JOURNAL OF VIROLOGY, July 1993, p. 3961–3968
0022-538X/93/073961-08$02.00/0

Vol. 67, No. 7

Identification of a New Transcriptional Unit That Yields a Gene Product within the Unique Sequences of the Short Component of the Herpes Simplex Virus 1 Genome

URANIA GEORGOPOULOU,[1] ATHINA MICHAELIDOU,[1] BERNARD ROIZMAN,[2] AND PENELOPE MAVROMARA-NAZOS[1]*

Department of Virology, The Hellenic Pasteur Institute, Athens, Greece,[1] and The Marjorie B. Kovler Viral Oncology Laboratories, The University of Chicago, Chicago, Illinois 60637[2]

Received 8 January 1993/Accepted 25 March 1993

The herpes simplex virus genome 1 consists of two unique stretches, long (U_L) and short (U_S), each flanked by inverted repeat sequences. The U_S sequence has been previously reported to contain 12 open reading frames designated U_S1 through U_S12. This report demonstrates the existence of a 13th open reading frame within the U_S sequence, designated $U_S8.5$. The $U_S8.5$ sequence is located between, and overlaps in part with, the domains of the U_S8 and U_S9 genes. Its transcription is initiated within the coding sequence of U_S8, and its transcript decays earlier than that of U_S8. On the basis of the size of its RNA (1.2 kb) and map position, it is likely that the $U_S8.5$ transcript is 3′ coterminal with the U_S8 and U_S9 mRNAs at the single polyadenylation signal which serves these genes. The nucleotide sequence of the $U_S8.5$ open reading frame predicts that its product is a 151-amino-acid basic, hydrophilic polypeptide. To determine whether the $U_S8.5$ encodes a protein, a sequence encoding 23 amino acids that contains an epitope reacting with a known monoclonal antibody to human cytomegalovirus protein was inserted in frame after the predicted fifth codon of the $U_S8.5$ gene. The recombinant virus carrying this epitope induced the synthesis of a protein reactive with the monoclonal antibody in immunoblots. The tagged protein localized in nucleoli of cells infected with the recombinant virus.

The herpes simplex virus 1 (HSV-1) genome consists of two covalently linked components, long (L) and short (S), each consisting of a unique sequence, U_L and U_S, respectively, flanked by inverted repeats (10, 38).

Analyses of the DNA sequence combined with transcriptional studies have led to the conclusion that the HSV-1 U_S sequence encodes 12 open reading frames designated U_S1 through U_S12. It was further reported that these open reading frames yield 13 mRNAs and that 11 of the 13 mRNAs are arranged into four nested families composed of two or three RNA species with common 3′ termini but distinct 5′ ends. Two other mRNAs have unique 5′ and 3′ ends (27, 33, 34). Concurrent and subsequent studies have shown that all but one gene, that encoding glycoprotein D (gD), are dispensable for viral replication in cells in culture (19, 20, 35).

In this article, we report the discovery of a new gene mapping in the S component. This gene, designated $U_S8.5$, yields an mRNA whose 3′ terminus may be colinear with those of U_S8 and U_S9 mRNAs.

MATERIALS AND METHODS

Cells and viruses. The origin, maintenance, and propagation of Vero, thymidine kinase-minus baby hamster kidney (BHKTK$^-$), 143TK$^-$, and rabbit skin cells were as previously described (2, 31). HSV-1(F) is a limited-passage wild-type isolate (7). The recombinant viruses R7030 and R7032 were constructed as detailed elsewhere (18, 28); R7030 contains a thymidine kinase gene under the control of the α27 promoter inserted into the unique *Hpa*I site of the U_S8 gene, whereas in R7032, a 1.8-kbp *Hin*cII-*Nru*I sequence was deleted from the U_S8 gene encoding gE.

Construction of plasmids. Plasmids pRB123 and pRB124 contain the *Bam*HI J and *Bam*HI X fragments of the HSV-1(F) DNA, respectively, inserted into the *Bam*HI site of pRB322 (30). Plasmid pRB417 contains the 2.7-kbp *Eco*RI-*Bgl*II subfragment of the HSV-1(F) U_S sequence inserted into the *Eco*RI-*Bam*HI sites of pUC8 (25). Plasmid pRB4321 was made by subcloning the 1.8-kbp *Sac*I-*Sac*I fragment from pRB123 into the *Sac*I site of the pGEM3Zf$^+$ vector. Plasmid pRB4319 contains the 1.8-kbp *Bam*HI X fragment derived from plasmid pRB124 inserted into the *Bam*HI site of pGEM3Zf$^+$. Plasmid pHPI321 was made by subcloning the 380-bp *Hin*cII-*Hin*cII fragment from pRB124 into the *Hin*cII site of pGEM3Zf$^+$. Plasmid pHPI400, kindly provided by Vivi Myriagou, contains the 2.1-kb *Nru*I-*Bam*HI fragment from pRB123 inserted into the *Bam*HI-*Hin*cII site of pGEM3Zf$^+$. Plasmid pHPI320, kindly provided by Evi Prousali, contains the 2.2-kbp *Sac*I-*Sac*I fragment from the HSV-1(F) *Eco*RI-H DNA inserted into the *Sac*I site of pGEM3Zf$^+$. Plasmid pHPI35 contains the *Eco*RI-H DNA fragment from the U_S component of the HSV-1(F) DNA, into which was inserted at the unique *Bgl*II site the 69-bp sequence 5′-GATCTAGAAAGGGACAGAAGCCCAACCT GCTAGACCGACTGCGACACCGCAAAAACGGGTAC CGACACA-3′. This sequence and its complement were synthesized in an Applied Biosystems DNA synthesizer 380A (Foster City, Calif.). The two oligonucleotides were annealed, and the double-stranded DNA oligomer was ligated into the unique *Bgl*II site of the *Eco*RI-H cloned DNA fragment to yield plasmid pHPI35. The oligonucleotide encodes an epitope recognized by monoclonal antibody CH28-2 to human cytomegalovirus (CMV) gB (16, 17); it was inserted in frame with the coding sequence of the predicted $U_S8.5$ open reading frame at codon 5. The monoclonal antibody to this epitope, a gift of L. Pereira, does not react with any HSV proteins in native or denatured form (16, 17).

* Corresponding author.

The DNA sequences encoding and flanking the epitope at the sites of insertion were sequenced to verify that the inserted epitope was intact and in frame with the open reading frame of the putative $U_S8.5$ gene. This oligonucleotide was designed to end in a *Bgl*II restriction enzyme site and to carry an *Xba*I site at the 5′ end and a *Kpn*I site at the 3′ end for convenient screening of the oligonucleotide insertion in the plasmids.

Plasmid pHPI324 contained the 2.2-kb *Sac*I-*Sac*I fragment derived from the pHPI35 inserted into the *Sac*I site of pGEM3Zf$^+$, and plasmid pHPI325 contained the same fragment but in the opposite orientation.

Construction of recombinant virus R35. Recombinant virus R35 was constructed as described in detail elsewhere (31). Briefly, intact R7030 DNA was cotransfected with pHPI35 plasmid DNA onto rabbit skin cells. The progeny of the transfection were plated on 143TK$^-$ cells in the presence of bromodeoxyuridine. In the presence of the drug, only the recombinant TK$^-$ viral progeny replicates and forms plaques. Individual plaques were picked, plaque purified again, and tested for the presence of the oligonucleotide insert by Southern blot analyses of electrophoretically separated *Kpn*I restriction endonuclease digests.

RNA analyses. Cytoplasmic RNA was prepared from confluent Vero or 143TK$^-$ cells grown in 150-cm flasks as previously described (39). Fifteen micrograms of RNA was subjected to electrophoresis on 1.4% agarose–formaldehyde gels as described by Maniatis et al. (22), transferred to a nitrocellulose membrane (Schleicher & Schuell, Dassel, Germany), and hybridized either with DNA probes nick translated with [^{32}P]dATP or with strand-specific ^{32}P-labeled RNA probes made with the Riboprobe system (Promega, Madison, Wis.) according to the procedure recommended by the manufacturer. The hybridization conditions for DNA were as previously described (30, 40). The hybridization with RNA probes was done at 80°C in 1 mM EDTA–0.5 M NaH_2PO_4 (pH 7.2)–1% sodium dodecyl sulfate (SDS) for about 12 h. Subsequently, the membranes were washed two times for 60 min each time in 1 mM EDTA–40 mM NaH_2PO_4 (pH 7.2)–1% SDS at 80°C and exposed to autoradiography. The S1 analyses were done as described elsewhere (13). The probe was a double-stranded 480-bp *Mlu*I-*Bam*HI DNA fragment derived from pHPI400, 5′ end labeled with polynucleotide kinase at the *Bam*HI site by using standard procedures.

In vitro transcription and translation. Plasmid templates were transcribed with Sp6 or T7 RNA polymerase in the presence of the cap analog m^{7}G(5′)ppp(5′)G (New England BioLabs) as recommended by the manufacturer (Promega). One microgram of synthetic RNA was translated in a 20-μl reaction mixture with a rabbit reticulocyte lysate in the presence of [^{35}S]methionine (Amersham International, Amersham, Buckinghamshire, United Kingdom) according to the protocol supplied by the company.

Transfections. Approximately 4×10^6 BHKTK$^-$ cells were transfected with 20 μg of plasmid DNA by the calcium phosphate precipitation method as previously described (2, 25). The cells were processed 48 h after transfection.

Immunofluorescence. The BHKTK$^-$ cell cultures transfected with plasmid pHPI35 or infected with R35 were washed twice with phosphate-buffered saline (PBS) and then fixed for 20 min at −20°C in methanol (37). After fixation, the cells were allowed to dry overnight; after a brief rinse in PBS, they were exposed to anti-CMV monoclonal antibody CH28-2. After 2 h, the cultures were washed three times for 10 min each time in PBS and then exposed to a 1:100 dilution of fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin G (Biosoft Laboratories, Paris, France) plus 1:10 volumes of Evans blue counterstain to overcome nonspecific staining. After four 10-min washes in PBS, the bottom of the flask was cut out, a coverslip was applied, and the cells were photographed under UV light illumination.

SDS-polyacrylamide gel electrophoresis and immunoblotting. The denatured solubilized polypeptides from mock-infected and R35-infected Vero cell lysates or the products of the in vitro transcription-translation described above were electrophoretically separated on 10 or 12% SDS-polyacrylamide gels (8, 9, 29). For immunoblotting, the electrophoretically separated polypeptides from infected cells were transferred electrically to nitrocellulose membranes (Schleicher & Schuell) and reacted in an enzyme-linked immunoassay first with monoclonal antibody CH28-2 against the CMV epitope and then with a biotinylated anti-mouse antibody as previously described (16, 17). The gels containing the electrophoretically separated in vitro-translated polypeptides were dried and exposed to Kodak X-Omat film.

RESULTS

Mapping of a new RNA transcript to the domain of the U_S8-U_S9 gene family in the S component of HSV-1 DNA. The experiments described below indicated that a novel, highly abundant RNA made relatively early in infection mapped to the domain of the U_S8-U_S9 gene family.

RNAs extracted at various times postinfection from 143TK$^-$ cells infected with 5 PFU of HSV-1(F) or with the recombinant R7032 (gE$^-$) per cell were separated on a formaldehyde-agarose gel, transferred by blotting to a nitrocellulose membrane, and probed with the ^{32}P-labeled pHPI321 plasmid DNA (probe 1 in Fig. 1). On the basis of the RNA maps published previously (27, 34), this probe should hybridize only to U_S8 and U_S9 RNAs. The autoradiograms shown in Fig. 2 indicated the following.

(i) Probe 1 (pHPI321) hybridized consistently to three HSV-1(F) transcripts approximately 2.7, 1.2, and 0.6 kb in size. On the basis of size, the 2.7- and 0.6-kb transcripts were identified as the U_S8 and U_S9 mRNAs. The presence of the 1.2-kb HSV-1 RNA, indicated by the arrow in Fig. 2, was totally unexpected. Similar results were obtained with the RNA extracted from HSV-1(F)-infected Vero cells at 4 and 7 h postinfection. As shown in Fig. 3, probe 1 recognized three HSV-1 RNAs of 2.7, 1.2, and 0.6 kb (Fig. 3, lanes 2 and 3). RNA extracted from mock-infected cells did not hybridize with this probe (Fig. 2 and 3, lanes 1).

(ii) The kinetics of accumulation of the 1.2-kb mRNA differed from that of the U_S8 and U_S9 mRNAs. None of the RNAs were detectable at 3 h postinfection. Whereas between 7 and 14 h postinfection the amounts of U_S8 and U_S9 RNAs increased, the amounts of $U_S8.5$ mRNAs decreased (Fig. 2, lanes 7 to 13).

(iii) Hybridization of the probe with RNA extracted from R7032 infected 143TK$^-$ cells identified only the 0.6-kb (U_S9) RNA. Both the 2.7-kb (U_S8) and 1.2-kb RNAs were absent (Fig. 2, lanes 5, 7, 9, 11, and 13). The RNA extracted from cells infected with the R7032 recombinant exhibited two additional RNA species of approximately 2.9 and 2.5 kb. Their origin remains uncertain; it is conceivable that they represent read-through transcription from the upstream U_S5-U_S6-U_S7 3′ coterminal mRNAs, since the deletion within the U_S component of the R7032 genome removed the predicted poly(A) site.

The results of these experiments indicate that the abun-

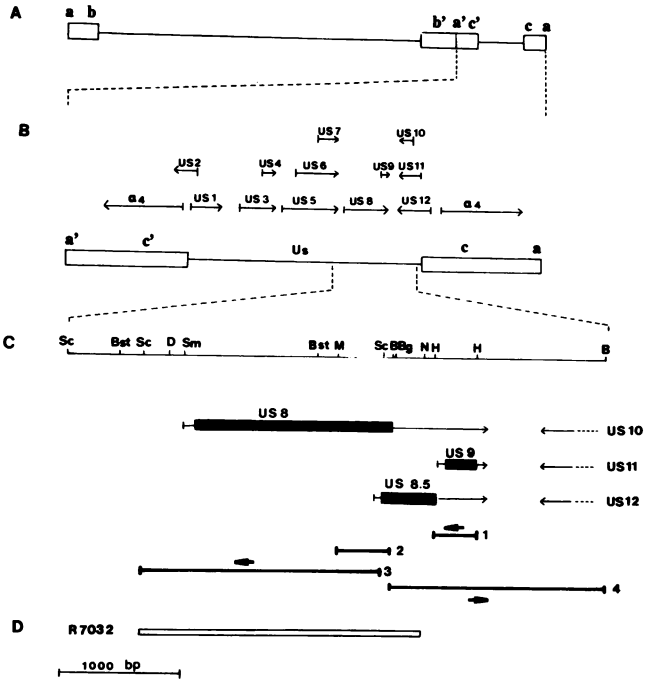

FIG. 1. Sequence arrangement of the HSV-1 genome and locations of the genes in the S component of the HSV-1 genome. (A) Schematic diagram of the HSV-1 genome in the prototype arrangement, showing the unique sequences (lines) flanked by the inverted repeats ab, b'a', a'c', and ca (rectangles). (B) Expanded representation of the S component of the HSV-1 genome depicting the gene arrangement (arrows) contained within these DNA sequences. (C) Restriction endonuclease map of the *Sac*I-*Bam*HI fragment from the unique sequence of the S component containing the U_S8, U_S8.5 (identified in this study), and U_S9 genes (Sc, *Sac*I; Bst, *Bst*EII; D, *Dde*I; M, *Mlu*I; B, *Bam*HI; Bg, *Bgl*II; N, *Nru*I; H, *Hin*cII). Arrows below represent the positions of the U_S8.5 and U_S9 gene sequences contained in this fragment. Heavy lines represent the coding sequences; thin lines represent transcribed untranslated RNA sequences. The lines below, numbered 1 to 4, represent DNA sequences used as probes for the RNA analysis. Line 1 represents the 380-bp *Hin*cII-*Hin*cII fragment cloned in pGEM3Zf$^+$ as pHPI321; line 2 represents the *Mlu*I-*Bam*HI fragment used as the probe for the S1 analysis; line 3 represents the 1.8-kb *Sac*I-*Sac*I fragment contained in plasmid pRB4321; line 4 represents the 1.8-kb *Bam*HI X fragment contained in plasmid pRB4319. The arrows show the direction of Sp6 transcription in vitro. (D) Schematic representation of the deletion in the R7032 recombinant. The open rectangle represents the DNA sequence deleted from the R7032 genome.

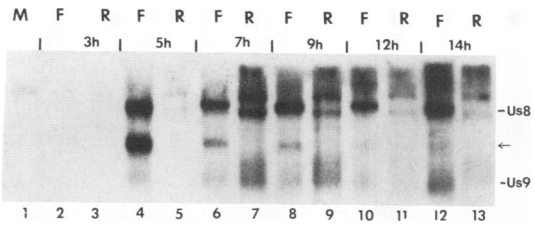

FIG. 2. Identification of a novel transcript overlapping the U_S8-U_S9 3' coterminal gene family. Shown is an autoradiographic image of a Northern (RNA) blot of formaldehyde-agarose gel-separated cytoplasmic RNA isolated from mock-infected (M), HSV-1(F)-infected (F), and R7032-infected (R) 143TK$^-$ cells hybridized with nick-translated ^{32}P-labeled DNA probe 1 (pPHI321; see Fig. 1 for location of the probe). Lanes: 2, 4, 6, 8, 10, and 12, 15 μg of RNA isolated from HSV-1(F)-infected 143TK$^-$ cells harvested at 3, 5, 5, 9, 12, and 14 h postinfection, respectively; 3, 5, 7, 9, 11, and 13, 15 μg of RNA isolated from R7032-infected 143TK$^-$ cells harvested at 3, 5, 7, 9, 11, and 13 h postinfection, respectively. Positions of the U_S8 and U_S9 transcripts are indicated by name; the arrow indicates a novel 1.2-kb transcript in the HSV-1(F)-infected 143TK$^-$ cells, designated U_S8.5.

dant 1.2-kb transcript represents a new, not previously identified RNA. This RNA maps, at least in part, within the domain of the U_S8-U_S9 3' coterminal transcriptional units, and its synthesis is not cell type specific inasmuch as it was present in infected cells derived from both African green monkeys (Vero) and humans (143TK$^-$). In anticipation of the results presented below, this RNA was designated U_S8.5 RNA.

Fine mapping of the U_S8.5 transcript. The results of two series of experiments described below suggest that the U_S8.5 RNA is colinear with the U_S8-U_S9 open reading frames.

In the first series, Vero cells were infected with HSV-1(F) at a multiplicity of 5 PFU per cell. Cytoplasmic RNAs harvested at 4 or 7 h postinfection were separated on formaldehyde-agarose gels, transferred to nitrocellulose membranes, and hybridized to six ^{32}P-labeled RNA probes derived by transcription in vitro of HSV-1 DNA fragments cloned in plasmids pHPI321, pRB4319, and pRB4321 (Fig. 1) with either Sp6 or T7 polymerase. The results (Fig. 3) were as follows.

(i) The labeled RNA transcribed in vitro with Sp6 polymerase from plasmid pHPI321 (probe 1; Fig. 1) hybridized to three RNA bands which correspond to the U_S8, U_S8.5, and U_S9 RNA species (Fig. 3B, lane 5). Labeled RNA generated in vitro with T7 polymerase from the same plasmid failed to hybridize to any RNA species.

(ii) The labeled RNA transcribed in vitro with T7 polymerase from plasmid pRB4319 (probe 4; Fig. 1) hybridized to all three RNAs (Fig. 3C, lanes 9 and 10), whereas the RNA transcribed in vitro with the Sp6 polymerase from the same plasmid hybridized with a broad RNA band averaging about 1.3 kb which corresponds to the U_S10-U_S11-U_S12 coterminal RNA family (Fig. 3C, lanes 12 and 13).

(iii) The labeled RNA derived by in vitro transcription of plasmid pRB4321 (probe 3; Fig. 1) with Sp6 hybridized only to the U_S8 RNA (Fig. 3D, lanes 18 and 19), whereas the T7 polymerase-transcribed RNA from the same plasmid failed to hybridize to any RNA species (Fig. 3D, lanes 14 to 16).

These results indicate that the U_S8.5 RNA is colinear with the U_S8 and U_S9 RNAs and is derived from rightward transcription of a portion of the U_S sequence of the HSV-1 genome in the prototype arrangement. These results also predict that the U_S8.5 RNA represents the third member of the U_S8-U_S9 3' coterminal RNA family.

The purpose of the second series of experiments was to determine the probable transcription initiation site of the U_S8.5 RNA. Relevant to the design of this experiment were the following data.

(i) Analyses of the RNA established that the size of the RNA was 1.2 kbp.

(ii) The bulk of the U_S8.5 RNA had to map downstream of the U_S8 open reading frame inasmuch as probe 3 (Fig. 1) did not hybridize to the U_S8.5 RNA.

(iii) As demonstrated in detail below, the domain to which the U_S8.5 RNA mapped contains an open reading frame which was expressed in infected cells and could be expressed in vitro in coupled transcription-translation systems, and therefore the U_S8.5 mRNA encoded the information for

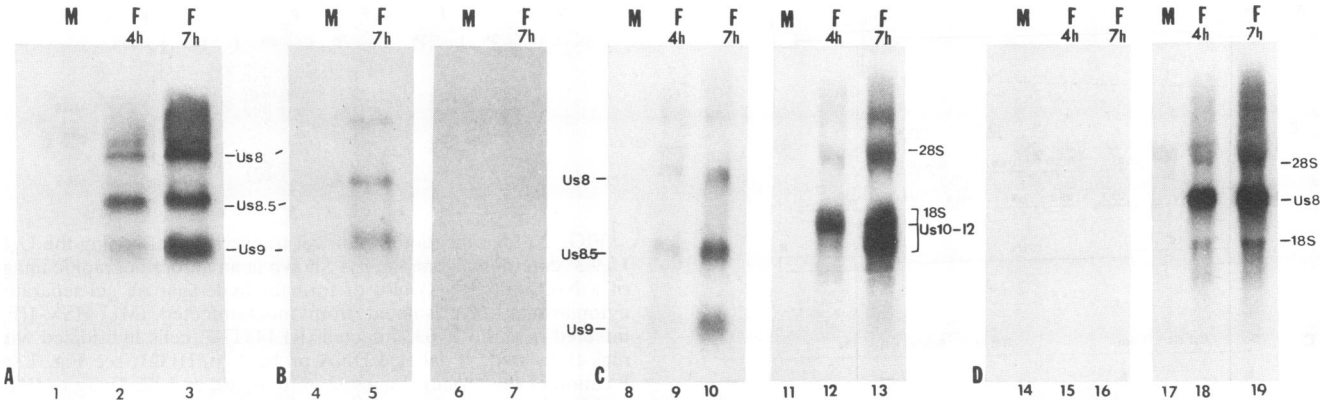

FIG. 3. Mapping of the $U_S8.5$ transcript. (A) Autoradiographic image of nick-translated ^{32}P-labeled DNA probe 1 (plasmid pHPI321; see Fig. 1 for location of the probe) hybridized to RNA extracted from mock-infected (M; lane 1) or HSV-1(F)-infected (F) Vero cells harvested at 4 h (lane 2) or at 7 h (lane 3) postinfection, electrophoretically separated in formaldehyde-agarose gels, and transferred to a nitrocellulose sheet. (B to D) Same as panel A except that the electrophoretically separated RNAs were hybridized with strand-specific Sp6 (lanes 4, 5, 11–13, and 17–19) or T7 (lanes 6, 7, 8–10, and 14–16)-derived ^{32}P-labeled RNA probes. The probes in panel B were derived from plasmid pHPI321 (line 1 in Fig. 1C); the probes in panel C were derived from plasmid pRB4319 (line 4 in Fig. 1C). The probes in panel D were derived from plasmid pRB4321 (line 3 in Fig. 1C). Lanes: 4, 6, 8, 11, 14, and 17, 15 μg of RNA from mock-infected Vero cells; 9, 12, 15, and 18, 15 μg of RNA from HSV-1(F)-infected Vero cells harvested at 4 h postinfection; 5, 7, 10, 13, 16, and 19, 15 μg of RNA from HSV-1(F)-infected Vero cells harvested at 7 h postinfection. Positions of the U_S8, $U_S8.5$, and U_S9 RNAs are indicated by name.

the $U_S8.5$ protein. The coding sequence was not spliced, and given the nature of the region, i.e., the presence of 3′ coterminal mRNAs (U_S8 and U_S9) which were the products of flanked genes, it seemed highly probable, as indicated above, that the polyadenylation signal for the U_S8 and U_S9 open reading frames was also shared by $U_S8.5$ RNA. These results predicted that the 5′ terminus of the $U_S8.5$ RNA would map within the last 400 bases of the U_S8 open reading frame close to the *Sac*I restriction endonuclease site. To test this prediction, the *Bam*HI-*Mlu*I DNA fragment which spans the *Sac*I site (probe 2; Fig. 1) was 5′ end labeled with T4 polynucleotide kinase at the *Bam*HI site and hybridized to cytoplasmic RNA extracted from Vero cells 7 h postinfection from HSV-1(F).

As shown in Fig. 4, lane 2, two bands of about 460 and 124 nucleotides were detected in the 6% polyacrylamide–urea gel containing the electrophoretically separated S1 digest. The 460-nucleotide top band is the result of protection from S1 digestion of the entire probe by the U_S8 RNA. The lower band corresponds to a fragment protected from S1 digestion by an RNA initiated within the U_S8 open reading frame as predicted above for the $U_S8.5$ RNA. The size of the protected fragment placed the transcription initiation site of the $U_S8.5$ within the U_S8 open reading frame approximately 27 bases downstream from the first T of a potential TATA box. DNA sequencing analyses revealed a potential open reading frame for the $U_S8.5$ RNA starting 123 bases downstream of the transcription initiation site for the $U_S8.5$ RNA (Fig. 5).

The $U_S8.5$ open reading frame specifies a protein. In the preceding section, we demonstrated that infected cells contain a previously unreported RNA that is made early in infection and initiates in the domain of the U_S8 gene. The purpose of the two series of experiments described in this section were to determine whether the $U_S8.5$ transcript specifies a protein.

As described in Materials and Methods, a DNA sequence encoding a CMV epitope was inserted in frame between codons 5 and 6 of the predicted coding domain of the $U_S8.5$ gene to yield a plasmid designated pHPI35. The chimeric

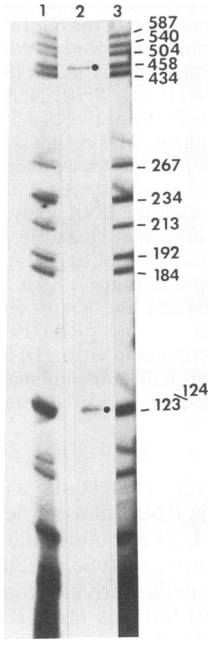

FIG. 4. Mapping of the 5′ termini of the $U_S8.5$ transcript by S1 analysis. One hundred micrograms of cytoplasmic RNA from HSV-1(F)-infected Vero cells harvested at 7 h postinfection was hybridized with a 460-bp *Mlu*I-*Bam*HI fragment labeled with ^{32}P at the *Bam*HI site (Fig. 1C, line 2). The hybridization mixture was digested with S1 nuclease, and S1-resistant fragments were separated in a denatured polyacrylamide gel (lane 2). For size markers (indicated in bases), plasmid pUC19 was digested with *Hae*III, 5′ labeled with ^{32}P, denatured, and electrophoresed (lanes 1 and 3). The positions of the protected bands are indicated by the dots.

```
    TGTGGGCGTGTATGACCTGTTGGCGCAGGCGTGCCTGGCGGGCGGTTAAAAGCAGGGCCT
1   ---------+---------+---------+---------+---------+---------+ 60
      TrpAlaCysMetThrCysTrpArgArgArgAlaTrpArgAlaValLysSerArgAlaSer -

    CGGGTAAGGGGCCCACGTACATTCGCGTGGCCGACAGCGAGCTGTACGCGGACTGGAGCT
61  ---------+---------+---------+---------+---------+---------+ 120
      GlyLysGlyProThrTyrIleArgValAlaAspSerGluLeuTyrAlaAspTrpSerSer -

                                                          SacI
    CGGACAGCGAGGGAGAACGCGACCAGGTCCCGTGGCTGGCCCCCCCGGAGAGACCCGACT
121 ---------+---------+---------+---------+---------+---------+ 180
      AspSerGluGlyGluArgAspGlnValProTrpLeuAlaProProGluArgProAspSer -

            BamHI       BglII
    CTCCCTCCACCAATGGATCCGGCTTTGAGATCTTATCACCAACGGCTCCGTCTGTATACC
181 ---------+---------+---------+---------+---------+---------+ 240
               MetAspProAlaLeuArgSerTyrHisGlnArgLeuArgLeuTyrThr   -
      ProSerThrAsnGlySerGlyPheGluIleLeuSerProThrAlaProSerValTyrPro -

                                          BamHI
    CCCGTAGCGATGGGCATCAATCTCGCCGCCAGCTCACAACCTTTGGATCCGGAAGGCCCG
241 ---------+---------+---------+---------+---------+---------+ 300
      ProValAlaMetGlyIleAsnLeuAlaAlaSerSerGlnProLeuAspProGluGlyPro  -
        ArgSerAspGlyHisGlnSerArgArgGlnLeuThrThrPheGlySerGlyArgProAsp -

    ATCGCCGTTACTCCCAGGCCTCCGATTCGTCCGTCTTCTGGTAAGGCGCCCCATCCCGAG
301 ---------+---------+---------+---------+---------+---------+ 360
      IleAlaValThrProArgProProIleArgProSerSerGlyLysAlaProHisProGlu  -
        ArgArgTyrSerGlnAlaSerAspSerSerValPheTrpEnd

    GCCCCACGTCGGTCGCCGAACTGGGCGACCGCCGGCGAGGTGGACGTCGGAGACGAGCTA
361 ---------+---------+---------+---------+---------+---------+ 420
      AlaProArgArgSerProAsnTrpAlaThrAlaGlyGluValAspValGlyAspGluLeu  -

    ATCGCGATTTCCGACGAACGCGGACCCCCCCCGACATGACCGCCCGCCCCTCGCCACGTCG
421 ---------+---------+---------+---------+---------+---------+ 480
      IleAlaIleSerAspGluArgGlyProProArgHisAspArgProProLeuAlaThrSer  -

    ACCGCGCCCTCGCCACACCCGCGACCCCCGGGCTACACGGCCGTTGTCTCCCCGATGGCC
481 ---------+---------+---------+---------+---------+---------+ 540
      ThrAlaProSerProHisProArgProProGlyTyrThrAlaValValSerProMetAla  -

    CTCCAGGCTGTCGACGCCCCCTCCCTGTTTGTCGCCTGGCTGGCCGCTCGGTGGCTCCGG
541 ---------+---------+---------+---------+---------+---------+ 600
      LeuGlnAlaValAspAlaProSerLeuPheValAlaTrpLeuAlaAlaArgTrpLeuArg  -

    GGGGCTTCCGGCCTGGGGGCCTCCTGTGTGGGATTGCGTGGTATGTGA
601 ---------+---------+---------+---------+--------- 649
      GlyAlaSerGlyLeuGlyAlaSerCysValGlyLeuArgGlyMetEnd  -
```

FIG. 5. Nucleotide and predicted amino acid sequences of the carboxyl terminus of the U_S8 open reading frame and the entire U_S8.5 open reading frame. The arrow indicates the approximate location of the transcription initiation site for the U_S8.5 RNA as determined from S1 nuclease analysis. The line indicates the position of a potential TATA box sequence upstream of the cap site of the U_S8.5 transcript.

gene was recombined into the viral genome, and the virus containing the tagged U_S8.5 gene was designated R35.

In the first series of experiments, replicate Vero cell monolayers grown in 25-cm^2 tissue culture flasks were mock infected or infected with the R35 recombinant virus at a multiplicity of 10 PFU per cell. Infected cell cultures were harvested at 4, 7, 13, and 24 h postinfection, solubilized in disruption buffer containing SDS, electrophoretically separated in denaturing gels, electrically transferred to nitrocellulose, and reacted with monoclonal antibody CH28-2 against CMV gB. Figure 6 shows that monoclonal antibody CH28-2 reacted with a specific R35-infected cell band with an apparent M_r of 23,000. The band was first apparent as a faint, barely visible band at 4 h postinfection and could be readily detected at 7 h. At later times, a less intense, higher-molecular-weight band reacted with the antibody. Neither band could be detected in mock-infected cells.

In the second series of experiments, BHKTK$^-$ cells were

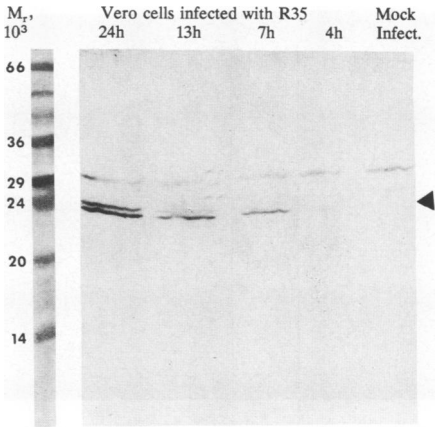

FIG. 6. Photograph of electrophoretically separated proteins from mock-infected Vero cells or cells infected with 10 PFU of R35 recombinant virus per cell and harvested at 4, 7, 13, or 24 h postinfection, solubilized in denaturing buffer, electrophoretically separated in a denaturing polyacrylamide gel, electrically transferred to nitrocellulose sheets, and reacted with monoclonal antibody CH28-2 to CMV. The left lane shows molecular weight markers. The band formed by the U_S8.5 protein is identified by the arrowhead.

either transfected with plasmid pHPI35 or infected with recombinant virus R35. Specific immunofluorescence was observed in both cases. Interestingly, in the case of the infected cells, the fluorescence was localized in the nucleoli of the cell, whereas in the case of the transfected cells, the fluorescence was localized reproducibly in rounded cells (Fig. 7). Nucleolar localization of the U_S8.5 protein in infected cells was verified by phase fluorescence microscopy.

Lastly, the U_S8.5 open reading frame was translated in vitro from RNA transcribed in vitro with the Promega transcription-translation system. Plasmid pHPI320 containing the U_S8.5 open reading frame cloned into the pGEM3Zf$^+$ expression vector was transcribed with Sp6 polymerase. The resulting RNA was then translated with the aid of a rabbit reticulocyte lysate. A protein which migrated in denaturing polyacrylamide gel electrophoresis as a single band with an apparent molecular weight of approximately 23,000 was produced (Fig. 8A). The transcription-translation of plasmid pHPI324 containing the CMV epitope inserted into the *Bgl*II site in frame with the U_S8.5 open reading frame with Sp6 polymerase resulted in a slower-migrating single band with an apparent M_r of 29,000 (Fig. 8B). No specific band was obtained following translation of the RNA obtained by transcription of pHPI325 with Sp6 polymerase.

DISCUSSION

The salient features of the results presented in this report are as follows.

(i) Several types of analyses, including mapping of viral mRNAs, mapping of temperature-sensitive lesions, mapping of domains encoding specific proteins, and particularly sequencing of the HSV-1 genome, led to the conclusion that the HSV-1 genome encodes 72 open reading frames (23, 24, 26, 27, 30, 32, 41). As McGeoch et al. (26, 27) had every reason to anticipate, this number formed a new lower limit inasmuch as less abundant RNAs were likely to be discovered subsequent to the publication of the sequence. Indeed,

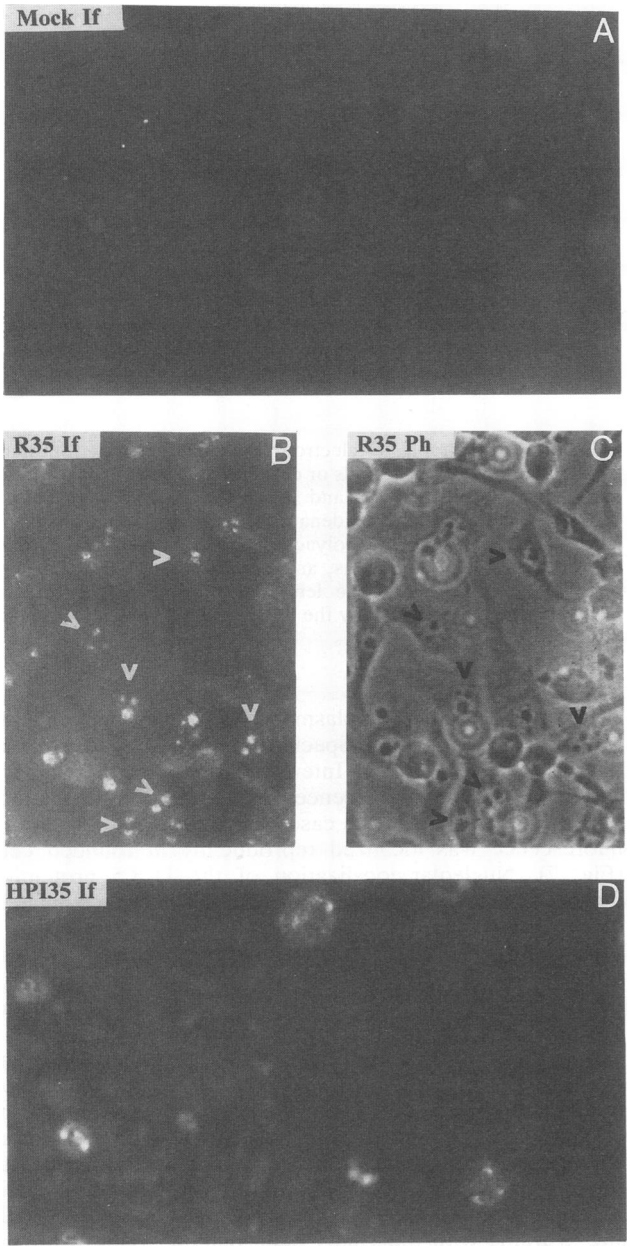

FIG. 7. Photographs of BHK cells mock-infected (A), infected with R35 recombinant virus (B and C), or transfected with plasmid pHPI35 (D). For immunofluorescence micrography the cells were stained with monoclonal antibody CH28-2 and counterstained with goat anti-mouse immunoglobulin G conjugated to fluorescein isothiocyanate. If, immunofluorescence; Ph, phase microscopy. The arrowheads in panels B and C show the correspondence of the $U_S8.5$ antigen seen by immunofluorescence with nucleoli seen by phase microscopy.

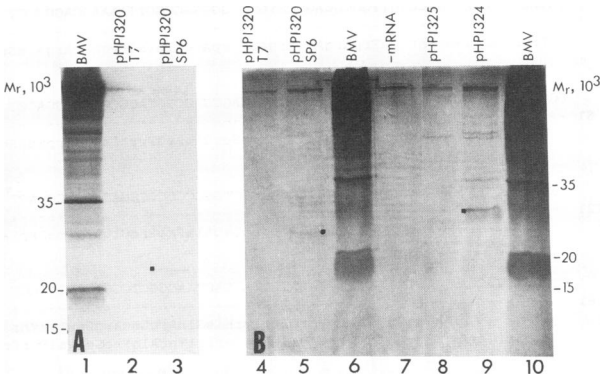

FIG. 8. Synthesis of the $U_S8.5$ protein in an in vitro transcription-translation system. Shown is an autoradiographic image of [^{35}S]methionine-labeled polypeptides encoded by the $U_S8.5$ open reading frame. The RNAs transcribed in vitro by the Sp6 or T7 polymerase of the 2.2-kbp, *Sac*I-*Sac*I DNA fragment (plasmid pHPI320) (A) and of the same 2.2-kb *Sac*I fragment but carrying the CMV epitope in the *Bgl*II site (plasmids pHPI325 and pHPI324) (B) were translated in a nuclease-treated rabbit reticulocyte lysate and electrophoretically separated in a 12% denaturing polyacrylamide gel. Lanes: 1, 6, and 10, translation product of the brome mosaic virus (BMV) template provided with the kit (Promega); 2 and 4, translation of the in vitro T7 polymerase-transcribed RNA from plasmid pHPI320; 3 and 5, translation of the in vitro Sp6 polymerase-transcribed RNA from plasmid pHPI320; 8 and 9, translation of the Sp6 polymerase-transcribed RNA from plasmids pHPI325 and pHPI324, respectively.

direct demonstration of a gene product (as opposed to computer-based analyses) led to the demonstration of four additional transcriptional units containing expressed open reading frames. These were the two copies of the $\gamma_2$34.5 gene mapping in the inverted repeats of the L component (1, 4–6) and the U_L26.5 (16, 17) and U_L49.5 (3) genes mapping in the U_L sequences of the L component. In this article, we report an additional gene which we have designated $U_S8.5$. We should note that unlike the transcript of the $\gamma_1$34.5 gene, the transcript of the $U_S8.5$ gene is quite abundant. The reason why it was missed in the initial studies of the U_S sequence (34) is unclear. The HSV-1(17) strain sequenced by McGeoch et al. (27) shows an open reading frame of a size consistent with the protein detected in our studies.

(ii) The S1 analyses reported here and the size of the $U_S8.5$ RNA suggest that transcription of the $U_S8.5$ gene is initiated within the coding domain of the U_S8 gene and that as noted in Results, the transcription initiation site is preceded at the appropriate distance by a TATA box. The first methionine codon which could serve as the translation initiation codon is located 123 bases from the transcription initiation site. Although the predicted 20th in-frame codon is a better methionine initiation codon than the first (14, 15), the observation that the sequence encoded by the CMV tag and inserted between codons 5 and 6 is expressed indicates that the first methionine is the initiator codon. On the basis of the size of the transcript and its initiation site, we predict that it is coterminal with the 3′ termini of the U_S8 and U_S9 transcripts. This conclusion is reinforced by the observation that a single polyadenylation signal is available for all three genes.

(iii) HSV genes form several groups (a, β, and γ) whose expression is coordinately regulated and sequentially ordered in a cascade fashion (11, 12). The γ genes are expressed late in infection and are further subdivided into two groups, γ_1 and γ_2, differing in their requirements for viral DNA synthesis for expression (35). Preliminary studies indicate that unlike the γ_2 U_S8 gene, the $U_S8.5$ gene is expressed earlier and is regulated as a β or γ_1 gene (data not shown).

(iv) The nucleotide sequence of the coding domain of

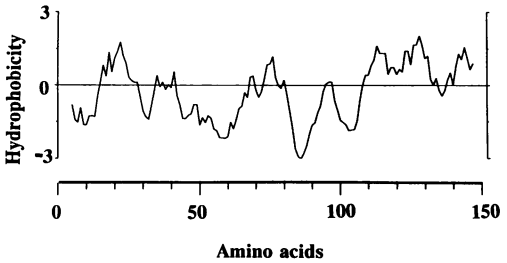

FIG. 9. Kyte-Doolittle hydropathic analysis of the protein predicted by the HSV-1 U_S8.5 open reading frame. The sequences were scanned with the University of Wisconsin Wisconsin Genetic Computer Group package in a DEC5000 work station, using a window of seven residues.

U_S8.5 predicts a basic protein of 151 amino acids. The hydrophobicity analysis (Fig. 9) predicts a hydrophilic protein. The apparent M_r of the tagged U_S8.5 protein (151 predicted amino acids plus the 23-amino-acid tag) as determined from electrophoretic mobility in lysates of cells infected with the R35 virus was 23,000. The electrophoretic mobility of the protein translated in vitro (Fig. 8) yielded an apparent M_r of 29,000. As a general rule, HSV proteins show an apparent molecular weight 15% to as much as 50% (e.g., ICP4) greater than that deduced from the coding sequence. Nevertheless, the difference in electrophoretic mobility of these two products suggest that the two products carry different posttranslational modifications. The nature of the posttranslational modifications of the U_S8.5 protein is under study.

(v) In immunofluorescence studies with the tagged virus, we detected the protein in the nucleoli of infected cells. We should note that another HSV-1 protein, the product of the U_S11 gene, also localizes in nucleoli (21, 37). However, the U_S11 protein also binds to the 60S ribosomal subunit and is an abundant constituent of the herpes simplex virion (34, 36). Whether the product of the U_S8.5 protein localizes solely to nucleoli, whether it is restricted to nucleoli because of a functional defect induced by the insertion of the epitope, and whether, like the protein product of the U_S11 gene, it performs multiple functions remains to be determined.

(vi) The PFU/50% lethal dose ratio for R7032 virus following intracerebral inoculation was higher for than that of the parent HSV-1(F) virus (28). R7032 multiplies well in cells in culture, and therefore we may conclude that U_S8.5 is not essential for viral replication in cultured cells tested to date. Inasmuch as R7032 lacks both U_S8 and U_S8.5, it is not known whether its phenotype in experimental animal systems is due to the absence of U_S8, U_S8.5, or both genes.

ACKNOWLEDGMENTS

We thank Lenore Pereira for the monoclonal antibody and V. Myriagou, R. J. Roller, and P. Ward for invaluable advice.

The studies at the Hellenic Pasteur Institute were aided by grants from the Greek Public Health Service and the National Secretariat for Research and Technology. These studies were aided by Public Health Service grants from the National Cancer Institute (CA47451) and the National Institute for Allergy and Infectious Diseases (AI24009) and by an unrestricted grant from Bristol-Myers Squibb Program in Infectious Diseases.

REFERENCES

1. **Ackermann, M., J. Chou, M. Sarmiento, R. A. Lerner, and B. Roizman.** 1986. Identification by antibody to a synthetic peptide of a protein specified by a diploid gene located in the terminal repeats of the L component of herpes simplex virus genome. J. Virol. **58**:843–850.
2. **Arsenakis, M., J. Hubenthal-Voss, G. Campadelli-Fiume, L. Pereira, and B. Roizman.** 1986. Construction and properties of a cell line constitutively expressing the herpes simplex virus glycoprotein B dependent on functional α4 protein synthesis. J. Virol. **60**:674–682.
3. **Barker, D. E., and B. Roizman.** 1992. The unique sequence of the herpes simplex virus 1 L component contains an additional translated open reading frame designated U_L49.5. J. Virol. **66**:562–566.
4. **Chou, J., E. R. Kern, R. J. Whitley, and B. Roizman.** 1990. Mapping of herpes simplex virus 1 neurovirulence to $\gamma_1$34.5, a gene nonessential for growth in cell culture. Science **252**:1262–1266.
5. **Chou, J., and B. Roizman.** 1986. The terminal *a* sequence of the herpes simplex virus genome contains the promoter of a gene located in the repeat sequences of the L component. J. Virol. **57**:629–637.
6. **Chou, J., and B. Roizman.** 1990. The herpes simplex virus 1 gene for ICP34.5, which maps in inverted repeats, is conserved in several limited passage isolates but not in strain 17syn+. J. Virol. **64**:1014–1020.
7. **Ejercito, P. M., E. D. Kieff, and B. Roizman.** 1968. Characterization of herpes simplex virus strains differing in their effect on social behavior of infected cells. J. Gen. Virol. **2**:357–364.
8. **Gibson, W., and B. Roizman.** 1972. Proteins specified by herpes simplex virus. VIII. Characterization and composition of multiple capsid forms of subtypes 1 and 2. J. Virol. **10**:1044–1052.
9. **Gibson, W., and B. Roizman.** 1974. Protein specified by herpes simplex virus. X. Staining and radiolabeling properties of B capsids and virion proteins in polyacrylamide gels. J. Virol. **13**:155–165.
10. **Hayward, G. S., R. J. Jacob, S. C. Wadsworth, and B. Roizman.** 1975. Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. Proc. Natl. Acad. Sci. USA **72**:4243–4247.
11. **Honess, R. W. and B. Roizman.** 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14**:8–19.
12. **Honess, R. W., and B. Roizman.** 1975. Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA **72**:1276–1280.
13. **Jenkins, F. J., and M. K. Howett.** 1984. Characterization of the mRNA mapping in the *Bgl*II N fragment of the herpes simplex virus type 2 genome. J. Virol. **52**:99–107.
14. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
15. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108**:229–241.
16. **Liu, F., and B. Roizman.** 1991. The promoter, transcriptional unit, and coding sequence of herpes simplex virus 1 family 35 proteins are contained within and in frame with the U_L26 open reading frame. J. Virol. **65**:206–212.
17. **Liu, F., and B. Roizman.** 1991. The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. J. Virol. **65**:5149–5156.
18. **Longnecker, R., S. Chatterjee, R. J. Whitley, and B. Roizman.** 1987. Identification of a novel herpes simplex virus 1 glycoprotein gene within a gene cluster dispensable for growth in cell culture. Proc. Natl. Acad. Sci. USA **84**:4303–4307.
19. **Longnecker, R., and B. Roizman.** 1986. Generation of inverting herpes simplex virus 1 mutant lacking the L-S junction a sequences, an origin of DNA synthesis including those specifying glycoprotein E and α47. J. Virol. **58**:583–591.
20. **Longnecker, R., and B. Roizman.** 1987. Clustering of genes dispensable for growth in cell culture in the small component of the herpes simplex virus 1 genome. Science **236**:573–576.
21. **MacLean, C. A., F. J. Rixon, and H. S. Marsden.** 1987. The

products of gene U_S11 of herpes simplex virus type 1 are DNA-binding and localize to the nucleoli of infected cells. J. Gen. Virol. **68**:1921–1937.

22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1990. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Marsden, H. S., N. D. Stow, V. G. Preston, M. C. Timbury, and N. M. Wilkie.** 1978. Physical mapping of herpes simplex virus-induced polypeptides. J. Virol. **28**:624–642.
24. **Matz, B., J. H. Subak-Sharpe, and V. G. Preston.** 1983. Physical mapping of temperature-sensitive mutations of herpes simplex virus type 1 using cloned restriction endonuclease fragments. J. Gen. Virol. **64**:2261–2270.
25. **Mavromara-Nazos, P., S. D. Silver, J. Hubenthal-Voss, J. L. C. McKnight, and B. Roizman.** Regulation of herpes simplex virus 1 genes: α gene sequence requirements for transient induction of indicator genes regulated by β or late (γ_2) promoters. Virology **149**:152–164.
26. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
27. **McGeoch, D. J., A. Dolan, S. Donald, and F. J. Rixon.** 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. **181**:1–13.
28. **Meignier, B., R. Longnecker, P. Mavromara-Nazos, A. E. Sears, and B. Roizman.** 1987. Virulence of and establishment of latency by genetically engineered mutants of herpes simplex virus 1. Virology **162**:251–254.
29. **Morse, L. S., L. Pereira, B. Roizman, and P. A. Schaffer.** 1978. Anatomy of herpes simplex virus (HSV) DNA. XI. Mapping of viral genes by analysis of polypeptides and functions specified by HSV-1 × HSV-2 recombinants. J. Virol. **26**:389–410.
30. **Post, L. E., A. J. Conley, E. S. Mocarski, and B. Roizman.** 1980. Cloning of reiterated and nonreiterated herpes simplex virus 1 sequences as BamHI fragments. Proc. Natl. Acad. Sci. USA **77**:4201–4205.
31. **Post, L. E., and B. Roizman.** 1981. A generalized technique for deletion of specific genes in large genomes: α gene 22 of herpes simplex virus 1 is not essential for growth. Cell **25**:227–232.
32. **Preston, V. G., A. J. Davison, H. S. Marsden, M. C. Timbury, J. H. Subak-Sharpe, and N. M. Wilkie.** 1978. Recombinants between herpes simplex virus types 1 and 2: analyses of genome structures and expression of immediate-early polypeptides. J. Virol. **28**:499–517.
33. **Rixon, F. J., and D. J. McGeoch.** 1984. A 3′ co-terminal family of mRNAs from the herpes simplex virus type 1 short region: two overlapping reading frames encode unrelated polypeptides one of which has a highly reiterated amino acid sequence. Nucleic Acids Res. **12**:2473–2487.
34. **Rixon, F. J., and D. J. McGeoch.** 1985. Detailed analysis of the mRNAs mapping in the short unique region of herpes simplex virus type 1. Nucleic Acids Res. **13**:953–973.
35. **Roizman, B., and A. E. Sears.** 1990. Herpes simplex viruses and their replication, p. 1795–1841. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Fields' virology, 2nd ed. Raven Press, New York.
36. **Roller, R. J., and B. Roizman.** 1990. The herpes simplex virus U_S11 open reading frame encodes a sequence-specific RNA-binding protein. J. Virol. **64**:3463–3470.
37. **Roller, R. J., and B. Roizman.** 1992. The herpes simplex virus 1 RNA-binding protein U_S11 is a virion component and associates with ribosomal 60S subunits. J. Virol. **66**:3624–3632.
38. **Sheldrick, P., and N. Berthelot.** 1975. Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harbor Symp. Quant. Biol. **39**:667–678.
39. **Silver, S., and B. Roizman.** 1985. γ_2-Thymidine kinase chimeras are identically transcribed but regulated as γ_2 genes in herpes simplex virus genomes and as β genes in cell genomes. Mol. Cell. Biol. **5**:518–528.
40. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
41. **Wagner, E. K.** 1985. Individual HSV transcripts: characterization of specific genes, p. 45–104. *In* B. Roizman (ed.), The herpesviruses, vol. 3. Plenum Press, New York.