An Unusual Spliced Herpes Simplex Virus Type 1 Transcript with Sequence Homology to Epstein-Barr Virus DNA

R. H. COSTA,[1] K. G. DRAPER,[1] T. J. KELLY,[2] AND E. K. WAGNER[1]*

Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, California 92717,[1] and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205[2]

High-resolution transcription mapping localized a spliced 2.7-kilobase herpes simplex virus type 1 mRNA. The 4-kilobase intron of this transcript encodes a nested set of transcripts on the opposite DNA strand. The nucleotide sequence of the DNA encoding the left-hand and right-hand exons of the spliced transcript was determined, and the salient features are presented here. Of major interest is that both exons contained regions within several hundred bases of the splice donor and acceptor sites which showed homology to two regions of the Epstein-Barr virus genome, which are themselves 3 kilobases apart. The spliced herpes simplex virus transcript encoded a translational reading frame which could encode a protein with an approximate size of 75,000 daltons. This value is in agreement with in vitro translation data. The predicted amino acid sequence of the herpes simplex virus protein had significant homology with putative amino acid sequences encoded by the homologous Epstein-Barr virus DNA sequences.

As recently reviewed (24; E. Wagner, in B. Roizman, ed., *Herpesviruses*, vol. 3, in press), correlation between the genetic maps and transcription maps for herpes simplex virus type 1 (HSV-1) is generally excellent. Detailed transcription maps for the virus indicate that the vast majority of transcripts are unspliced and have promoter-control regions immediately 5' to the mRNA cap sites. Nested sets of transcripts can be identified. These are partially overlapping transcripts which can share polyadenylation sites yet encode unique polypeptides. Other partially overlapping clusters have been identified which have a common 5' cap site but which differ in the polyadenylation site. Such a situation leads to mRNAs which are "redundant" in that they each encode the same protein. Variations on these patterns are common.

Splicing of mRNA is not common in HSV-1 mRNA metabolism. The reason for this is unclear, since other herpesviruses encode spliced mRNAs. Furthermore, some spliced mRNAs are expressed during HSV infection. In HSV-1 infection, most splices are short. It was therefore surprising to note the occurrence of a late HSV-1 spliced mRNA which is encoded by a DNA sequence containing a 4-kilobase (kb) intron. This transcript is located in a region (0.185 to 0.225 map units [m.u.]) with no reported genetic markers (25; P. Schaffer, personal communication). As described in this communication, analysis of the DNA sequence encoding the transcript showed that it had significant homology with an analogous region encoded by Epstein-Barr virus (EBV) DNA, the sequence of which has been reported by Baer et al. (3) Such homology is compelling evidence that this unusual (for HSV) spliced transcript encodes a function important for the replication of herpesviruses in general.

MATERIALS AND METHODS

Cells and virus. For RNA preparation, plaque-purified isolates of the KOS strain of HSV-1 were used to infect

HeLa cells. Monolayer cultures of HeLa cells were grown at 37°C in Eagle minimal essential medium containing 10% calf serum, penicillin, and streptomycin.

Enzymes. All restriction enzymes and bacterial alkaline phosphatase were obtained from Bethesda Research Laboratories. Digestions were carried out in buffers recommended by the supplier. Bacteriophage T4 polynucleotide kinase (Bethesda Research Laboratories) was used for 5' end labeling as described by Maxam and Gilbert (20). *Escherichia coli* DNA polymerase I (Klenow fragment; Boehringer-Mannheim Biochemicals) was used to generate 3'-end-labeled DNA by the method of Maniatis et al. (19).

Isolation, labeling, and size fractionation of polyribosomal RNA. Monolayer cultures of HeLa cells (2 × 10^7 cells per flask) were infected for 30 min at a multiplicity of 10 PFU of virus per cell in phosphate-buffered saline containing 0.1% glucose and 1.0% fetal calf serum. Polyribosomes were isolated from the cytoplasm of HSV-1-infected cells by the magnesium precipitation method of Palmiter (22). Polyadenylic acid-containing [poly(A)] mRNA was isolated from total rRNA by oligodeoxythymidylic acid-cellulose (Collaborative Research, Inc.) chromatography. This is referred to as HSV poly(A) mRNA. Details of this procedure have been presented elsewhere (1, 2, 10–14, 16). RNA was isolated at 6 h postinfection. RNA was size-fractionated by electrophoresis on 1.4% agarose gels containing 10 mM methylmercury hydroxide (4) as previously described (1, 2, 14, 16).

Recombinant DNA. All recombinant DNA clones described in this paper were derived from either BamHI-*Hind*III fragment A-IO (0.151 to 0.182 m.u.), *Hind*III fragment J (0.182 to 0.262 m.u.), *Hind*III-BamHI fragment J-A (0.182 to 0.223 m.u.), or *Bgl*II fragment P (0.201 to 0.233 m.u.) of the KOS strain of HSV-1 cloned in pBR322. Several subclones were used: XhoI-*Hind*III fragment C'-IO (0.171 to 0.182 m.u.), *Hind*III-*Eco*RI fragment J-D (0.182 to 0.190 m.u.), *Eco*RI-*Sal*I fragment G-D (0.190 to 0.195 m.u.), *Eco*RI-*Xho*I fragment G-U (0.190 to 0.207 m.u.), *Sal*I fragment I' (0.195 to 0.202 m.u.), *Sal*I fragment X (0.202 to 0.219 m.u.), XhoI-BamHI fragment V-A (0.207 to 0.223 m.u.), *Sal*I-BamHI fragment U-A (0.219 to 0.223 m.u.), BamHI-*Xho*I

* Corresponding author.

fragment A-V (0.223 to 0.226 m.u.), XhoI-BglII fragment L-P (0.226 to 0.233 m.u.), and SalI-XhoI fragment U-V (0.219 to 0.226 m.u.). Procedures for cloning HSV-1 DNA fragments in the pBR322 vector have been described previously (1, 7). Cloned DNA fragments were named as described previously and located by their map coordinates on the prototype arrangement of the HSV-1 genome (7).

In situ Northern RNA blots. Unless noted otherwise, 10-μg samples of HSV poly(A) mRNA were fractionated on methylmercury gels and dried onto Whatman 3 MM paper under vacuum as previously described (10, 14, 16). The agarose film was floated off the paper in water and hybridized with appropriate nick-translated ^{35}P-labeled DNA probes in 50% formamide containing 0.4 M Na$^+$, 0.1 M HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid), pH 8.0, 0.005 M EDTA, and Denhardt solution (9) at 50°C for 36 h. Blots were rinsed at 50°C. The first two rinses were in 50% formamide–2× SSC (1× SSC = 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate (SDS). The last rinse was in 0.1× SSC–0.1% SDS. Autoradiography was on Kodak XRP film with or without intensifying screens as needed.

In vitro ^{32}P-labeled DNA was made by nick-translating appropriate DNA clones with DNA polymerase I, DNase I (Boehringer-Mannheim), and 50 μCi of [α-^{32}P]dCTP (3,000 Ci/mmol; Amersham Corp.).

Isolation of restriction fragment-specific mRNA. Restriction fragment-specific mRNA was isolated from HSV poly(A) mRNA by preparative hybridization to the appropriate DNA covalently coupled to cellulose. Details of coupling DNA to cellulose and preparative hybridization have been described previously (2, 7).

Nucleotide sequencing. As described previously (10, 14), nucleotide sequence analysis was carried out by the method of Maxam and Gilbert (20).

Nuclease mapping of HSV-1 mRNA. S1 nuclease and exonuclease VII analysis of RNA was carried out essentially as described by Berk and Sharp (5) and as described previously (1, 6–8, 10–14, 16). Appropriate HSV-1 DNA clones (10 μg) were cleaved at the desired site with the appropriate restriction enzyme. The DNA then was 5' end labeled with [γ-^{32}P]ATP (3,000 Ci/mmol; ICN) with polynucleotide kinase (Bethesda Research Laboratories) to a specific activity of 100,000 cpm/μg of DNA. Alternatively, the DNA was 3' end labeled to the same specific activity by using DNA polymerase I (Klenow fragment) (Boehringer-Mannheim).

The DNA fragments were then denatured and strand separated on 5% acrylamide gels as described by Maxam and Gilbert (20). The strand-separated DNA (from 10 μg of cloned DNA) was hybridized with 10 μg of infected-cell mRNA in 0.1 M Na$^+$–0.1 M Herpes (pH 8.0)–0.01 M EDTA at 65°C for 6 to 16 h in a 30-μl volume. Hybrids were subjected to S1 nuclease (Boehringer-Mannheim) or exonuclease VII (Bethesda Research Laboratories) digestion as described previously (10, 14). Material was fractionated on a denaturing 5% acrylamide gel, with 5'-end-labeled HinfI-digested pBR322 DNA fragments used as size standards. Alternatively, S1-protected products were fractionated on a 1.5% alkaline agarose gel, with 5'-end-labeled HindIII-digested lambda DNA fragments used as size standards (14, 16).

In vitro translation. Translation of size-fractionated viral mRNA was carried out in vitro with a micrococcal nuclease-treated rabbit reticulocyte system (New England Nuclear Corp.), with [^{35}S]methionine (>800 Ci/mmol) as the radio-active amino acid. Details of the procedure and fractionation of polypeptides in SDS-acrylamide gels by the method of Laemmli (18) have been described in several previous papers (8, 10, 14, 16). Gels were treated with En^{3}Hance (New England Nuclear Corp.) and dried under vacuum at 60°C, and radioactive bands were localized by autoradiography with Kodak XRP film. Exposure was for 3 to 5 days at −70°C.

RESULTS

Location of transcripts between 0.15 and 0.27 m.u. The map (Fig. 1) illustrates pertinent restriction endonuclease sites and the location of HSV-1 transcripts in the 18-kilobase-pair region between 0.151 and 0.272 m.u. Characterization of the transcripts between 0.16 and 0.185 m.u. and between 0.225 and 0.27 m.u. is also shown (6, 8). The preliminary characterization of the transcripts between 0.185 and 0.225 m.u. is the subject of this paper. Temporal classification of individual transcripts was based on the measured abundance of transcripts in the presence and absence of HSV-1 DNA synthesis inhibitors as described previously (17). These data are not shown here.

Individual cloned HSV-1 DNA fragments which subdivided the region of interest were used as probes for in situ RNA hybridization (Fig. 2). The sizes of transcripts were calculated from nucleotide sequence data (see below; Draper and Wagner, unpublished data). These do not include the length of poly(A) tails (ca. 200 bases [23]). Of particular interest is that a transcript of ca. 2.7 kb was seen when probes mapping between 0.19 and 0.195 m.u. and between 0.22 and 0.23 m.u. were used, but not when probes in the 4-kb region, between 0.195 and 0.22 m.u., were used. Each probe that hybridized to the 2.7-kb transcript extended only about 1,200 to 1,300 base pairs on either side of the 4-kb segment of DNA. The 4-kb region of DNA between 0.195 and 0.22 m.u. encoded a group of nested transcripts 1.5, 3.65, and 3.8 kb in length. The two larger transcripts were not readily resolvable (Fig. 2), but were clearly resolved in other in situ RNA hybridization experiments and exposures (not shown). Data from hybridization experiments with smaller probes spanning the region between 0.195 and 0.22 m.u. indicated that the 1.5-kb transcript was confined to the left-hand half of this region. Thus, no evidence for noncontiguous portions of this mRNA was seen (data not shown).

In vitro translation products of the 1.5-, 2.7-, and 3.65-3.8-kb transcripts. The transcripts of interest all mapped with BglII fragment P (0.201 to 0.233 m.u.). This fragment was used previously to isolate HSV-1 transcripts for in vitro translation (see Fig. 4a in reference 7). These earlier data indicated that the 3.65-3.8-kb transcript encodes an 80,000-dalton (Da) polypeptide, the 2.7-kb transcript encodes one or two polypeptides between 77,000 and 85,000 Da, and the 1.5-kb transcript uniquely encodes a 40,000-Da polypeptide. These results were confirmed by using various cloned HSV-1 DNA fragments that mapped between 0.19 and 0.225 m.u. Since the data fully confirm earlier work, they are not shown. Hybrid selection was used to isolate the transcripts, and this was followed by in vitro translation as described above. In another experiment (not shown), the 2.7-kb mRNA was isolated by using either EcoRI-SalI fragment G-D (0.190 to 0.195 m.u.) or BamHI-XhoI fragment M-V (0.222 to 0.225 m.u.), and each sample was translated separately. Both isolates yielded the 75,000- and 85,000-Da polypeptides. The 3.65-3.8-kb transcript encoded an 80,000-Da polypeptide, and the 1.5-kb transcript unambiguously encoded a 40,000-Da translation product.

Map Units   0.151   0.164   0.182 0.190 0.202   0.222   0.241   0.258   0.272
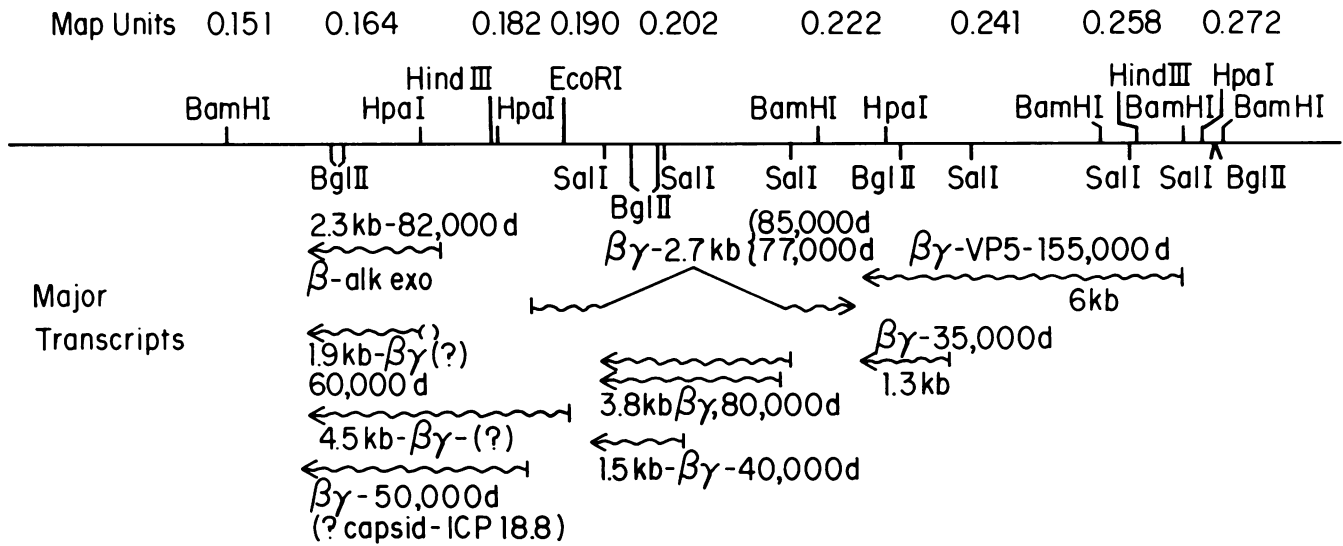
Major

Transcripts

FIG. 1. High-resolution transcription map of HSV-1 KOS in the region between 0.151 and 0.272 m.u. Locations of selected restriction endonuclease sites on the prototype arrangement of the HSV-1 genome are shown. The position in map units is shown for several restriction sites (0.01 m.u. is equivalent to 1,500 base pairs). Locations of the mRNA transcripts within the region are illustrated with respect to restriction endonuclease sites. Arrows represent the 3' ends, vertical lines represent the cap sites, and the caret delineates the position of the 2.7-kb spliced mRNA intron. The indicated mRNA sizes do not include the poly(A) tail. The molecular sizes of the in vitro translation products for each mRNA were determined by comigration with size standards on SDS-9% acrylamide gels. Identification of transcripts mapping between 0.16 and 0.185 m.u. was described previously (8), as was characterization of the transcripts between 0.225 and 0.27 m.u. (6). Identification of transcripts between 0.185 and 0.225 m.u. is described in the text.

Nuclease mapping of the transcripts. The termini of the 2.7-, 1.5-, and 3.65-3.8-kb transcripts were precisely located by S1 nuclease and exonuclease VII mapping with hybrids between infected-cell polyribosomal poly(A) RNA (viral mRNA) and strand-separated 5'- or 3'-end-labeled DNA restriction fragments as described in detail previously (8, 14). Evidence that the 2.7-kb mRNA transcript contains an

intro of ca. 4 kb is based on analysis of hybrids between HSV-1 viral mRNA and HindIII-BamHI fragment J-A DNA (0.182 to 0.223 m.u.) 5' end labeled at the BamHI site (Fig. 3). In this case, S1 nuclease digestion yielded a fragment 500 bases long, whereas exonuclease VII digestion resulted in a fragment ca. 6.2 kb long (Fig. 3A, tracks X and S). A major 5' end of the mRNA was found to lie 530 bases to the left (5')

FIG. 2. In situ RNA hybrids localizing mRNAs between 0.17 and 0.23 m.u. of the HSV-1 KOS genome. Samples (10 μg) of late polysomal poly(A) RNA isolated from infected cells were fractionated on methylmercury-containing agarose gels and immobilized by drying in vacuo. RNA was detected by hybridization with specific nick-translated probes, as indicated by brackets below the restriction endonuclease map of this region. Sizes shown were determined by the position of HeLa cell rRNA (not shown). However, the length of the poly(A) tail (ca. 200 bases [23]) was subtracted to remain consistent with the sizes shown in the transcription map (Fig. 1). No effort was made to standardize the specific radioactivity of various probes. Therefore, relative mRNA abundance should not be compared between tracks.
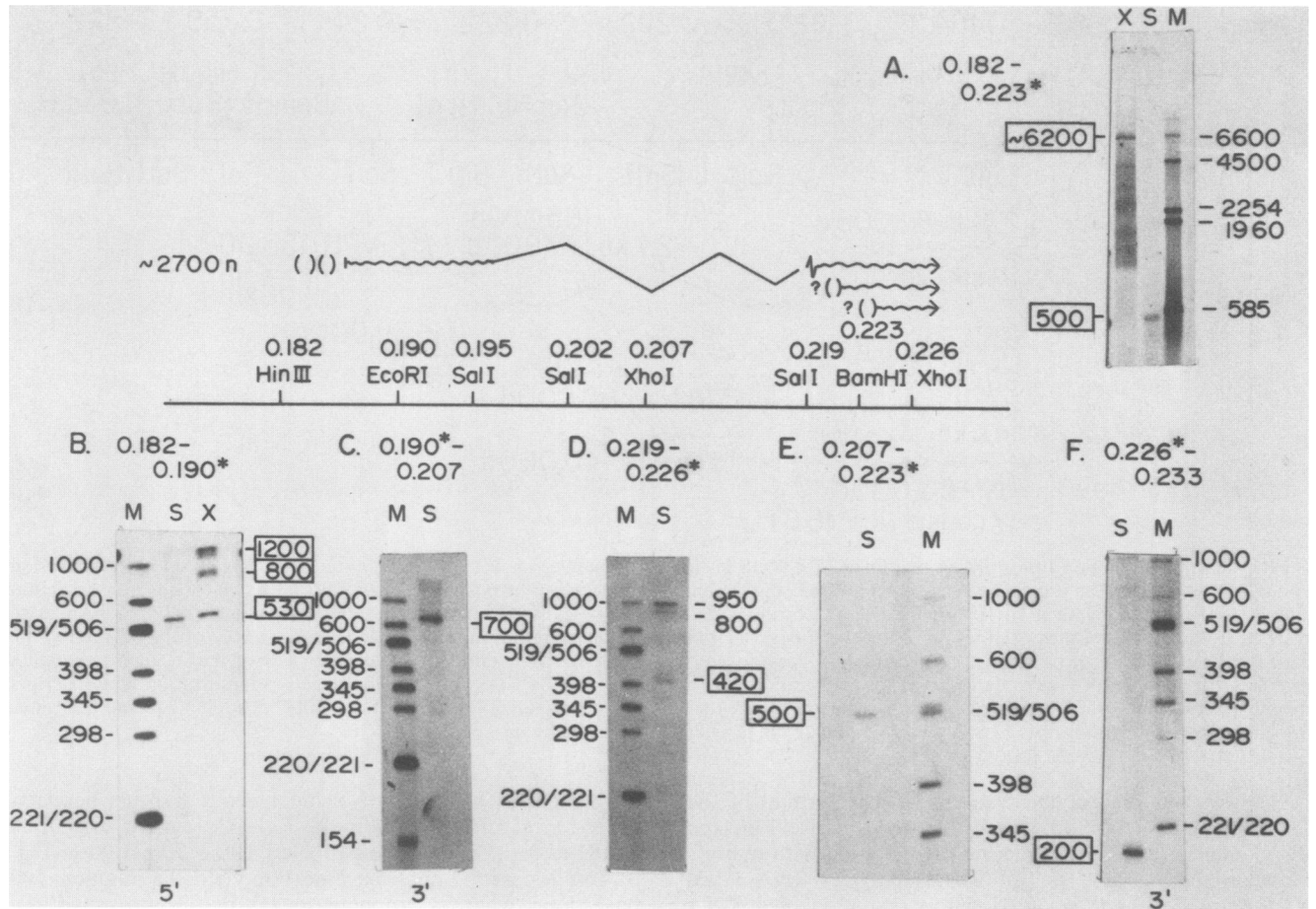
FIG. 3. Localization of the 2.7-kb spliced mRNA by S1 nuclease and exonuclease VII mapping. The diagram summarizes mapping data presented in the text by positioning the spliced mRNA with respect to the restriction endonuclease sites. The 4-kb intron is represented by a wavy line, with possible minor splice acceptor sites indicated by a question mark (?). Single-stranded DNA 5' or 3' end labeled at specific restriction sites was hybridized with HSV-1 poly(A) mRNA, and the DNA protected from nuclease digestion was size-fractionated on an alkaline 1.5% agarose gel (A) or on denaturing 5% acrylamide gels (B through F). Tracks: S, S1 nuclease-digested material; X, exonuclease VII-digested material; M, size marker of HinfI-EcoRI-digested pBR322 DNA or HindIII-digested lambda DNA. Sizes are indicated in daltons. n, Nucleotides. Map positions of the restriction fragments used in the nuclease protection experiments are shown. Asterisks denote the labeled restriction site. (A) HindIII-BamHI fragment J-A DNA; (B) HindIII-EcoRI fragment J-D DNA; (C) EcoRI-XhoI fragment G-U DNA; (D) SalI-XhoI fragment U-V DNA; (E) XhoI-BamHI fragment V-A DNA; (F) XhoI-BglII fragment L-P DNA.

of the EcoRI site at 0.190 m.u. by using HindIII-EcoRI fragment J-D DNA (0.182 to 0.190 m.u.) 5' end labeled at the EcoRI site (Fig. 3B). Exonuclease VII digestion yielded two other bands, 1,200 and 800 bases in size, which are probably artifacts since they were also observed in exonuclease VII digestion experiments with uninfected HeLa cell RNA (data not shown). The 5' contiguous end of the 2.7-kb mRNA (spliced donor portion) extended only about 700 bases to the 3' side of the EcoRI site at 0.190 m.u. (Fig. 3C), where mRNA hybridized to EcoRI-XhoI fragment G-U DNA (0.190 to 0.207 m.u.) 3' end labeled at the EcoRI site.

The right-hand portion of the 2.7-kb mRNA was located in a similar manner. S1 nuclease digestion of hybrids between viral mRNA and SalI-XhoI fragment U-V DNA (0.219 to 0.226 m.u.) 5' end labeled at the XhoI site yielded fully protected DNA (950 bases long) and minor amounts of DNA fragments 800 and 420 bases long (Fig. 3D). This result suggested that a major splice acceptor is located near the SalI site at 0.219 m.u. and that other minor sites may be utilized. However, these minor bands were not seen when DNA 5' end labeled at the BamHI site at 0.223 m.u. was

used in such experiments. The major acceptor was located 500 bases to the left (5') of the BamHI site at 0.223 m.u. by S1 nuclease digestion of hybrids between viral mRNA and XhoI-BamHI fragment V-A DNA (0.207 to 0.223 m.u.) 5' end labeled at the BamHI site (Fig. 3E). The 3' end of the 2.7-kb mRNA was located 200 bases to the right (3') of the XhoI site at 0.226 m.u. by carrying out S1 nuclease digestion of hybrids between viral mRNA and XhoI-BglII fragment L-P DNA (0.226 to 0.233 m.u.) 3' end labeled at the XhoI site (Fig. 3F).

The nested 1.5-kb and 3.65-3.8-kb mRNAs were transcribed from the opposite DNA strand. These RNAs were precisely mapped in a manner similar to that shown in Fig. 4. The 3' end of the clustered transcripts was mapped very near the SalI site at 0.195 m.u. This was accomplished by S1 nuclease analysis of hybrids between viral mRNA and SalI fragment I' DNA (0.195 to 0.202 m.u.) 3' end labeled at 0.202 m.u. (Fig. 4A, track S), where nearly full-length DNA was protected from digestion. Hybridization between viral mRNA and EcoRI-SalI fragment G-D DNA (0.190 to 0.195 m.u.) 3' end labeled at the SalI site gave no S1 nuclease-re-
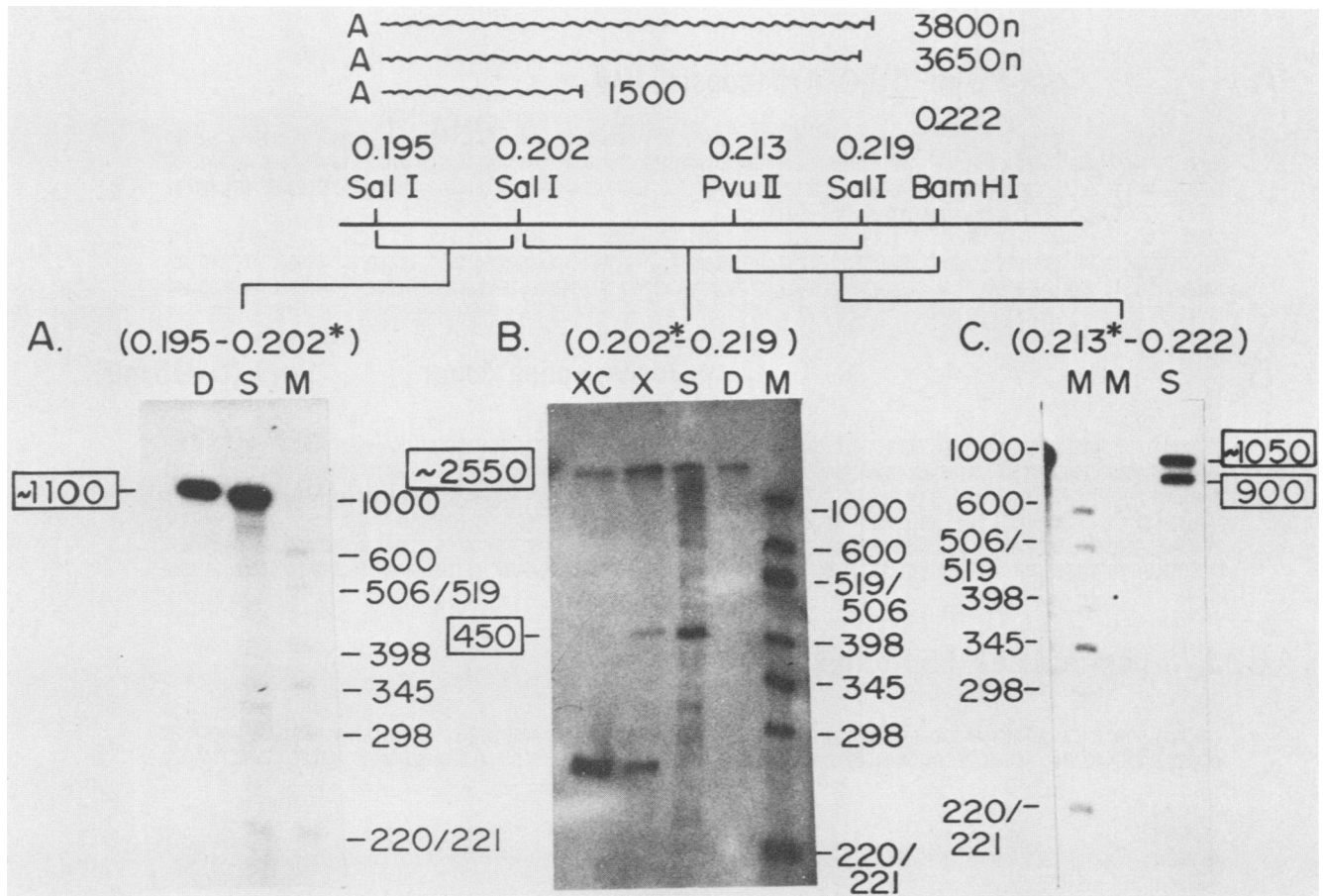
FIG. 4. Mapping of mRNAs located between 0.195 and 0.219 m.u. by S1 nuclease and exonuclease VII analysis. The diagram is a transcription map summarizing the data presented in the text. Details of the nuclease digestions are described in the legend to Fig. 3 and in the text. All tracks are denaturing 5% acrylamide gels. Brackets span the restriction fragments utilized in the nuclease protection experiment. See the legend to Fig. 3 for details. Tracks: D, undigested DNA; XC, exonuclease VII control in which uninfected cell mRNA was used for hybridization. (A) SalI fragment I' DNA; (B) SalI fragment X DNA; (C) PvuII-BamHI fragment P-A DNA.

sistant material (data not shown). Therefore, the 3' end of the mRNA group was mapped to the right (5') of the SalI site at 0.195 m.u.

The 5' end of the 1.5-kb mRNA was mapped 450 bases to the right (5') of the SalI site at 0.202 m.u. S1 nuclease and exonuclease VII digestion of hybrids between viral mRNA and SalI fragment X DNA (0.202 to 0.219 m.u.) 5' end labeled at 0.202 m.u. both yielded fragments 450 bases long (Fig. 4B, tracks S and X). A smaller fragment (ca. 250 bases) seen in the exonuclease VII-digested material was probably an artifact of digestion since it was also seen in a control track of exonuclease VII-digested hybrids between this DNA and uninfected cell mRNA (Fig. 4B, track XC).

The 5' ends of the 3.8- and 3.65-kb mRNAs were shown to map 150 bases apart by S1 nuclease digestion of hybrids between viral mRNA and PvuII-BamHI fragment P-A (0.213 to 0.222 m.u.) 5' end labeled at the PvuII site. It was found (Fig. 4C) that two S1 nuclease-resistant fragments, 900 and 1,050 bases long, were produced.

Correlation between transcript mapping and nucleotide sequence data. The nucleotide sequence of the DNA encoding the left-hand and right-hand portions of the spliced 2.7-kb mRNA, as well as the DNA containing the 5' end of the 1.5-kb mRNA, was determined. These data are available

upon request. Specific features related to the transcripts characterized herein are shown in Fig. 5.

The precise 5' end of the 2.7-kb mRNA was localized by using denaturing acrylamide gels to fractionate S1 nuclease digests of hybrids between viral mRNA and an appropriate 5'-end-labeled 380-base AvaI restriction fragment found in HindIII-EcoRI fragment J-D DNA (0.182 to 0.190 m.u.). This fragment was electrophoresed next to a sequence ladder of DNA labeled at the same site (Fig. 6A). The 5' end of the 3.9-kb mRNA encoding the putative 50,000-Da capsid protein characterized previously (8) is also shown. These data indicate that cap site for these two transcripts, encoded from opposite DNA strands, lie within 100 bases of each other. Note that the sequence ladder is of the DNA strand complementary to the mRNA sequence. The data are summarized in Fig. 5A. There are reasonable TATA and CAT box homologies within 100 bases of the cap sites for both mRNAs, as is normally found for HSV-1 transcripts (24; Wagner, in press).

The putative splice donor was located ca. 700 bases to the right (3') of the EcoRI site at 0.190 m.u. (Fig. 3C). It was located (data not shown) ca. 40 bases to the left (5') of the SalI site at 0.195 m.u. by S1 nuclease analysis of hybrids between viral mRNA and a 3'-end-labeled 130-base AvaI

A   Cap-3.9kb-50,000d (capsid) MW

```
     10        20        30        40        50        60        70        80        90       100       110       120
CGCGAATTTTAACACAGGTCGCGCGCTGTCGCATCATCTCTAAGCGCGCGCGGGACTTTAGCCGCGCCTCCAATTCCAAGTGGGCCGCCTTTGCAGCCATAAAGGCGCCAACAAACCGAG
GCGCTTAAAATTGTGTCCAGCGCGCGACAGCGTAGTAGAGATTCGCGCGCGCCCTGAAATCGGCGCGGAGGTTAAGGTTCACCCGGCGGAAACGTCGGTATTTCCGCGGTTGTTTGGCTC
```

Cap 2.7 kb spliced mRNA

```
    130       140       150       160       170       180       190       200       210       220       230       240
GATCTTGGGTGCTGACGCCCTCCCGGTGCAGCTGCAGGGTCTTGTCCTTGTAAATCTCGGCTCGGAGGTGCGTCTCGGCCAGGCGTCGGCGCAGGGCCGCGTGGGCGGCATCTCGGTCCA
CTAGAACCCACGACTGCGGGAGGGCCACGTCGACGTCCCAGAACAGGAACATTTAGAGCCGAGCCTCCACGCAGAGCCGGTCCGCAGCCGCGTCCCGGCGCACCCGCCGTAGAGCCAGGT
```

B   putative splice donor   Sal I (0.195 mu)

```
     10        20        30        40        50        60   ↓  70        80        90       100       110       120
CCTTCTCGTTTCCGGACGGGTCGCGCAGTACCATCGTGTTTGCCTCCAGCCACAACACAAACGTAAGTCCTCTTTTCTTTCGCATGGCTCTCCCAAGGGGCCCCGGGTCGACCCGACCCA
GGAAGAGCAAAGGCCTGCCCAGCGCGTCATGGTAGCACAAACGGAGGTCGGTGTTGTGTTTGCATTCAGGAGAAAAGAAAGCGTACCGAGAGGGTTCCCCGGGGCCCAGCTGGGCTGGGT
```

poly(A)sig. (1.5-3.85 kb m RNA cluster)

```
    130       140       150       160       170       180       190       200       210       220       230       240
CACCCACCCACCCACATACACACACAACCAGACGCGGGAGGAAAGTCTGCCCCGTGGGACTGATTTTTATTCGGGATCGCTTGAGGAGGCCCGGGCAACGGCCAGCCGGTGGGGCAACTC
GTGGGTGGGTGGGTGTATGTGTGTGTTGGTCTGCGCCCTCCTTTCAGACGGGGCACCCTGACTAAAAATAAGCCCTAGCGAACTCCTCCGGGCCCGTTGCCGGTCGGCCACCCCGTTGAG
```

C   nominal cap 1.5 kb mRNA

```
     10        20        30        40        50        60        70        80        90       100       110       120
CGTCACCATCCATCGCCACGGCGATTGGACGATTGTTAAGCCGCAGCGTGTCTCCGCTTGTGCTGTAGTAGTCAAAAACGTAATGCCGTCGGAGTCGGCAAAGCGGGCCGGGAGGTCGTC
GCAGTGGTAGGTAGCGGTGCCGCTAACCTGCTAACAATTCGGCGTCGCACAGAGGCGAACACGACATCATCAGTTTTTGCATTACGGCAGCCTCAGCCGTTTCGCCCGGCCCTCCAGCAG
```

```
    130
TTCTCGAG
AAGAGCTC
```

D   possible splice acceptor

```
     10        20  ↓   30        40        50        60        70        30        90       100       110       120
GGGGGTGGGTGTGGGTGTGTAGGCGATGCGGGTCCCCCGAGTCCGCGGGAAGGGCGTGGGTTTGGCGCGCGTATGCGTATTCGCCAACGGAGGCGTGCGTGCTTATGCGCGGCGCGTTTC
CCCCCACCCACACCCACACATCCGCTACGCCCAGGGGGCTCAGGCGCCCTTCCCGCACCCAAACCGCGCGCATACGCATAAGCGGTTGCCTCCGCACGCACGAATACGCGCCGCGCAAAG
```

splice acceptor   Sal I (0.219 mu)

```
    130 ↓    140       150       160       170       180       190       200       210       220       230       240
TTCTGTCTCCAAGGAATCCGAGGCCAGGACTTTAACCTGCTCTTTGTCGACGAGGCCAACTTTATTCGCCCGGATGCGGTCCAGACGATTATGGGCTTTCTCAACCAGGCCAACTGCAAG
AAGACAGAGGTTCCTTAGGCTCCGGTCCTGAAATTGGACGAGAAACAGCTGCTCCGGTTGAAATAAGCGGGCCTACGCCAGGTCTGCTAATACCCGAAAGAGTTGGTCCGGTTGACGTTC
```

Nominal cap 3.6 kb mRNA

```
    250       260       270       280       290       300       310       320       330       340       350       360
ATTATCTTCGTGTCGTCCACCAACACCGGGAAGGCCAGTACGAGCTTTTTGTACAACCTCCGCGGGCCGCCGACGAGCTTCTCAAC... ..TGACCTATATATGCGATGATCACATGCCG
TAATAGAAGCACAGCAGGTGGTTGTGGCCCTTCCGGTCATGCTCGAAAAACATGTTGGAGGCGCCCCGGCGGCTGCTCGAAGAGTTGCACCACTGGATATATACGCTACTAGTGTACGGC
```

```
    370       380       390       400       410       420
AGGGTGGTGACGCACACAAACGCCACGGCCTGTTCTTGTTATATCCTCAACAAGCCCGTT
TCCCACCACTGCGTGTGTTTGCGGTGCCGGACAAGAACAATATAGGAGTTGTTCGGGCAA
```

Cap-3.85 kb mRNA

E   XhoI (0.226 mu)

```
     10        20        30        40        50        60        70        80        90       100       110       120
TCATGGCCTCCCAGGAGGTCGTTTCCGCGACGGTGCGCCTGCAGACCGACCCGGTCGAGTATCTGCTCGAGCAGCTGAATAACCTCACCGAAACCGTCTCCCCCAACACTGACGTCCGTA
AGTACCGGAGGGTCCTCCAGCAAAGGCGCTGCCACGCGGACGTCTGGCTGGGCCAGCTCATAGACGAGCTCGTCGACTTATTGGAGTGGCTTTGGCAGAGGGGGTTGTGACTGCAGGCAT
```

translation terminator

```
    130       140       150       160       170       180       190       200       210       220       230       240
CGTATTCCGGAAAACGGAACGGCGCCTCGGATGACCTTATGGTCGCCGTCATTATGACCATCTACCTCGCGACCCAGGCCGACCTCCGCACACATTCGCTCCCATCACACGCGTTTCGTG
GCATAAGGCCTTTTGCCTTGCCGCGGAGCCTACTGGAATACCAGCGGCAGTAATACTGGTAGATGGAGCGCTGGGTCCGGCTGGAGGCGTGTGTAAGCGAGGGTAGTGTGCGCAAAGCAC
```

poly(A)sig. (2.7 kb spliced mRNA)

```
    250       260       270       280       290       300       310       320
AGCGCCCAATAAACACACCTAGGTATGCTACGCACGACCACGGTGTCGCCTGCCAAGGGGGGGGGGAAGGGGGTGCCGGC
TCGCGGGTTATTTGTGTGGATCCATACGATGCGTGCTGGTGCCACAGCGGACGGTTCCCCCCCCCCTTCCCCCACGGCCG
```
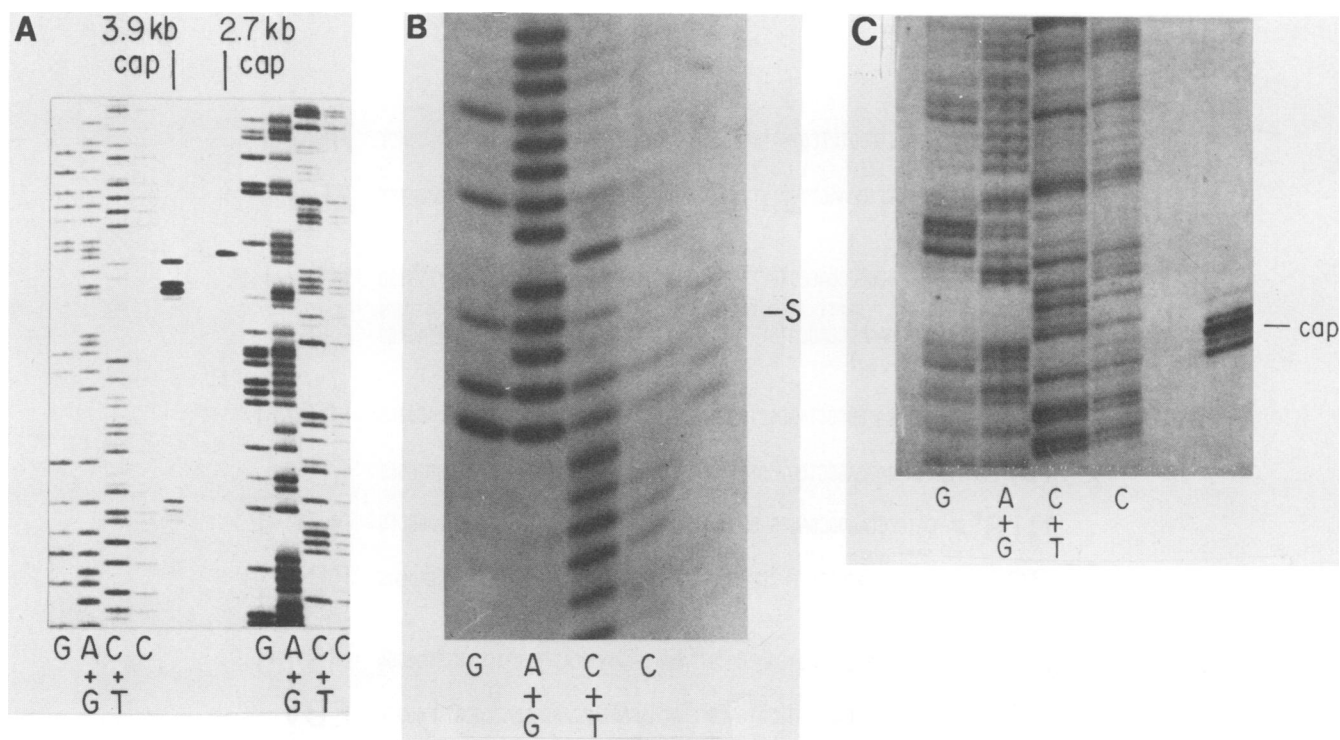
FIG. 6. Precise localization of the 5' ends of the 3.9-, 2.7-, and 3.8-kb mRNAs and of a splice acceptor site of the 2.7-kb mRNA. HSV-1 mRNA was hybridized with 5'-end-labeled single-stranded DNA, followed by digestion with S1 nuclease. The nuclease-resistant material was fractionated against a sequence ladder of the DNA fragment 5' end labeled at the same restriction site. (A) 5'-end-labeled 380-base-pair *Ava*I restriction fragment DNA located in *Hind*III-*Eco*RI fragment J-D (0.182 to 0.190 m.u.) localizing the mRNA cap sites of the 2.7- and 3.9-kb mRNAs on the DNA sequence. (B) *Xho*I-*Sal*I fragment V-X DNA (0.210 to 0.219 m.u.) 5' end labeled at the *Sal*I site to position the 2.7-kb splice acceptor(s) to the DNA sequence. (C) *Sal*I-*Bam*HI fragment U-A DNA (0.219 to 0.222 m.u.) 5' end labeled at the *Sal*I site to precisely localize the 5' end of the 3.85-kb mRNA (cap).

restriction fragment whose right-hand (3') end corresponded to the *Sma*I site at nucleotides 102 through 107 (Fig. 5B). This potential splice donor sequence (Fig. 5B) shares nominal sequence similarity with other eucaryotic splice donor sequences (21).

A potential polyadenylation signal, AATAAAAA, was observed 80 bases to the right of the *Sal*I site at 0.195 m.u. (Fig. 5B). The location of this site agrees well with the 3' end of the 1.5- and 3.65-3.8-kb mRNA nested group (Fig. 4A). This location was confirmed by S1 nuclease analysis of hybrids between viral mRNA and a 110-base *Sal*I-*Sma*I restriction fragment extending to the right of the *Sal*I site at 0.195 m.u. and 3' end labeled at the *Sma*I site after cleavage with *Xma*I to generate a 5' overhang (data not shown).

The cap site for the 1.5-kb member of the cluster of mRNAs encoded by the strand opposite the 2.7-kb spliced mRNA intron was located 450 bases to the right (5') of the *Sal*I site at 0.203 m.u. (Fig. 4B). The nucleotide sequence of DNA in this region is shown in Fig. 5C. Tentative location of the cap site is ca. 25 to 30 bases to the left (3') of the nominal

TATA box homology (bases 31 through 39, Fig. 5C). We suggest that this region is a normal late HSV-1 promoter region on the basis of comparative sequence analysis of HSV-1 promoters (reviewed in Wagner, in press).

One splice acceptor for the 2.7-kb mRNA was precisely located 34 bases to the left (5') of the *Sal*I site at 0.219 m.u. by S1 nuclease analysis of hybrids between viral mRNA and *Xho*I-*Sal*I fragment V-X DNA (0.210 to 0.219 m.u.) 5' end labeled at the *Sal*I site. An example of fractionated S1 nuclease-resistant material compared with a sequence ladder proceeding from the *Sal*I site at 0.219 m.u. is shown in Fig. 6B. The contiguous 5' end of the RNA lies within bases 130 through 136 (Fig. 5D), whose complement is shown in the ladder (Fig. 6B). The position of this acceptor is 960 bases to the left (5') of the *Xho*I site at 0.226 m.u., a result in excellent agreement with the data shown in Fig. 3D.

Other potential splice acceptors were suggested from the data (Fig. 3D). As discussed above, these correspond to regions ca. 150 and 530 bases to the right (3') of the major acceptor indicated here. There are potential sites at nucleo-

FIG. 5. Selected DNA sequences in the region between 0.19 and 0.25 m.u. The positions of the mRNA cap sites, polyadenylation sites, and putative splice donor and acceptor sites were determined by S1 nuclease analysis (6, 8, 10, 14) as discussed in the text. (A) DNA sequence around the mRNA cap sites of both the 3.9-kb 50,000-Da (d) capsid protein (ICP 18.8) and the 2.7-kb spliced mRNA. (B) Nucleotide sequence in the vicinity of the 2.7-kb mRNA splice donor site and the polyadenylation site [poly(A) sig.] of the nested 1.5-kb and 3.65-3.8-kb mRNAs. (C) DNA sequence data 5' of the 1.5-kb mRNA cap site. (D) DNA sequences in the region of the splice acceptor site and a possible alternate acceptor site for the 2.7-kb spliced mRNA. Sequences upstream from the 5' ends of both the 3.65- and 3.8-kb mRNAs are included. (E) Nucleotide sequence at the 3' end of the 2.7-kb spliced mRNA; the potential translation terminator (bracket, top row) and the poly(A) site are indicated.

A.

tides 172 and 203 (Fig. 5D). Both sites would change the translational reading frame from that generated from the major acceptor. Another potential acceptor is located at nucleotide 22 (Fig. 5D) to the left (5') of the major acceptor. This site does yield the same translational reading frame as the major acceptor. Its significance will become apparent in the next section, where homology between this HSV-1 DNA sequence and that of EBV DNA is discussed. Finally, several potential splice acceptor sequences are seen in the region around the BamHI site at 0.223 m.u. (not shown). These could function to generate the putative splice acceptor which was observed 420 bases 5' of the XhoI site at 0.226 m.u. (Fig. 3D).

The cap site of the 3.8-kb mRNA was determined on gels by using sequence ladders as size standards for the S1 nuclease-resistant material generated by the S1 nuclease digestion of hybrids between viral mRNA and SalI-BamHI fragment U-A DNA (0.219 to 0.223 m.u.) 5' end labeled at the SalI site (Fig. 6C). The sequence around this cap site is shown in Fig. 5D. A TATA box homology 25 bases to the right (5') of this cap site is clearly visible (nucleotides 337 through 343, Fig. 5D). The sequence 155 bases to the left (3')

of this site (nucleotide 165) is indicated as the nominal cap site for the 3.65-kb mRNA, since a good TATA box homology is seen ca. 30 bases to the right (5') of this site (nucleotides 181 through 187, Fig. 5D).

The sequence in the vicinity of the XhoI site at 0.226 m.u. is shown in Fig. 5E. Of particular note are nucleotides 75 through 78, which are a translation terminator signal (TGA) that could close a long open reading frame for the 2.75-kb spliced mRNA (see below), and a poly(A) signal at nucleotides 248 through 258, 185 bases to the right (3') of the XhoI site. The position of this signal sequence is in excellent agreement with the position determined for the 3' end of the 2.7-kb spliced mRNA (200 bases to the right [3'] of the XhoI site at 0.226 m.u., Fig. 3F).

Homology between the 2.7-kb spliced HSV-1 mRNA sequence and sequences of EBV DNA. Because spliced mRNAs are so unusual in HSV-1 transcripts, it seemed possible that the 2.7-kb mRNA might be conserved among other herpesviruses. Therefore, the sequence of HSV-1 DNA encoding the spliced 2.7-kb mRNA was compared with the complete EBV DNA sequence (3). The computer search algorithm used was developed by D. J. Lipman and W. R. Pearson

B.

GAACCCGGCCGTGGCCGAGATCGTACTGGAC **HSV-1**

129,001 CAATCGGCCTTGGTCTGACATTCCACCCCGGCCAGGTCCAGGAGGGTGCAAATATTCTCC **EBV**

(base No.23,fig.6D) HSV-1294

CTCGTTGGCCGGGTGCGCGTTCATGGTTCGGGGGTGGGTGTGGGTGTGTAGGCGATGCGG **HSV-1**

:: ::::

129,061 AGGCGCTGCACCTCAGAGACCTCCTGCTCAAAGAGACCTCCCACCGCCACGTAGACGCGG **EBV**

1303' GTCCCCCGAGTCCGCGGGAAGGGC-GTGGGTTTGGCGCGCGTATGCGTATTCGCCAACGG **HSV-1**

: :: ::: :::: :::::::: : ::::: : :: : :: :::: ::::

129,121 G-CCACCG--TCCG-GGGAAGGTCAGTGGGGTCCCAGCTC---AGC-AATTCTCCAA--- **EBV**

(HSV-1 splice acceptor-base 134-Fig 6D)

↓|

1362' AGGCGTGCGTGCTTAT-GCGCGGCGCGT-TTCTTCTGTC-TCCAAGGAATCCGAGGCCAG **HSV-1**

: : : : :: : :: ::: : :: : ::::: : :: : ::::: :: :::

129,170 ATTCTCTCTCCCCAATAGTGCCTCGCTTCTTATCCTGTCTTTCAGAGCATCCGGGGGCAG **EBV**

HSV-1 SalI (0.219 mu)　　|

1419' GACTTTAACCTGCTCTTTGTCGACGAGGCCAACTTTATTCGCCCGGATGCGGTCCAGACG **HSV-1**

::: :::: : ::::: :::::::::: :::::::: ::: :: : : : ::

129,230 ACATTTCACCTCTTGTTTGTGGACGAGGCTAACTTTATCAAGAAGGAGGCCCTGCCGGCG **EBV**

1479' ATTATGGGCTTTCTCAACCAGGCCAACAGCAAGATTATCTTCGTGTCGTCCACCAACACC **HSV-1**

:: :::::::: : ::: : :::::::::::::: : ::::: ::: :

129,290 ATCCTGGGCTTTATGCTTCAGAAGGATGCCAAGATTATCTTCATCTCGTCTGTGAACTCG **EBV**

1539' GGGAAGGCCAGTACGAGCTTTTTGTACAACCTCCGCGGGGCCGCCGACGAGCTTCTCAAC **HSV-1**

: : :: :::::: : :: :: :: : :: ::: :::

129,350 GCTGACCAGGCCACCAGCTTTCTTTATAAGCTGAAGGATGCTCAGGAGCGGCTGCTGAAC **EBV**

1599' GTGGTGACCTATATATGCGATGATCACATGCCGAGGGTGGTGACGCACACAAACGCCACG **HSV-1**

::::: : ::: : :: : :: :: :: : : : :: : :

129,410 GTGGTAAGTTATGTGTGTCAGGAGCATCGGCAAGATTTTGACATGCAGGACAGCATGGTC **EBV**

|

1659' GCCTGTTCTTGTTATATCCT-CACAAGCCCGTTTTCATCACGATGGACGGGGCGGTTCGC **HSV-1**

: :: : :: : : ::: :::: :: : ::::::: :::::: : : ::

129,470 TCATGCCCCTGCTTTCGCCTGCACATCCCGTCCTACATCACCATGGACAGCAACATCCGA **EBV**
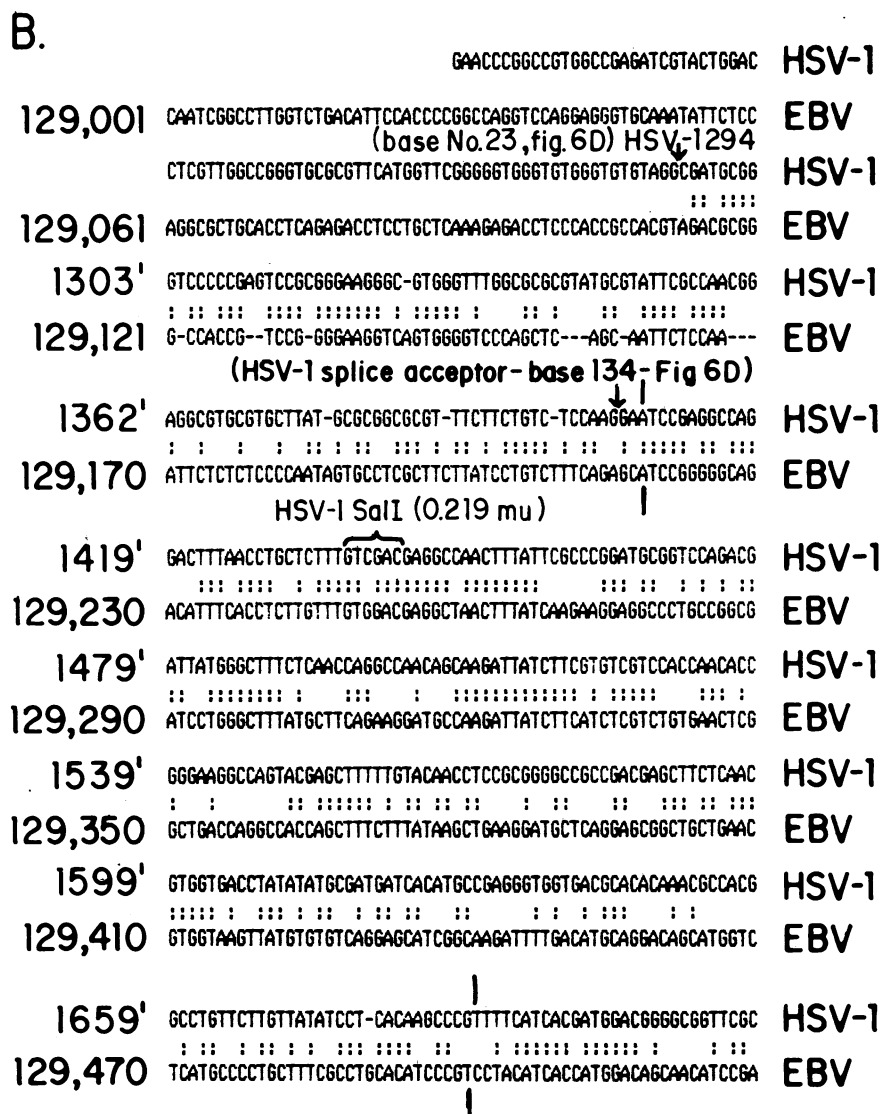
|

FIG. 7. Nucleotide sequence homology between HSV-1 KOS DNA encoding the spliced 2.7-kb transcript and EBV B95-8 DNA. The splice donor (A) and splice acceptor (B) regions of HSV-1 are shown, as are the SalI sites at 0.195 and 0.219 m.u. Regions of homology are designated by :; - indicates regions where gaps were introduced by the computer homology program to maximize overall sequence fit. Regions of strongest computer-generated homology are bracketed by vertical lines. HSV-1 DNA is numbered from a site 120 bases 5' of the cap site of the spliced 2.7-kb transcript. EBV DNA is numbered as described previously (3).

(Science, in press). Significant nucleic acid homology was found to exist between the HSV-1 DNA sequence and two regions of the DNA of EBV B95-8. These regions are separated by ca. 3,000 bases in the EBV genome (Fig. 7).

The 783 bases of sequence to the left (5') of the HSV-1 splice donor site near the SalI site at 0.195 m.u. was found to be homologous to EBV DNA between bases 125,118 and 125,901. An optimized match between 375 of 780 bases was found, compared with a "random" expected value of less than 190 matches for nonhomologous EBV DNA. The sequence of EBV DNA is numbered as described previously (3). The region of greatest homology (Fig. 7A) lay within 300 bases of the splice donor sequence of HSV-1 (182 matches). Homology was lost 3' of the splice donor site in the HSV-1 sequence (nucleotide 1287, Fig. 7A). Potential splice donor sites in the EBV DNA sequence are seen at base 125,875 as well as at several other places.

Considerable homology was seen between the HSV-1

DNA at the splice acceptor region and EBV DNA beginning at base 129,055. This region showed a match for 476 of 900 bases. The homology began at a potential splice acceptor site in HSV-1 DNA lying 145 bases to the left (5') of the SalI site at 0.219 m.u. This site (nucleotide 23, Fig. 5D) is numbered HSV-1 base 1294 in Fig. 7B. Homology extended ca. 820 bases to the right (3') of this site, where it was abruptly lost (not shown). The position where homology was lost corresponds to base 129,956 of the EBV DNA sequence. The 450-base region of highest homology near the identified splice acceptors is shown in Fig. 7B (251 matches). Potential splice acceptor sites are evident in the EBV DNA sequence very near those identified in HSV-1 DNA.

The nucleic acid sequence homology seen between the HSV-1 DNA encoding the spliced 2.7-kb mRNA and several regions of EBV DNA suggested that strong protein homology should also exist. The HSV-1 DNA sequence encoding the spliced mRNA contains a translation start signal (ACAA-

A.

201　　211　　221　　231　　241　　251　　261

YGTLELFQKMILMHATYFLAAVLLGDHAEQVNTFLRLVFEIPLFSDAAVRHFRQRATVFLVPRRHGKTWF　HSV-1

　:: :::　　:　:::　　　　　　　　:　　　　　　　: : : ::: :::::::::

PARLEAFQKQVVLHSFYFLISIKSLEITDTMFDIFQSAFGLEEMTLEKLHIFKQKASVFLIPRRHGKTWI　EBV

169　　179　　189　　199　　209　　219　　229

271　　281　　291　　301　　311　　321　　331

LVPLIALSLASFRGIKIGYTAHIRKATEPVFEEIDACLRGWFGSARVDHVKGE TISFSFPDGSRSTIVF　HSV-1

: : : :　　　::: ::　　　:: ::　:　: : ::　:　:: :　　　::

WAIISLILSNLSWQIGYVAHQKHVASAVFTEIIDTLTKSFDSKRVEVNKETSTITFRHSGKISSTVMC　EBV

239　　249　　259　　269　　279　　289　　299

340

ASSHNTWV　←splice donor site

: : ::

ATCFNKWV

309

putative alternate splice acceptor

|

identified splice acceptor of fig. 6D

↓

B.

↓315　　325　　335　　345↓　　355　　365　　375

AGPPSPREGRGFGARMRIRQRRRACLCAARFFCLQGIRGQDFNLLFVDEANFIRPDAVQTIMGFLNQANC　HSV-1

::::　　::　:　　: :　　　: :　　　: :::: : :::::::::::　　:　　: :: :

RGPPSG EGQ WGPSSAILQILSPQXCLASYPVFQSIRGQTFHLLFVDEANFIKKEALPAILGFMLQKDA　EBV

42　　52　　62　　72　　82　　92　　102

385　　395　　405　　415　　425　　435　　445

KIIFVSSTNTGKASTSFLYNLRGAADELLNVVTYICDDHMPRVVTHTNATACSCYILNKPVFITMDGAVR　HSV-1

::::: :: :　　:::::: :　:　　::::: : : :　　　　　: : : : :::: :

KIIFISSVNSADQATSFLYKLKDAQERLLNVVSYVCQEHRQDFDMQDSMVSCPCFRLHIPSYITMDSNIR　EBV

110　　120　　130　　140　　150　　160　　170

455　　465　　475　　485　　495　　505　　515

RTADLFLAESCMQEIIGGQAR ETGDDRPVLTKSAGERFLLYRPST TTN SGLMAPDLYVYVDPAFTAN　HSV-1

: :::　　: :　　　:　　　　:　　　:: :　　　: : ::::::::: : :

ATTNLFLDGAFSTELMGDTSSLSQGSLSRTVRDDAINQLELCRVDTLNPRVAGRLASSLYVYVDPAYTNN　EBV

180　　190　　200　　210　　220　　230　　240

522　　532　　542　　552　　562　　572　　582

TRASGTGVAVVGRYR D D YIIFALEHFFLRALTGSAPADIARCVVHSPTQVLALHPGAFRGVRVAVEG　HSV-1

: :::::: : :　　: :　　:　::::::　::: :　:: :::　　:::　　: :::::

TSASGTGIAAVTHDRADPNRVIVLGLEHFFLKDLTGDAALQIATCVVALVSSIVTLHP HLEEVKVAVEG　EBV

250　　260　　270　　280　　290　　300　　310

589　　599

NSSQDSAVAIATHV　HSV-1

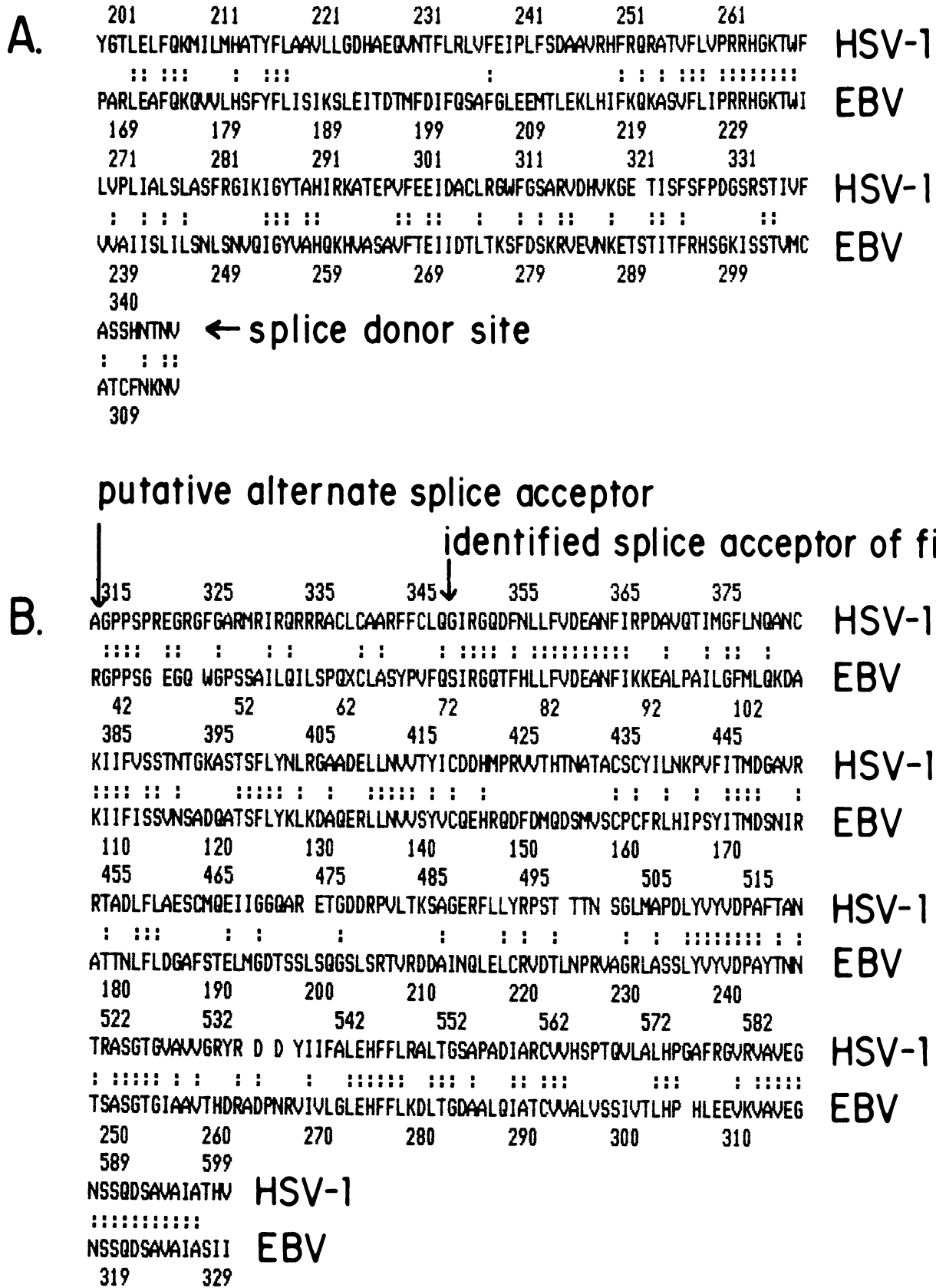::::::::::::

NSSQDSAVAIASII　EBV

319　　329

FIG. 8. Comparison of predicted amino acid homology between the protein encoded by the spliced HSV-1 transcript and the putative EBV polypeptide. As described in the text, the HSV protein is initiated at an ATG triplet 331 bases 3' of the mRNA cap site and extends 646 or 683 amino acids, depending on the splice acceptor used. (A) Amino acid sequence encoded around the splice donor site of the mRNA. The EBV polypeptide here is initiated at base 124,939. (B) Amino acid sequence near the splice acceptor sites. The reading frame used for EBV here was arbitrarily begun at base 128,994. There is a terminator (X, number 62 in the amino acid sequence) at base 129, 185. The reading frame then stays open between bases 129,188 and 130,348. Both EBV reading frames were noted previously (3).

TGG) 331 bases to the right (3') of the cap site (Fig. 5A). This translational frame remains open through the splice donor site (base 1287, Fig. 7A) and through both the alternate and primary splice acceptor sites (nucleotides 23 and 134, Fig. 6D, corresponding to HSV-1 bases 1294 and 1405, Fig. 7B). It terminates at the TGA codon near the XhoI site at 0.226 m.u. (base 76, Fig. 5E). This translational reading frame predicts proteins of 70,750 (646 amino acids) or 74,822 (683 amino acids) Da or both, depending on the splice acceptor used. The codon use frequency of the proteins was similar to that seen for other HSV-1 proteins predicted from nucleotide sequence data (reviewed in Wagner, in press). Furthermore, the predicted molecular weight of the protein encoded by the spliced 2.7-kb mRNA was in good agreement with that predicted by the in vitro translation data.

The EBV DNA sequence predicted a protein initiated at base 124,938 (AACATGC) and continuing through the potential splice donors noted in Fig. 7A. A protein translational reading frame is also seen in the EBV sequence corresponding to the major splice acceptor homology shown in Fig. 7B (EBV DNA base 129,114). There is, however, a translation terminator in the EBV DNA sequence to the left (5') of this splice acceptor site and to the right of the site homologous to the putative alternate HSV-1 acceptor site 111 bases to the left (5') of the major one (EBV base 129,186, Fig. 7B). The translation reading frame for the EBV DNA sequence terminates at base 130,351. A residue molecular weight of ca. 82,000 is predicted for the protein encoded by the combined translation frames if the chain terminator noted above is ignored. If the major splice acceptor is used, then the predicted protein would be ca. 79,000 Da.

The search algorithm of Lipman and Pearson (in press) was used to examine homology between the predicted HSV-1 protein encoded by the spliced 2.7-kb mRNA and the putative protein encoded by EBV (Fig. 8). Regions of striking homology were seen in the predicted amino acid sequence of the proteins both on the N-terminal side of the predicted splice donor site (52 of 146 amino acids matched; Fig. 8A) and on the C-terminal side of the two splice acceptor sites (135 of 287 amino acids matched; Fig. 8B). In this latter case, it is seen that homology extends to the left of the terminator signal in the EBV sequence (X, corresponding to amino acid 62 of EBV, Fig. 8B). The predicted amino acid homology for the proteins encoded between the potential alternate splice acceptor sites of HSV-1 suggests that there may be alternate splicing patterns in the EBV transcript encoding the protein homologous to that of HSV-1. At any rate, the very striking amino acid sequence homology is compelling evidence for the conservation of this protein between these two rather distantly related herpesviruses.

DISCUSSION

Detailed location of the transcripts described in this report revealed their general similarity to HSV-1 transcripts mapped in other regions. Thus, sequences recognizable as promoter-regulatory regions exist just 5' to the transcript cap sites. Polyadenylation signals are the nominal eucaryotic ones, and nested sets of transcripts are present. The current data do show one striking variation on this theme: the existence of a transcript containing a 4-kb intron. Full characterization of this transcript will require cDNA sequencing through the potential splice donor and acceptor sites. Availability of the DNA sequence within this region will allow the synthesis of specific DNA primers to carry out the necessary synthesis and cloning of the cDNA required.

Although full characterization of splicing patterns for the

HSV transcript remains to be done, current data clearly reveal that this spliced transcript, which is so unusual for HSV-1, encodes a protein conserved between HSV-1 and EBV. This suggests that the protein may be important in the biology of herpesviruses as a group. The sequence homology is of the same order as that seen between the HSV-1 and EBV genes for ribonucleotide reductase (15; Wagner and Draper, unpublished data). In the case of the spliced transcript, however, there was somewhat less homology between the regions. There are, at this date, two other HSV proteins which have been reported to have homology with EBV ones: DNA polymerase and thymidine kinase (3). For thymidine kinase, the homology is minimal and is not seen in the DNA sequence (15).

Although it cannot yet be said what the functions are of the HSV genes encoding the spliced mRNA and its intron, it is clear that the combination of precise transcript mapping and nucleotide sequence analysis has revealed a potentially interesting herpesvirus genetic marker. The predicted amino acid sequence homology with a putative EBV protein is striking. Furthermore, sequence data will allow the synthesis of specific synthetic polypeptides to be used as immunogens for generating antibodies which can be used to examine the properties and potential function of this protein.

Finally, it is interesting that the arrangement of the transcripts, vis-à-vis the nested set of transcripts encoded from the opposite DNA strand of the spliced 2.7-kb mRNA intron, may also be generally conserved in the EBV genome. Analysis of potential translational reading frames of the EBV DNA indicates several long open frames in the region between bases 126,000 and 129,000 (3). These translation frames are on the opposite coding strand from the EBV frames homologous to those of HSV-1 and could be expressed as a nested set of transcripts analogous to those seen in HSV-1. Further sequence analysis may reveal homologies within such translational reading frames.

LITERATURE CITED

1. Anderson, K., R. Frink, G. Devi, B. Gaylord, R. Costa, and E. Wagner. 1981. Detailed characterization of the mRNA mapping in the HindIII fragment K region of the herpes simplex virus type genome. J. Virol. 37:1011–1027.
2. Anderson, K., J. Stringer, L. Holland, and E. Wagner. 1979. Isolation and localization of herpes simplex virus type 1 mRNA. J. Virol. 30:805–820.
3. Baer, R., A. Bankier, M. Biggin, P. Deininger, P. Farrell, T. Gibson, G. Hatfull, G. Hudson, S. Satchwell, C. Seguin, P. Tuffnell, and B. Barrell. 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) 310:207–211.
4. Bailey, J. M., and N. Davidson. 1976. Methylmercury as a reversible denaturing agent for agarose gel electrophoresis. Anal. Biochem. 70:75–85.
5. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-

digested hybrids. Cell **12**:721–732.

6. Costa, R., G. Cohen, R. Eisenberg, D. Long, and E. Wagner. 1984. Direct demonstration that the abundant 6-kilobase herpes simplex virus type 1 mRNA mapping between 0.23 and 0.27 map units encodes the major capsid protein VP5. J. Virol. **49**:287–292.

7. Costa, R., B. Devi, K. Anderson, B. Gaylord, and E. Wagner. 1981. Characterization of a major late herpes simplex virus type 1 mRNA. J. Virol. **38**:483–496.

8. Costa, R., K. Draper, L. Banks, K. Powell, G. Cohen, R. Eisenberg, and E. Wagner. 1983. High-resolution characterization of herpes simplex virus type 1 transcripts encoding alkaline exonuclease and a 50,000-dalton protein tentatively identified as a capsid protein. J. Virol. **48**:591–603.

9. Denhardt, D. T. 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641–646.

10. Draper, K., R. Costa, G. Lee, P. Spear, and E. Wagner. 1984. Molecular basis of the glycoprotein-C-negative phenotype of herpes simplex virus type 1 macroplaque strain. J. Virol. **51**:578– 585.

11. Draper, K., R. Frink, and E. Wagner. 1982. Detailed characterization of an apparently unspliced beta herpes simplex virus type 1 gene mapping in the interior of another. J. Virol. **43**:1123–1128.

12. Frink, R., K. Anderson, and E. Wagner. 1981. Herpes simplex virus type 1 HindIII fragment L encodes spliced and complementary mRNA species. J. Virol. **39**:559–572.

13. Frink, R., K. Draper, and E. Wagner. 1981. Uninfected cell polymerase efficiently transcribes early but not late herpes simplex virus type 1 mRNA. Proc. Natl. Acad. Sci. U.S.A. **78**:6139–6143.

14. Frink, R., R. Eisenberg, G. Cohen, and E. Wagner. 1983. Detailed analysis of the portion of the herpes simplex virus type 1 genome encoding glycoprotein C. J. Virol. **45**:634–647.

15. Gibson, T., P. Stockwell, M. Ginsburg, and B. Barrell. 1984. Homology between two EBV early genes and HSV ribonucleotide reductase and 38K genes. Nucleic Acids Res. **12**:5087–5099.

16. Hall, L., K. Draper, R. Frink, R. Costa, and E. Wagner. 1982. Herpes simplex virus mRNA species mapping in EcoRI fragment I. J. Virol. **43**:594–607.

17. Holland, L., K. Anderson, C. Shipman, Jr., and E. Wagner. 1980. Viral DNA synthesis is required for the efficient expression of specific herpes simplex virus type 1 mRNA species. Virology **101**:10–24.

18. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.

19. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning—a laboratory manual. Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.

20. Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–559.

21. Mount, S. 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.

22. Palmiter, R. D. 1974. Mg^{++} precipitation of ribonucleo-protein complexes. Expedient techniques for the isolation of undegraded polysome and messenger ribonucleic acid. Biochemistry **13**:3606–3614.

23. Stringer, J., L. Holland, R. Swanstrom, K. Pivo, and E. Wagner. 1977. Quantitation of herpes simplex virus type 1 RNA in infected HeLa cells. J. Virol. **21**:889–901.

24. Wagner, E. 1983. Transcription patterns in HSV infections. Adv. Viral Oncol. **3**:239–270.

25. Weller, S., D. Aschman, W. Sacks, D. Coen, and P. Schaffer. 1983. Genetic analysis of temperature-sensitive mutants of HSV-1: the combined use of complementation and physical mapping for cistron assignment. Virology **130**:290–305.