The Herpes Simplex Virus Origins of DNA Synthesis in the S Component Are Each Contained in a Transcribed Open Reading Frame

JEFF HUBENTHAL-VOSS,¹ LISA STARR,² AND BERNARD ROIZMAN^{1*}

The Marjorie B. Kovler Viral Oncology Laboratories, the University of Chicago, Chicago, Illinois 60637,¹ and Syntro Corporation, San Diego, California 92121²

Received 27 May 1987/Accepted 13 July 1987

In the herpes simplex virus 1 genome, the origins of viral DNA synthesis are located in the unique sequences of the L component (Ori_L) and in the reiterated sequences of the S component (Ori_S) located between the 5' terminus of the α 4 gene and the 5' terminus of either the α 22 (left terminus of the S component) or the α 47 (right terminus of the S component) gene. Studies prompted by the finding that only one, but not both, Ori_S sequence is dispensable for growth in cell culture indicate that each Ori_S sequence is contained in an open reading frame designated as Ori_SORF. The transcription of Ori_SORF is initiated approximately 860 nucleotides upstream from that of the α 4 gene and 162 nucleotides downstream, but on the opposite strand from the transcription initiation site of the α 22 or α 47 genes within the inverted repeat c sequence. The Ori_SORF transcript is 3' coterminal with the mRNA of the α 4 gene, polyadenylated but not spliced, transported into the cytoplasm, and capable of directing the synthesis of a 330-amino-acid protein with a translated molecular weight of approximately 34,000. Transcription is cycloheximide but not phosphonoacetate sensitive and is therefore regulated as either a β or a γ_1 gene. The implications of the transcription of Ori_S and of possible functions of the product of Ori_SORF are discussed.

The herpes simplex virus 1 (HSV-1) genome consists of two components, L and S, which can invert relative to each other such that wild-type viral DNA extracted from virions or from infected cells consists of four equimolar populations differing solely in the relative orientation of the two components (4, 9). Each of the components consists of unique sequences, *UL* and *US*, flanked by inverted repeats (35). The inverted repeats of the L component were designated as *ab* and *a'b'*, whereas the inverted repeats of the S component were designated as *a'c'* and *ca* (39). The entire S component including the unique sequences and the inverted repeats has been sequenced (20, 21). Analyses of the sequence confirmed previous studies showing that the inverted repeats contain in their entirety the α 4 gene, an origin of viral DNA replication, and the promoter-regulatory domains of the α 22 and α 47 genes (19, 22, 23, 38). Analyses of the nucleotide sequence of the *US* domain of the S component revealed the presence of 12 open reading frames designated as US1, the coding sequences of the α 22 gene, through US12, the coding sequences of the α 47 gene (21). In this paper we report that the inverted repeat sequences *a'c'* and *ca* each express a hitherto unsuspected open reading frame which encompasses the S component origins (Ori_S) of viral DNA replication. The circumstances that led to the recognition of this open reading frame, designated as Ori_SORF, were as follows.

Several studies reported during the past several years have shown that the S component contains clusters of viral genes and sequences of unknown function that could be deleted without affecting viral replication in cultured cells. The deleted domains include an 800-base-pair (bp) stretch between the terminus of the α 4 gene and the terminal *a*

sequence (10), one origin of viral DNA replication (15), and all of the genes encoded in the unique sequences with the exception of the gene encoding glycoprotein D (15-17). Attempts to make insertional mutations sequentially at or near both origins of viral DNA replication were not, however, successful. In the course of analyses to determine why both origins or nearby domains cannot be interrupted, we discovered that the origins are a part of a transcriptional unit contained entirely within the inverted repeat sequences and therefore present in two copies per genome. The domain of the open reading frame containing the origin of DNA synthesis of the S component was designated Ori_SORF. Of particular interest is the observation that the Ori_SORF is contained entirely within the promoter-regulatory domains of the α genes 4, 22, and 47, and that the mRNA transcribed from Ori_SORF could direct the synthesis of a protein 330 amino acids in length.

Although there are a number of examples in which an origin of DNA replication is contained within a transcribed region, there has as yet been no evidence that this is the case for HSV-1; in fact the Ori_SORF was not deduced from analyses of the nucleotide sequence of the S component. The possible significance of this open reading frame stems from the observation that the transcription of genes encompassing an origin of replication may serve as a determinant of the onset of DNA replication (2, 7, 18, 26, 33).

MATERIALS AND METHODS

Viruses and cells. The isolation and properties of HSV-1 strain F, the procedures for maintenance and growth of Vero (African green monkey kidney) cells, and the propagation of virus stocks have been previously described (6, 32). Vero cells were used for the preparation of viral RNA.

* Corresponding author.

FIG. 1. Sequence arrangement of the HSV-1 genome showing the positions of relevant transcripts and DNA probes used in this study. (A) Schematic representation of the sequence arrangement of the HSV-1 genome. L and S refer to the long and short components of the genome which consist of unique sequences flanked by terminal inverted repeats *ab* and *b'a'*, *a'c'*, and *ca*, respectively. (B) Positions of the RNA transcripts relevant to this study. $\alpha 4$ and $\alpha 47$ or $\alpha 22$ represent the transcripts of the α genes 4, 47 and 22 which initiate in the *c* terminal repeats of the S component. The S component origins of DNA synthesis are designated Ori_s . The initiation site and terminus of the Ori_sORF transcript are also shown. Abbreviations: B, *Bam*HI; Sa, *Sall*; E, *Eco*RI; S, *Sph*I; Nc, *Nco*I; Nr, *Nru*I; H, *Hind*III. Only the sites relevant to the results are shown. The two *Bam*HI cleavage sites show the location of the *Bam*HI Y fragment mentioned in Results. (C) Positions of the DNA probes, the cDNA, and the cDNA primer described in the text on the physical map of the HSV-1 genome. The asterisk identifies the terminus of the probe C that was labeled with ^{32}P .

cDNA cloning. A synthetic DNA primer was synthesized to prime first-strand cDNA synthesis with an Applied Biosystems 380A DNA synthesizer. The position of the primer in the Ori_sORF is given in Fig. 1. For first-strand DNA synthesis 2 μg of poly(A)⁺ RNA, isolated from Vero cells late during infection, was denatured in 20 mM methyl mercuric hydroxide (Alfa, Danvers, Mass.), neutralized with 75 mM 2-mercaptoethanol (Eastman Kodak Co., Rochester, N.Y.), and then annealed with 2 μg of synthetic primer in a reaction volume of 25 μl that contained 28 U of placental ribonuclease inhibitor (Bethesda Research Laboratories, Gaithersburg, Md.), 100 mM Tris (pH 8.3 at 42°C), 140 mM KCl, 10 mM MgCl₂, 0.8 mM each of dCTP, dATP, dTTP, dGTP (Pharmacia Fine Chemicals, Piscataway, N.J.), and 100 μCi of [α - ^{32}P]dCTP (New England Nuclear Corp., Boston, Mass.). The primer was extended with 180 U of avian myeloblastosis virus reverse transcriptase (Molecular Genetics Resources, Tampa, Fla.) at 42°C for 90 min. The second strand was synthesized by a modification of the Gubler and Hoffman technique (8) with 50 μg of deoxynucleotide triphosphates, 5.4 U of DNA polymerase I (Boehringer Mannheim Biochemicals, Indianapolis, Ind.), and 100 U of *Escherichia coli* DNA ligase (New England BioLabs, Beverly, Mass.) per ml. RNase H was purchased from Bethesda Research Laboratories. Oligo(dC) tails were added to the double-stranded cDNA by using 25 U of terminal deoxynucleotidyl transferase (Molecular Genetics Resources) in a 25- μl reaction mixture that contained 100 mM cacodylate (pH 7.2), 0.2 mM dithiothreitol, 2 mM CoCl₂, and 80 μM dCTP and was incubated for 30 min at

37°C. The tailed cDNA was annealed to 200 ng of dG-tailed, *Pst*I-cleaved pBR322 (Bethesda Research Laboratories) and used to transform competent *E. coli* RR1 purchased from Bethesda Research Laboratories.

Isolation and analyses of RNA. Cytoplasmic RNA was purified as described by Jenkins and Howett (11) from mock-infected cells and cells infected at a multiplicity of 10 PFU per cell. For 5'-end analysis 50 μg of total cytoplasmic RNA was hybridized to approximately 0.02 pmol of 5' end-labeled HSV-1 DNA fragment and digested with S1 nuclease as described previously (11). For identification of RNA homologous to the probes, 30 μg of total cytoplasmic RNA extracted from mock-infected or infected cells was separated by electrophoresis in a 1.4% SeaKem ME agarose gel (FMC Corp., Marine Colloids Div., Rockland, Maine) that contained 3% (wt/vol) formaldehyde (Fisher Scientific Co., Fair Lawn, N.J.) and transferred to a nitrocellulose sheet (Schleicher & Schuell Co., Keene, N.H.) in the presence of 3 M NaCl by capillary action. Hybridization probes were labeled to high specific activity with a kit obtained from New England Nuclear Corp. by procedures recommended by the manufacturer, and the hybridization conditions were as described elsewhere (11).

DNA sequence analysis. For DNA sequence analysis, cDNA fragments were subcloned into M13 mp18 (40). The sequence was determined by extending a ^{32}P -labeled 5'-end primer fragment (New England BioLabs) in the presence of dideoxynucleotide triphosphates with avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim Biochemicals).

FIG. 2. Autoradiographic images of labeled HSV-1 DNA probes hybridized to cytoplasmic RNA electrophoretically separated in denaturing gels and transferred to a nitrocellulose sheet. The RNAs were extracted from the cytoplasm of mock-infected and infected cells as described in the text. The location of the DNA probes on the physical map of the HSV-1 genome is shown in Fig. 1.

Computer analysis of DNA sequence. Open reading frame usage probabilities were determined by the positional base preference method of Staden (36). Computations were done on a Hewlett Packard 9000 minicomputer.

RESULTS

Ori_sORF RNA is present in HSV-1 cells at 14 h postinfection. As shown diagrammatically in Fig. 1, the two copies of Ori_s map between the 5' termini of the $\alpha 4$ genes and those of either the $\alpha 22$ or $\alpha 47$ gene corresponding to the left or right reiterated sequence *c* of the S component, respectively. To determine whether the sequences 5' to the transcription initiation site of the $\alpha 4$ gene are transcribed, probes derived from the domains of the $\alpha 4$ gene were hybridized to 30- μ g samples of total cytoplasmic RNA electrophoretically separated in the presence of 3% (vol/vol) formaldehyde and then transferred to a nitrocellulose sheet. The cytoplasmic RNAs tested were extracted from Vero cells infected and maintained for 4 h in the presence of 50 μ g of cycloheximide per ml and from untreated cells either mock infected or 14 h postinfection. Probe A was a 450-bp *Sall*-*Sma*I DNA fragment that contains portions of both the transcribed noncoding and coding sequence and promoter-regulatory region of the $\alpha 4$ gene, whereas probe B was a 220-bp *Eco*RI-*Sma*I fragment that contains only the 5' nontranscribed, regulatory sequence of the $\alpha 4$ gene. Probe A hybridized to a band formed by RNA extracted from cells infected and maintained in the presence of cycloheximide and from untreated cells harvested 14 h postinfection (Fig. 2). The electrophoretic mobility of this band corresponded to that of the $\alpha 4$ mRNA. In contrast, probe B hybridized to a transcript similar in size to that of the $\alpha 4$ gene but present only in RNA isolated from cells at 14 h postinfection. This experiment suggested the possibility that either the transcript hybridizing with probe B represented an mRNA of an as yet unknown gene that was 3'

coterminous with the mRNA of the $\alpha 4$ gene or the transcription initiation site of the $\alpha 4$ gene late in infection was further upstream from that utilized early in infection. Additional S1 analyses with a 3' end-labeled fragment at the *Nco*I site (Fig. 1) indicated that the transcript extended beyond the stop codons present in the *Bam*HI Y fragment 5' to the transcription initiation site of the $\alpha 4$ gene (data not shown). This finding is consistent with the observation that probe B hybridized with an RNA species long enough to encompass both the Ori_sORF and the $\alpha 4$ open reading frames.

Transcription initiation site of the Ori_sORF RNA. Figure 3 shows the results of S1 nuclease tests designed to identify the initiation site for the RNA transcript extending upstream from the 5' terminus of the $\alpha 4$ gene designated for convenience as Ori_sORF RNA. To map the transcription initiation site of the Ori_sORF RNA, 50- μ g amounts of total cytoplasmic RNAs extracted from mock-infected Vero cells, cells infected and maintained for 14 h in the presence of phosphonoacetate (300 μ g/ml), a specific inhibitor of viral DNA synthesis, and from untreated cells at 14 h postinfection

FIG. 3. Determination of the transcription initiation site of the Ori_sORF RNA transcript. (A) Results of S1 nuclease digestion of cytoplasmic RNAs hybridized to probe C (Fig. 1) and electrophoretically separated under denaturing conditions. The RNAs were prepared as described text and are labeled as follows: 14HR, extracted from untreated cells 14 h postinfection; PAA, extracted from cells infected and maintained for 14 hours in the presence of phosphonoacetate (300 μ g/ml of medium); MOCK, extracted from mock-infected cells. The lane A+G contains pooled A terminated and G terminated dideoxy sequencing reactions. PROBE indicates undigested probe C DNA. (B) Autoradiographic image of a 14-h RNA-probe C hybrid digested with S1 as described above but subjected to electrophoresis under nondenaturing conditions. The lanes labeled as pUC in panels A and B contain 5' end-labeled fragments obtained from *Hpa*II-digested pUC19.

FIG. 4. Position of methionine and termination codons, and positional base preference analysis of the three reading frames of Ori_sORF. The figure shows a plot of the positional base value as a function of base position calculated by the method of Staden (36) for each of the three reading frames. Nucleotide 1 is the 5'-terminal nucleotide of the cDNA described in the text. The last methionine codon shown in frame one is the translation initiation site of the $\alpha 4$ protein. Dotted lines indicate the position of a 50% coding probability. Symbols: |, positions of stop codons; +, position of methionine codons. The domain of the Ori_s sequence defined by the arrows in frame 1 indicates the position of the smallest sequence shown (37) to be able to function as an origin of DNA synthesis.

tion were each hybridized to probe C DNA uniquely 5' end labeled at the *Nru*I site (Fig. 1) and then digested with S1 nuclease as described in Materials and Methods. The digests were electrophoretically separated in a 6% polyacrylamide gel in the presence of 8 M urea. By using the A+G sequencing ladder and the end labeled DNA fragments as size standards, the start site of transcription was determined to be in the 5' transcribed noncoding domain of the $\alpha 22$ or $\alpha 47$ genes, but on the opposite strand. Furthermore, the apparent abundance of probe protected by RNA isolated in the presence of phosphonoacetate suggests that transcription of this region may be insensitive to PAA.

Ori_sORF RNA is not spliced. Figure 3B shows the autoradiographic image of an S1 digest prepared and electrophoretically separated as above, except that 8 M urea was omitted from the gel and the samples were not subject to denaturing conditions before electrophoresis. Since the protected portion of the probe as determined by electrophoresis under these conditions was identical in size, relative to the marker DNA, to the protected fragment seen under denaturing conditions, it may be concluded that the Ori_sORF transcription start site is as described above and does not start within a noncontiguous DNA sequence.

Transcription initiation site of Ori_sORF RNA based on the sequence of a cDNA clone made from poly(A)⁺ cytoplasmic RNA. A cDNA library was constructed as described in Materials and Methods from approximately 2 μ g of cytoplasmic RNA and twice selected over oligo(dT)-cellulose. The position of a synthetic oligonucleotide which was used to prime first-strand synthesis is shown in Fig. 1. Of the approximately 1,000 colonies screened by hybridization with probe B, 3 were positive. The insert from one positive clone was recloned in M13 mp18 and sequenced by a modification of the chain-termination technique (34). The results of the sequence analysis placed the transcription initiation site of Ori_sORF within the 5'-transcribed noncoding sequences of

the $\alpha 47$ and $\alpha 22$ genes 162 nucleotides from their transcription initiation sites, and we refer to this as the transcription initiation site for Ori_sORF RNA. S1 analyses on hybrids of genomic and cDNA probes cleaved and labeled at the same *Nru*I site (Fig. 1) with cytoplasmic RNA of cells harvested 14 h postinfection yielded comigrating protected DNA bands, consistent with the conclusion that the cDNA clone contains the 5' terminus of the transcript (data not shown). This would place the initiation site 860 nucleotides upstream from the $\alpha 4$ transcription initiation site.

Characteristics of ORI_sORF. The transcription initiation site and the 5' domains of the Ori_sORF overlap, but in the opposite sense, the transcribed domains of the $\alpha 22$ and $\alpha 47$ genes. The transcription initiation site is approximately 25 bp from the sequence TGTGTT, the nearest analog to a TATAA box. The three open reading frames (Fig. 4) that could be used by Ori_s have the following characteristics. (i) The first methionine codon is in frame 1, 25 bp downstream from the transcription initiation site of Ori_sORF. The coding capacity of frame one is a protein 330 amino acids in length, corresponding to 34,000 in translated molecular weight. The Ori_s sequence is contained within the coding sequence of this open reading frame. (ii) The first methionine codons in frames 2 and 3 are approximately in the middle of the Ori_sORF sequence. They are preceded by at least two methionine codons in frame 1, and because of stop codons their coding capacity is relatively small and does not include the Ori_s sequence. Positional base preference analysis done as described by Staden (36) also suggests that open reading frame 1 is most likely to code for a protein. It is noteworthy that the Ori_s sequences represent the only domain of reading frame 1 in which the positional base preference value is below the predicted value for coding domains. This, however, may be expected inasmuch as the base utilization within the Ori_s sequence may be severely restricted by its functional requirements.

DISCUSSION

Characteristics of the Ori_sORF, its transcript, and its predicted protein product. The studies described in this report have established the following.

(i) The nucleotide sequences located within the repeated sequences *a'c'* and *ca* flanking the unique sequences of the S component each contain an open reading frame designated as Ori_sORF which encompasses an origin of DNA replication. This open reading frame is located between the transcription initiation sites of the α genes 22 or 47 and the transcription initiation site of the $\alpha 4$ genes.

(ii) Ori_sORF is transcribed later in infection than the α genes between which it is sandwiched. The transcription initiates from the transcribed noncoding sequences of the $\alpha 22$ or 47 genes approximately 162 bp from the transcription initiation sites of these genes. The site of termination of this transcript was not specifically determined, but from the size of the transcript we deduced that it is coterminal with that of the $\alpha 4$ gene. Analyses of the published coding sequences of the $\alpha 4$ gene revealed no other polyadenylation signals other than those used by the $\alpha 4$ gene (20).

(iii) S1 analyses indicate that the transcript of the Ori_sORF is not spliced. The open reading frame could encode a protein of a maximum size of 330 amino acids with a translated molecular weight of 34,000. The coding sequence contains the domain of Ori_s.

Implications of Ori_sORF transcription with respect to the regulation of expression of the $\alpha 4$ genes. The discovery of the transcribed open reading frame encompassing the promoter-regulatory domains of the coding strand of the $\alpha 4$ gene and of the noncoding strands of the $\alpha 22$ and $\alpha 47$ genes has implications with respect to the regulation of the α genes. Available evidence indicates that the α genes are regulated both positively and negatively. Specifically, the α genes are *trans*-activated by the α transinducing factor, a structural component of the virion (1, 3, 24, 27, 29) and are turned off by a function that involves the $\alpha 4$ gene product (23, 28, 30). Recent studies have shown that the $\alpha 4$ protein binds to promoter-regulatory domains of α genes and specifically to the regulatory domain and across the transcription initiation site of the $\alpha 4$ gene (12–14; T. M. Kristie, N. Michael, D. Specter, and B. Roizman, in *Mechanisms of Control of Gene Expression*, in press). It has been postulated that the binding of the $\alpha 4$ protein at the cap site blocks the transcription of the $\alpha 4$ gene (14). The problem stems from the results reported in this paper that the polyadenylated RNA homologous to the Ori_sORF and $\alpha 4$ genes are 3' coterminal and that the Ori_sORF transcription is regulated as either a β or a γ gene. These observations suggest either that the transcription of the $\alpha 4$ gene is shut off simultaneously with that of the Ori_sORF or that shutoff of $\alpha 4$ gene transcription does not occur by occlusion of transcription by factors binding to the viral DNA, inasmuch as this mechanism would also turn off the transcription of Ori_sORF. If in fact the transcription of Ori_sORF is asynchronous from that of the $\alpha 4$ gene, the repression of $\alpha 4$ specific transcription would have to be mediated by a mechanism that prevents the formation of a transcriptionally active complex on the $\alpha 4$ promoter-regulatory element rather by physically blocking the movement of the polymerase complex across the $\alpha 4$ gene transcription initiation site.

The significance of the inclusion of Ori_s in a transcribed open reading frame. The significant aspects of this study concern the potential role of Ori_sORF in viral DNA synthesis. The relevant issues are as follows.

(i) Of the three origins of viral DNA synthesis, the single origin identified in the L component is dispensable without affecting viral growth (25). Of the two identical origins of the S component, one but not both is dispensable for growth in cell culture (15). The evidence presented in this report is that the origin is contained in an open reading frame. Because we were not able to mutagenize by insertional mutagenesis the domain of the Ori_sORF at some distance from the mapped "minimal origin" (37) in a virus containing a single origin of the S component, at least one copy of the Ori_sORF appears to be essential for viral replication. These studies leave unanswered the question of whether the three mapped origins are freely interchangeable or whether at least one origin of the S component must be present for viral DNA replication.

(ii) Transcription has been functionally associated with the regulation of a number of procaryotic and eucaryotic replicons. The effect of local transcription or transcriptional activation can be either positive or negative depending on the replication origin, but it is clear that regulation of origin activity by this means is common if not ubiquitous.

A pertinent example in which replication origin-associated transcription negatively regulates the activity of an origin is the γ origin of the bacterial plasmid R6K (26). In this instance, transcription from a promoter designated P_{sil} either competes for authentic primer formation or disrupts the structure of *ori γ* . It is noteworthy that the effect of the P_{sil} promoter is polar; it acts only in *cis*, indicating that it is transcription per se and not a diffusible product that regulates replication. R6K, like HSV, has three sequences capable of initiating DNA replication, but the activity of *ori γ* is seldom detected in the presence of R6K *ori α* and β . Which HSV origin of replication is used in vivo is not known, but it is interesting to speculate that the virus may use different origins depending on its environment or physiologic requirements.

In contrast to the R6K γ origin, the origin of bacteriophage λ DNA synthesis is stimulated by local transcription (7). Mutations that block rightward transcription across the λ origin region are *cis* dominant but can be suppressed by *ri^c* mutations which permit constitutive rightward transcription from the right side of the λ origin. These findings suggest that in addition to viral and host replication factors there is a requirement for transcriptional activation of DNA sequences near the λ origin to permit its function. Furthermore, in many respects the HSV Ori_s domain is similar to that of the bacteriophage λ . Specifically, both replication origins lie within transcribed open reading frames, the λ origin within the *O* gene and Ori_s within Ori_sORF, and the expression of each of these genes is potentially affected by transcription of the opposite DNA strand. The product of the *O* gene has been demonstrated to be a 34,000 molecular weight, λ origin-specific DNA binding protein.

Among eucaryotic viruses there have been a number of demonstrations of replication origin-associated transcription and its relevance to the regulation of DNA synthesis. The genome of bovine papillomavirus, for example, contains two homologous DNA sequences capable of supporting DNA replication designated plasmid maintenance sequences (PMS) I and II (18). PMS I is functional in wild-type bovine papillomavirus genomes, whereas PMS II functions as an origin of replication in bovine papillomavirus genomes that contain mutations in PMS I. PMS II lies within the E1 open reading frame of the bovine papillomavirus genome; the 5' portion has been shown to act as a negative regulator of DNA replication initiated at PMS I, and the 3' portion of the

E1 reading frame has been demonstrated to have a positive regulatory effect on bovine papillomavirus replication (2). Likewise, the plasmid replication origin of the Epstein-Barr virus, a member of the herpesvirus family, is contained within a complex transcriptional unit that gives rise to the EBNA-1 protein (33). As is the case for bacteriophage λ , the product of the Epstein-Barr virus origin-associated transcript is required for the function of the replication origin (31).

These relationships serve to exemplify precedents for encoding *trans*-acting regulatory factors for DNA replication near the replication origin at which they act. Although we present no evidence that this may be the case for the translation product of the HSV Ori_sORF, it is noteworthy that Elias et al. (5) have identified an Ori_s-binding protein activity in infected cell extracts.

(iii) Evidence that the transcription of Ori_sORF is not a *cis* requirement for HSV DNA replication is suggested by the report that a plasmid containing as little as 89 bp of the S component origin resident in cells is amplified in *trans* by superinfection with HSV (37). These studies do not conclusively exclude a *cis*-acting function to the transcription of Ori_sORF, since host or plasmid factors may complement that function. The experiments do not exclude, however, the possibility that the putative protein encoded by Ori_sORF acts in *trans* or provides a regulatory function in a natural host.

Studies on the product and function of the Ori_sORF are in progress.

ACKNOWLEDGMENTS

These studies were supported by Public Health Service grants CA-08494 and CA-19264 from the National Cancer Institute and by grant MV-2T from the American Cancer Society. J.H.-V. is a predoctoral trainee receiving support from the Training Program in Genetics and Regulation (GM 07197).

LITERATURE CITED

- Batterson, W., and B. Roizman. 1983. Herpes simplex virus virion-associated factor responsible for the induction of α genes. *J. Virol.* **46**:371-377.
- Berg, L., M. Lusky, A. Stenlund, and M. R. Botchan. 1986. Repression of bovine papilloma virus replication is mediated by a virally encoded *trans*-acting factor. *Cell* **46**:753-762.
- Campbell, M. E. M., J. W. Palfreyman, and C. M. Preston. 1984. Identification of herpes simplex virus sequences which encoded a *trans*-acting polypeptide responsible for stimulation of immediate early transcription. *J. Mol. Biol.* **180**:1-19.
- Delius, S. J., and J. B. Clements. 1976. A partial denaturation map of herpes simplex virus type 1 DNA: evidence for inversions of the unique DNA regions. *J. Gen. Virol.* **33**:125-134.
- Elias, P., M. E. O'Donnell, E. Mocarski, and I. R. Lehman. 1986. A DNA binding protein specific for an origin of replication of herpes simplex virus type 1. *Proc. Natl. Acad. Sci. USA* **83**:6322-6326.
- Ejercito, J. M., E. D. Keiff, and B. Roizman. 1968. Characterization of herpes simplex virus strains differing in their effects on social behavior of infected cells. *J. Gen. Virol.* **2**:357-364.
- Furth, E. F., and S. E. Wickner. 1985. Lambda DNA replication, p. 145-170. In R. W. Hendrix, J. W. Robert, F. W. Stahl, and R. Weisberg (ed.), *Lambda II*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Gubler, U., and B. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene* **25**:263-269.
- 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 segments. *Proc. Natl. Acad. Sci. USA* **72**:4243-4247.
- Hubenthal-Voss, J., and B. Roizman. 1985. Herpes simplex virus 1 reiterated S component sequences (*c_i*) situated between the *a* sequence and $\alpha 4$ gene are not essential for virus replication. *J. Virol.* **54**:509-514.
- Jenkins, F. J., and M. K. Howett. 1984. Characterization of the mRNAs mapping in the *Bg*/II N fragment of the herpes simplex virus type 2 genome. *J. Virol.* **52**:99-107.
- Kristie, T. M., and B. Roizman. 1986. $\alpha 4$, the major regulatory protein of herpes simplex virus type 1, is stably and specifically associated with promoter-regulatory domains of α genes and of selected other viral genes. *Proc. Natl. Acad. Sci. USA* **83**:3218-3222.
- Kristie, T. M., and B. Roizman. 1986. DNA-binding site of major regulatory protein $\alpha 4$ specifically associated with promoter-regulatory domains of α genes of herpes simplex virus type 1. *Proc. Natl. Acad. Sci. USA* **83**:4700-4704.
- Kristie, T. M., and B. Roizman. 1987. *Cis*-acting elements and *trans*-acting factors involved in the expression and regulation of the α genes of herpes simplex virus type 1, p. 161-174. In N. O. Kjeldgaard and J. Forchhammer (ed.), *Viral carcinogenesis: functional aspects*. Munksgaard, Copenhagen.
- Longnecker, R., and B. Roizman. 1986. Generation of an inverting herpes simplex virus 1 mutant lacking the L-S junction *a* sequences, an origin of DNA synthesis, and several genes including those specifying glycoprotein E and the $\alpha 47$ gene. *J. Virol.* **58**:583-591.
- Longnecker, R., and B. Roizman. 1986. Clustering of genes dispensable for growth in culture in the S component of the HSV-1 genome. *Science* **236**:573-576.
- Longnecker, R., and B. Roizman. 1987. Identification of a new glycoprotein gene of herpes simplex virus 1 within a gene cluster dispensable for growth in cell culture. *Proc. Natl. Acad. Sci. USA* **84**:4303-4307.
- Lusky, M., and M. Botchan. 1984. Characterization of bovine papilloma virus plasmid maintenance sequences. *Cell* **36**:391-401.
- Mackem, S., and B. Roizman. 1980. Regulation of herpes virus macromolecular synthesis: transcription initiation sites and domains of α genes. *Proc. Natl. Acad. Sci. USA* **77**:7122-7126.
- McGeoch, D. J., A. Dolan, S. Donald, and H. K. Bauer-Dieter. 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. *Nucleic Acids Res.* **14**:1727-1745.
- McGeoch, D. J., A. Dolan, S. Donald, and F. 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.
- Mocarski, E. S., and B. Roizman. 1982. Herpes virus-dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral *a* sequences and linked to an origin of viral DNA replication. *Proc. Natl. Acad. Sci. USA* **79**:5626-5630.
- Morse, L. S., L. Pereira, B. Roizman, and P. A. Schaffer. 1978. Anatomy of herpes simplex virus (HSV) DNA. X. Mapping of viral genes by analysis of polypeptides and functions specified by HSV-1 $\times$ HSV-2 recombinants. *J. Virol.* **26**:389-410.
- O'Hare, P., and G. S. Hayward. 1985. Three *trans*-acting regulatory proteins of herpes simplex virus modulate immediate early gene expression in a pathway involving positive and negative feedback regulation. *J. Virol.* **56**:723-733.
- Orberg, P. K., and P. A. Schaffer. 1987. Expression of herpes simplex type 1 major DNA-binding protein, ICP8, in transformed cell lines: complementation of deletion mutants and inhibition of wild-type virus. *J. Virol.* **61**:1136-1146.
- Pate, I., and D. Bastia. 1986. A replication origin is turned off by an origin "silencer" sequence. *Cell* **47**:785-795.
- Pellet, P. E., J. L. C. McKnight, F. 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.
- Pizer, L. I., D. G. Tedder, J. L. Betz, K. W. Wilcox, and P. Beard. 1986. Regulation of transcription in vitro from herpes

- simplex virus genes. *J. Virol.* **60**:950-959.
29. **Post, L. E., S. Mackem, and B. Roizman.** 1981. The regulation of α genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. *Cell* **24**:555-565.
30. **Preston, C. M.** 1979. Control of herpes simplex virus 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *tsK*. *J. Virol.* **29**:275-284.
31. **Reisman, D., J. Yates, and B. Sugden.** 1985. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two *cis*-acting components. *Mol. Cell. Biol.* **5**:1822-1832.
32. **Roizman, B., and P. G. Spear.** 1968. Preparation of herpes simplex virus of high titer. *J. Virol.* **2**:83-84.
33. **Sample, J., M. Hummel, D. Braun, M. Birkenbach, and E. Kieff.** 1986. Nucleotide sequence of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcription initiation site. *Proc. Natl. Acad. Sci. USA* **83**:5096-5100.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
35. **Sheldrick, P., and N. Berthelot.** 1975. Inverted repetitions in the chromosome of herpes simplex virus. *Cold Spring Harbor Symp. Quant. Biol.* **39**:667-678.
36. **Staden, R.** 1984. Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. *Nucleic Acids Res.* **12**:551-567.
37. **Stow, N. D.** 1982. Localization of an origin of DNA replication within the *TR_s/IR_s* repeated region of the herpes simplex virus type 1 genome. *EMBO J.* **1**:863-867.
38. **Vlazny, D. A., and N. Frenkel.** 1981. Replication of herpes simplex virus DNA: location of replication recognition signals within defective virus genome. *Proc. Natl. Acad. Sci. USA* **78**:742-746.
39. **Wadsworth, S., G. S. Hayward, and B. Roizman.** 1976. Anatomy of herpes simplex virus DNA. V. Terminally repetitive sequences. *J. Virol.* **17**:503-512.
40. **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**:103-109.