of human cytomegalovirus genes: identification of DNA sequences coding for a virion phosphoprotein of 71 kDa and a viral 65-kDa polypeptide. Virology **134**:91–102.
22. **Pande, H., S. W. Baak, A. D. Riggs, B. R. Clark, J. E. Shively, and J. A. Zaia.** 1984. Cloning and physical mapping of a gene fragment coding for a 64-kilodalton major late antigen of human cytomegalovirus. Proc. Natl. Acad. Sci. USA **81**:4965–4969.
23. **Pellett, P. E., J. L. C. McKnight, F. J. Jenkins, and B. Roizman.** 1985. Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of trans-inducing α genes. Proc. Natl. Acad. Sci. USA **82**:5870–5874.
24. **Rautmann, G., H. W. D. Matthes, M. J. Gait, and R. Breathnach.** 1984. Synthetic donor and acceptor splice sites function in an RNA polymerase B (II) transcription unit. EMBO J. **3**:2021–2028.
25. **Rigby, P.W. J., M. Dickmann, C. Rhodes, and P. Berg.** 1977. Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Med. Biol. **113**:237–251.
26. **Roby, C., and W. Gibson.** 1986. Characterization of phosphoproteins and protein kinase activity of virions, noninfectious enveloped particles, and dense bodies of human cytomegalovirus. J. Virol. **59**:714–727.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
28. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
29. **Staden, R.** 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. **10**:4731–4752.
30. **Staden, R.** 1984. A computer program to enter DNA gel reading data into a computer. Nucleic Acids Res. **12**:499–504.
31. **Talbot, P., and J. D. Almeida.** 1977. Human cytomegalovirus: purification of enveloped virions and dense bodies. J. Gen. Virol. **36**:345–349.
32. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, and M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
33. **Wagner, E. K.** 1983. Transcription patterns in HSV infections. Adv. Viral Oncol. **3**:239–270.
34. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103–119.