JOURNAL OF VIROLOGY, Apr. 1988, p. 1364–1372
0022-538X/88/041364-09$02.00/0

Vol. 62, No. 4

Location, Transcript Analysis, and Partial Nucleotide Sequence of the Cytomegalovirus Gene Encoding an Early DNA-Binding Protein with Similarities to ICP8 of Herpes Simplex Virus Type 1

DAVID G. ANDERS†* AND WADE GIBSON

Virology Laboratories, Department of Pharmacology and Molecular Sciences, the Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 2 November 1987/Accepted 20 December 1987

The results presented here locate the gene encoding an early, nonvirion, single-stranded DNA-binding protein of human and simian strains of cytomegalovirus (CMV) [HCMV(Towne) DB140 and SCMV(Colburn) DB129, respectively] and provide additional evidence that this protein is the CMV homolog of the herpes simplex virus type 1 (HSV-1) major DNA-binding protein (ICP8), as proposed earlier (D. G. Anders, A. Irmiere, and W. Gibson, J. Virol. 58:253–262). The ICP8 gene was used as a probe in Southern analyses done at moderate stringency as an approach to locating similar sequences in the CMV genome. The *Bam*HI K and *Eco*RI V fragments from the center of the long unique segment of HCMV(Towne) hybridized with the ICP8 probe and were in turn used to identify corresponding sequences in the *Eco*RI D fragment of SCMV(Colburn). RNA prepared from SCMV(Colburn)-infected cells directed the in vitro synthesis of DB129. If the RNA was first hybridized with the cloned 12.5-kilobase *Eco*RI D fragment, in vitro synthesis of DB129 was specifically inhibited. Additional hybrid-arrested in vitro translation experiments with subclones spanning the *Eco*RI D fragment demonstrated that the DB129 gene is located in the left half of that fragment, approximately bisected by a *Sal*I site. RNA analyses identified 3.9- , 8.9- , and 10.0-kilobase RNA species expressed from this region. A partial nucleotide sequence of the Colburn region mapping within the boundaries of the 3.9-kilobase transcript, suspected to be the primary coding species, showed significant sequence similarity to the major DNA-binding protein gene homolog identified in B95-8 Epstein-Barr virus.

Cytomegalovirus (CMV)-infected cells contain an early, nonvirion, single-stranded DNA-binding protein (DBP) whose production is selectively enhanced by inhibiting viral DNA synthesis, e.g., simian CMV (SCMV) Colburn DB129 and human CMV (HCMV) Towne DB140 (2, 2a, 24). On the basis of extensive biochemical similarities between this CMV protein species and the major DBP of herpes simplex virus (HSV) (i.e., HSV-1 ICP8, HSV-2 ICSP11, and ICSP12), we concluded that these CMV proteins are probable homologs of the more extensively studied HSV DBP (4, 7, 10, 25, 26, 31–33, 39, 40, 43, 44). The function of the HSV major DBP is unknown, but the findings that it is required for viral DNA replication (5, 7, 36, 37, 42, 47) and may be involved in regulating viral gene expression (7, 15–17) underscore the importance of trying to learn more about it and its counterparts in other systems.

As an approach to establishing whether HCMV DB140 and SCMV DB129 are homologs to the HSV DBP, we used the cloned HSV-1 ICP8 gene to search for cross-hybridizing sequences in the CMV genome. This general strategy has been useful in many systems (1, 19, 21). The results presented in this report show that the cloned HSV ICP8 gene cross-hybridized with a specific region of the HCMV genome. Additional experiments confirmed that the corresponding region from SCMV(Colburn) encodes DB129 and identified possible coding transcripts.

* Corresponding author.

† Present address: Virology Laboratories, the Wadsworth Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, Albany, NY 12201.

MATERIALS AND METHODS

Cells and viruses. Human foreskin fibroblasts were prepared, passaged, and infected as previously described (11). Also described previously are the source and conditions of propagation (11, 13) and the methods of virion isolation (13) and DNA preparation (34) for HCMV(Towne) and SCMV (Colburn).

Plasmids. Plasmid pSG18:*Sal*A (44) containing the *Sal*I A subfragment of the HSV-1 KOS fragment *Eco*RI F (18) was obtained from D. Knipe. Plasmids pTJ148, containing the SCMV(Colburn) immediate-early region as the *Hin*dIII H fragment (22), and p*Sal*G161 containing the Colburn *Sal*I G fragment were provided by G. Hayward. pDGA1 and pDGA8 were made by excising the HCMV(Towne) *Bam*HI K fragment (28) and the SCMV(Colburn) *Eco*RI D fragments (23), respectively, from a 1.0% agarose gel and ligating the purified fragment into dephosphorylated pUC18 by standard methods (34). The identities of the resulting clones were confirmed by restriction analyses and Southern blotting. The *Eco*RI Q and V fragments (see Fig. 1) were then subcloned from pDGA1 into appropriately prepared pUC18 by the same technique to yield pDGA2 and pDGA3, respectively. The gel-purified subfragments *Xba*I B, C, D, and E (see Table 2) of the *Eco*RI D fragment (i.e., pDGA8) were subcloned into the *Xba*I site of pBluescribe M13t (Stratagene, San Diego, Calif.), as listed in Table 1. pDGA14, containing the 4.1-kilobase (kb) *Eco*RI-to-*Sal*I subfragment of *Eco*RI-D, was made by excising from pDGA8 the fragment defined by the unique *Sal*I sites in *Eco*RI-D (see Fig. 1C) and pUC18 and then religating the plasmid. The *Sal*I site in *Eco*RI-D lies very close to the leftmost *Xba*I site, and thus

TABLE 1. Description of clones

Clone	Origin	Description	Size (kb)
pSG18:*Sal*I A[a]	HSV-1(KOS)	*Sal*I-A from *Eco*RI-F	4.6
pDGA1	HCMV(Towne)	*Bam*HI-K	8.0
pDGA2	HCMV(Towne)	*Eco*RI Q subfragment of pDGA1	2.4
pDGA3	HCMV(Towne)	*Eco*RI V subfragment of pDGA1	4.1
pDGA8	SCMV(Colburn)	*Eco*RI D subfragment of pDGA1	12.1
pDGA9	SCMV(Colburn)	*Xba*I B subfragment of pDGA8	3.0
pDGA10	SCMV(Colburn)	*Xba*I C subfragment of pDGA8	1.9
pDGA11	SCMV(Colburn)	*Xba*I D subfragment of pDGA8	1.6
pDGA12	SCMV(Colburn)	*Xba*I E subfragment of pDGA8	1.0
pDGA14	SCMV(Colburn)	*Sal*I B subfragment of pDGA8	4.1
p*Sal*G161[b]	SCMV(Colburn)	*Sal*I G fragment	11.9
pTJ148[b]	SCMV(Colburn)	*Hin*dIII H fragment	10.5

[a] Obtained from D. Knipe (44).
[b] Obtained from G. Hayward.

the *Sal*I B subclone is equivalent to the *Xba*I A fragment in the experiments described.

Southern blot hybridization. Restriction enzyme-treated DNA was subjected to electrophoresis in 1% agarose gels and transferred to nitrocellulose (49) with 10X SSC (1X SSC is 0.15 M NaCl, 0.015 M sodium citrate). Hybridization stringency was controlled by varying the formamide concentration. Following 4-h prehybridizations in the same solution without probe, filters were allowed to hybridize overnight at 42°C with about 10^7 cpm of ^{32}P-labeled probe in solutions containing 1.0 M NaCl, 0.01 M Tris (pH 7.4), 0.2% sodium dodecyl sulfate (SDS), 0.01 M EDTA, 100 μg of salmon sperm DNA per ml, 5X Denhardt solution (1X Denhardt solution is 0.02% Ficoll–0.02% polyvinylpyrrolidone–0.02% bovine serum albumin), and deionized formamide, at the concentration given in the text for each experiment. Probes were ^{32}P-labeled by nick translation (46) or by the random oligomer-primed synthesis method (8, 9), as indicated in the text. After hybridization, filters were washed twice for 10 min in 2X SSC–0.1% SDS at room temperature, and then washed twice for 1 h at 62°C in 6X SSC–0.1% SDS (low stringency) or twice for 1 h at 62°C in 0.1X SSC (more stringent). The nitrocellulose filter was air dried, and the bound probe was visualized by exposure to Kodak X-Omat AR film (Eastman Kodak Co., Rochester, N.Y.) under Cronex intensifying screens (DuPont Co., Wilmington, Del. [30]).

RNA isolation and poly(A)$^+$ selection. Total RNA was extracted from SCMV(Colburn)-infected cells at 72 h after infection or at 96 h after infection for cultures containing phosphonoformic acid (200 μg/ml), by the guanidinium isothiocyanate method, followed by ultracentrifugation through a CsCl cushion (6). Polyadenylated RNA was selected by oligo(dT)-cellulose column chromatography (34).

In vitro translation and hybrid-arrested in vitro translation. Cell-free translation, directed by total infected-cell RNA, was performed by using a rabbit reticulocyte lysate (Amersham Corp., Arlington Heights, Ill.) with [^{35}S]methionine (translation grade; New England Nuclear Corp., Boston, Mass.), according to the instructions of the supplier. Hybrid-arrested in vitro translation was done as described by Paterson et al. (38). Hybridization was carried out for 2 h at 42°C in 25 μl of a solution containing 80% deionized formamide, 0.01 M PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4), and 0.4 M NaCl. The reactions were terminated by adding 200 μl of ice-cold water and 25 μg of calf liver tRNA (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) and then the mixture was divided into two equal portions. One of these was heat denatured by incubation for 1 min in a boiling water bath and quick-chilled in a dry ice-isopropanol bath; the other was not heat denatured. Following ethanol precipitation, the reaction products from both the heat-treated and untreated sets were suspended in 1.0 μl of water and translated as described above. Portions of the in vitro translation reactions were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by fluorography, as previously described (29), and the remaining volume was immunoprecipitated.

Immunoprecipitation. The production and specificity of the rabbit monospecific polyclonal anti-DB129 serum used for these studies has been described previously (2a). Immunoprecipitation was performed by mixing 10 μl of the translation reaction with 20 μl of 2X RIPA buffer (1X RIPA is 0.05 M Tris [pH 7.4]–0.05 M NaCl–0.1% SDS–0.1% Nonidet P-40–0.1% sodium deoxycholate–0.001 M phenylmethylsulfonyl fluoride) and 10 μl of immune serum. Following incubation overnight at 4°C, immune complexes were collected on protein A-Sepharose beads as described before (2a) and subjected to SDS-PAGE.

Northern analysis. RNA was resolved in 1% agarose gels (11 by 14 cm) which were prerun for 30 min at 60 V and contained formaldehyde (34), 0.05 M morpholinepropanesulfonic acid (MOPS) (pH 7.0), and 0.001 M EDTA. The electrode buffer was mixed periodically during electrophoresis at 100 V for 2 h. RNA was transferred to nitrocellulose in 20X SSC, and the nitrocellulose was baked at 80°C for 2 h under vacuum. After prehybridization in the same buffer, hybridization was done for 24 h at 42°C in a buffer containing 50% formamide, 5X SSC, 0.02 M sodium phosphate (pH 6.5), 5X Denhardt solution, 0.05% SDS, and 100 μg of salmon sperm DNA per ml. Approximately 10^7 cpm of probe, ^{32}P-labeled as described in the text, was used for each hybridization. After hybridization, the filter was rinsed three times for 10 min each at room temperature in 2X SSC containing 0.1% SDS and then twice for 1 h each at 50°C in 0.1X SSC containing 0.1% SDS and was finally air dried. Bound probe was visualized as described above.

DNA sequencing. The 1.0-kb *Xba*I E subfragment of SCMV(Colburn) *Eco*RI-D was subcloned into M13mp19, and the sequence of about 300 base pairs (bp) from the right-hand end was determined by using the dideoxy method (48) with ^{35}S-labeled nucleotides.

RESULTS

Identification of CMV genomic DNA homologous to the HSV-1 major DBP gene. DNA from HCMV(Towne) virions was cleaved separately with *Bam*HI, *Hin*dIII, and *Eco*RI; resolved by electrophoresis through a 1.0% agarose gel; and transferred to nitrocellulose (49). The filter was then probed with ^{32}P-labeled pSG18:*Sal*IA insert DNA containing more than 90% (i.e., all but 3′ 243 bp) of the HSV-1 ICP8 gene (44; Table 1), by using conditions of moderate stringency (30% formamide, T_m −41°C [21]), as described in Materials and Methods. Weak but specific hybridization was seen within the *Hin*dIII A, *Bam*HI K, and *Eco*RI V fragments (Fig. 1B),

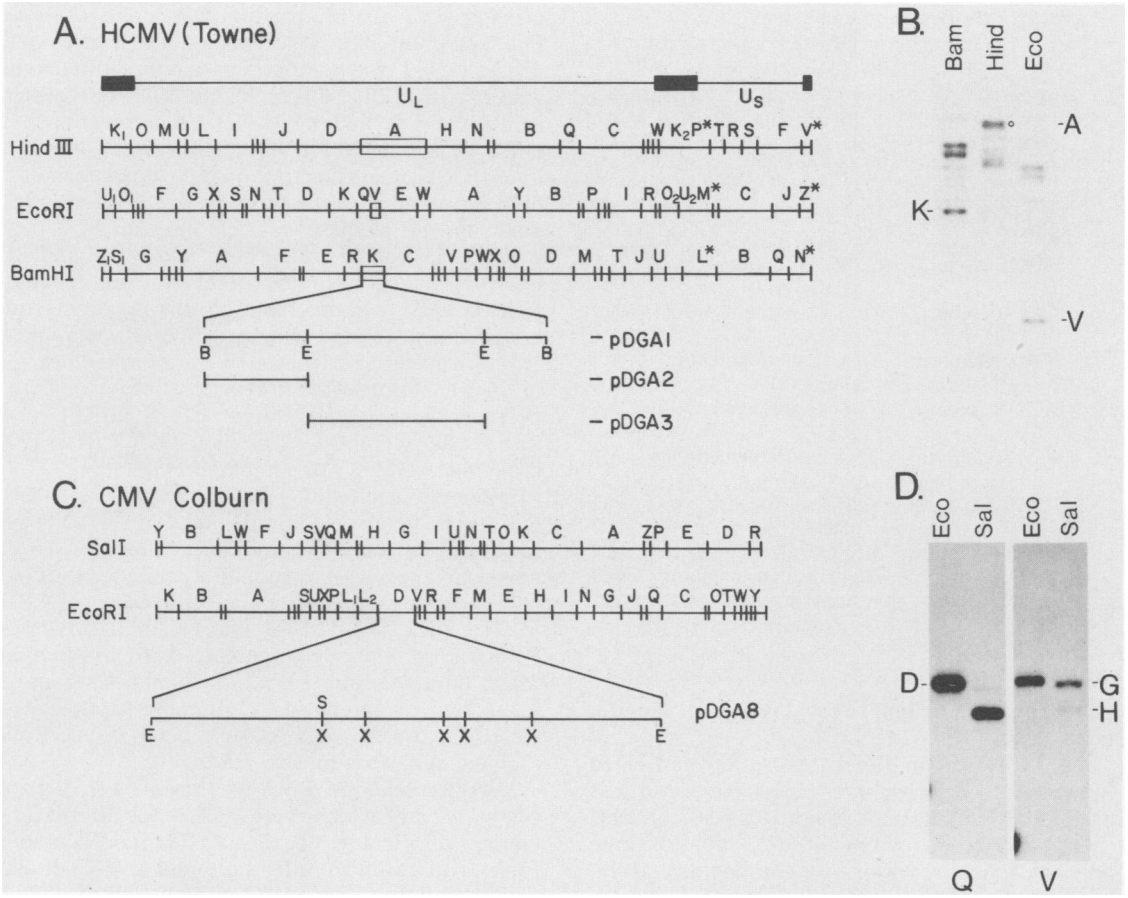

FIG. 1. HCMV sequences which cross-hybridize with the HSV-1 ICP8 gene and corresponding sequences in the SCMV(Colburn) genome. (A) *Hin*dIII, *Eco*RI, and *Bam*HI restriction maps of HCMV(Towne) (28). The top line depicts the general organization of the HCMV genome, indicating the long and short unique regions (U_L and U_S, respectively). Flanking inverted repeats are shown as solid boxes. The 8.0-kb *Bam*HI K fragment is expanded, and the positions of *Eco*RI sites defining the subcloned *Eco*RI Q and *Eco*RI V subfragments are marked. Fragments showing consistent specific hybridization in this experiment and others mentioned in the text are boxed. (B) Southern blot analysis of HCMV DNA. Southern blots of HCMV DNA which was cleaved with the indicated enzymes were probed, as described in Materials and Methods, with the HSV-1 ICP8 gene (i.e., pSG18:*Sal*A insert DNA) ^{32}P-labeled by the random oligomer-primed synthesis method (8, 9). The hybridization buffer contained 30% formamide, corresponding to T_m −41°C (using a G+C percentage of 57 [45]). An autoradiogram is shown here. The fragments which most strongly and reproducibly hybridized with the ICP8 probe (i.e., *Bam*HI-K, *Hin*dIII-A, and *Eco*RI-V) are indicated. The circle adjacent to the *Hin*dIII lane denotes the A fragment. (C) *Sal*I and *Eco*RI restriction maps of SCMV(Colburn) (23). The *Eco*RI D fragment is expanded to indicate the positions of the *Xba*I and *Sal*I sites (see also Table 2). (D) Southern blot analysis of SCMV(Colburn) DNA. Identical Southern blots of SCMV(Colburn) DNA, cleaved with the indicated enzymes, were probed as described in Materials and Methods with either nick-translated pDGA2 (Q), or nick-translated pDGA3 (V). The hybridization buffer contained 40% formamide (T_m −34°C, calculated as described in the legend to panel B). The positions of *Eco*RI-D, *Sal*I-G, and *Sal*I-H are indicated. Abbreviations: E, *Eco*RI; B, *Bam*HI; S, *Sal*I; and X, *Xba*I.

each of which is located around 0.4 on the physical map of HCMV(Towne) (Fig. 1A). Although some nonspecific hybridization was observed the specificity of hybridization to the fragments noted above is supported by the following considerations. (i) Hybridization at this position was consistent in all three restriction maps. (ii) Several other bands that appeared to hybridize well with the ICP8 gene probe (e.g., those above the *Bam*HI K fragment) contained multiple restriction fragments and thus a greater concentration of nonspecifically hybridized probe. (iii) At higher hybridization or wash stringencies, nonspecific hybridization was reduced, relative to specific hybridization with *Bam*HI-K and *Eco*RI-V; hybridization with these fragments was observed up to T_m −34°C (not shown). These results suggested that ICP8-like sequences are present in the HCMV(Towne) *Eco*RI V fragment. The HCMV(Towne) *Bam*HI K fragment encompassing the cross-hybridizing regions was cloned into the *Bam*HI site of pUC18, and subclones containing the left-hand *Bam*HI-*Eco*RI fragment (*Eco*RI Q subfragment) and *Eco*RI-V were constructed. These clones were designated pDGA1, pDGA2, and pDGA3, respectively (Fig. 1A; Table 1).

To identify the corresponding region of the SCMV(Colburn) genome, DNA from Colburn virions was cleaved with *Eco*RI or *Sal*I, sets of both digests were subjected to electrophoresis through a 1.0% agarose gel and transferred to nitrocellulose, and the resulting blots were probed at T_m −34°C with either of two contiguous subclones from HCMV (Towne) *Bam*HI-K (i.e., clones pDGA2 or pDGA3; Fig. 1D). Both of these HCMV(Towne) probes hybridized selectively with the SCMV(Colburn) *Eco*RI D fragment, centered around 0.4 on the physical map (Fig. 1C). They showed different specificities, however, with respect to the *Sal*I fragments comprising the right and left halves of SCMV(Col-

burn) *Eco*RI-D. Specifically, the rightward *Eco*RI V fragment of HCMV(Towne) hybridized preferentially with the rightward side of SCMV(Colburn) *Eco*RI-D (i.e., the *Sal*I-G fragment), and the leftward *Eco*RI Q fragment of HCMV (Towne) hybridized preferentially with the leftward side of SCMV(Colburn) *Eco*RI-D (i.e., *Sal*I-H) (Fig. 1D). This result demonstrates that these regions of the Towne and Colburn genomes have the same orientation.

Hybrid-arrested in vitro synthesis of DB129 with cloned Colburn *Eco*RI-D. RNA used in this experiment was prepared from SCMV(Colburn)-infected cells grown in phosphonoformate (50) in an effort to selectively increase synthesis of the DB129 message (2). This treatment did not appear to cause any major qualitative changes in the RNA species seen in Northern analyses (see Fig. 3, lane 2, and compare with Fig. 4, Sal-G). The RNA was translated in vitro in the presence of [^{35}S]methionine, and the products were analyzed by SDS-PAGE, either before (Fig. 2A) or after (Fig. 2B) immunoprecipitation with monospecific polyclonal anti-DB129 serum (2a). Anti-DB129 serum, but not preimmune serum (not shown), precipitated a single band which corresponded to DB129 in extracts prepared from infected cells (Fig. 2B, No Plasmid). In vitro synthesis of DB129 was nearly eliminated when the infected-cell RNA was hybridized with the entire *Eco*RI D fragment (i.e., pDGA8) before translation (Fig. 2B and D, Eco D). The location of sequences which inhibited in vitro synthesis of DB129 was better defined by including in this experiment a series of subclones spanning the Colburn *Eco*RI D fragment (Fig. 1C and 4; Table 1). Three of the subclones, comprising the left-hand end of the Colburn *Eco*RI D fragment (i.e., *Xba*I C, *Xba*I E, and *Sal*I B), inhibited synthesis of DB129 by six- to eightfold compared with controls (Fig. 2B and D). Clones containing fragments from the right-hand end of Colburn *Eco*RI-D (i.e., *Xba*-B and *Xba*-D) did not affect translation of the DB129 message (Fig. 2B and D). Synthesis of other proteins was unaffected by hybridization with any of the subclones (Fig. 2A). As a control, cloned DNA from another region of the genome was tested. pTJ148, a plasmid containing the Colburn immediate-early region, was used and found to have no effect on the synthesis of DB129 (TJ148, Fig. 2B and D) but selectively reduced synthesis of the Colburn major immediate-early protein, IE94 (TJ148, Fig. 2A). Synthesis of other proteins appeared to be unaffected. Heating all hybrids before translation restored synthesis of the respective proteins (compare Fig. 2B and C; data for pTJ148 not shown).

Northern analysis. To identify RNA species that were transcribed from the *Eco*RI D region of the SCMV(Colburn) genome, the *Eco*RI D fragment (i.e., pDGA8) was nick translated and used as a probe in Northern analyses of polyadenylated RNA prepared from Colburn-infected cells (Fig. 3). As a control, the blot was probed with pTJ148, which contains the Colburn major immediate-early region. pTJ148 has been shown to detect a 2.5-kb Colburn transcript in immediate-early RNA from cycloheximide-treated cells

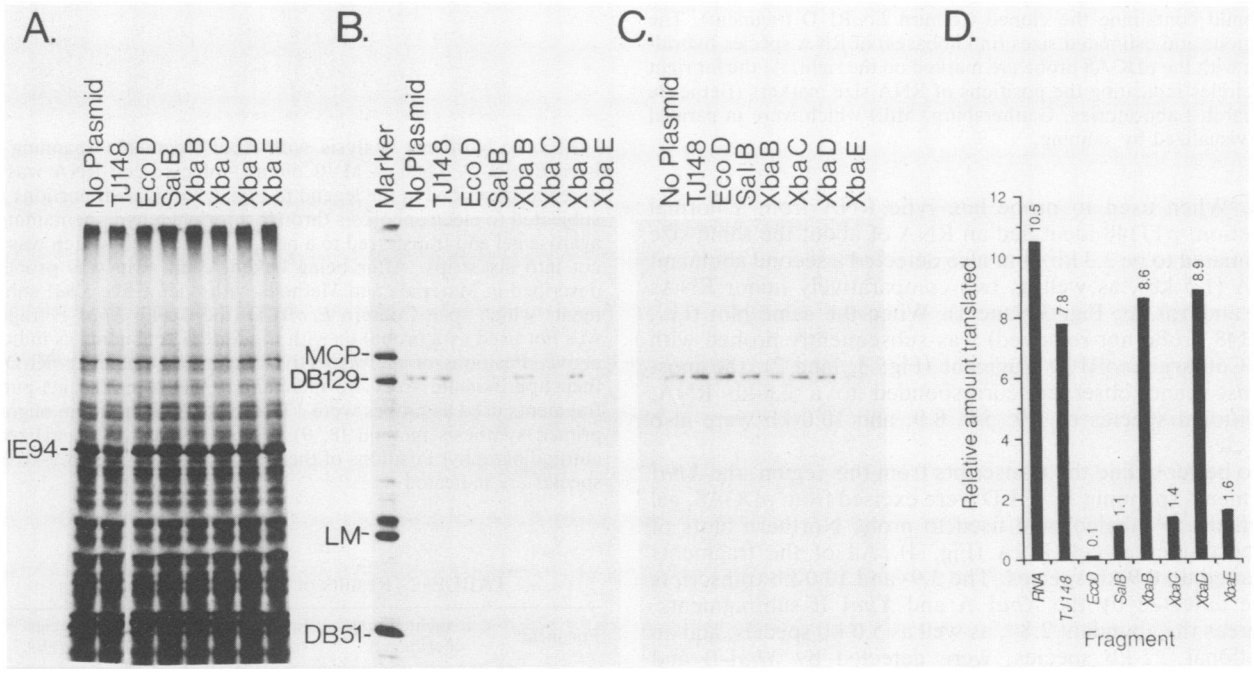

FIG. 2. Verification that Colburn *Eco*RI-D contains DB129-coding sequences. Total SCMV(Colburn) infected-cell RNA used for these experiments was isolated as described in Materials and Methods at 96 h after infection from cultures containing phosphonoformic acid (200 μg/ml). Portions of the RNA were incubated under hybridization conditions with either no plasmid or a plasmid containing the cloned fragment indicated above each lane. After hybridization, each reaction mixture was divided in half; half was heated at 100°C for 1 min to denature the hybrids, and the other half was left unheated. After precipitation, all preparations were translated in vitro as described in Materials and Methods. (A) A sample from each translation reaction was resolved by SDS-PAGE. Only the products of the unheated set of reactions are shown; the position of the major immediate-early protein (IE94; [12, 13]) is indicated. A sample from each translation reaction done using nondenatured (B) or denatured (C) RNA was incubated with rabbit monospecific anti-DB129 serum (2a), and the resulting immune precipitates were subjected to SDS-PAGE. The marker lane contained a [^{35}S]methionine-labeled extract of SCMV(Colburn)-infected cells. The positions of the major capsid protein (MCP), DB129, the lower matrix protein (LM), and a delayed-early DBP, DB51 (14), are indicated (11). (D) To estimate the relative amount of DB129 which was immunoprecipitated from each translation reaction, the fluorogram shown in panel B was analyzed by densitometry as previously described (2), and the results are presented as a bar graph.

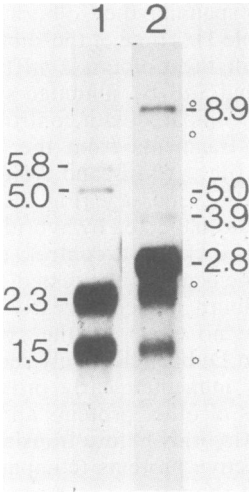

FIG. 3. RNA species detected by the Colburn *Eco*RI D fragment. (Lane 1) RNA was isolated from SCMV(Colburn)-infected cells at 72 h after infection, and polyadenylated RNA was selected on oligo(dT)-cellulose. Following electrophoresis through a formaldehyde-containing 1% agarose gel and transfer to nitrocellulose, the RNA was probed with nick-translated TJ148 plasmid DNA (i.e., the region of the Colburn IE94 gene), the blot was washed, and bound probe was visualized as described in Materials and Methods. The positions and estimated sizes of RNA species identified by this control probe are indicated on the left. (Lane 2) The blot, without removing the pTJ148 probe, was reprobed with approximately the same number of counts per minute of nick-translated pDGA8 (plasmid containing the cloned Colburn *Eco*RI D fragment). The positions and estimated sizes (in kilobases) of RNA species hybridizing with the pDGA8 probe are marked on the right. At the far right are circles indicating the positions of RNA size markers (Bethesda Research Laboratories, Gaithersburg, Md.) which were in parallel and visualized by staining.

(22). When used to probe late lytic RNA from a normal infection, pTJ148 identified an RNA of about the same size (estimated to be 2.3 kb), but also detected a second abundant RNA (1.5 kb), as well as two comparatively minor RNAs (5.8 and 5.0 kb; Fig. 3, lane 1). When the same blot (i.e., pTJ148 probe not removed) was subsequently probed with the Colburn *Eco*RI D fragment (Fig. 3, lane 2), the most intense band observed corresponded to a 2.8-kb RNA. Additional species of 3.9, 5.0, 8.9, and 10.0 kb were also detected.

To better define the transcripts from this region, the *Xba*I fragments spanning *Eco*RI-D were excised from pDGA8, gel purified, ^{32}P-labeled, and used to probe Northern blots of Colburn-infected-cell RNA (Fig. 4). All of the fragments detected an 8.9-kb species. The 3.9- and 10.0-kb transcripts were detected by the *Xba*I A and *Xba*I E subfragments, whereas the abundant 2.8- , as well as 5.0-kb species, and an additional 2.2-kb species, were detected by *Xba*I-B and *Xba*I-D. *Xba*I-C hybridized with both sets of RNA species. The results of this experiment are summarized in Table 2.

Partial nucleotide sequence analysis. The results of the experiments described above indicated that the DB129 gene was located in the left half of Colburn *Eco*RI-D, spanning the *Xba*I E fragment. To establish whether this region was indeed similar to the ICP8 coding region and to aid in defining its organization, a 315-bp sequence from the right end of pDGA12 (i.e., beginning at the junction of the *Xba*I E and *Xba*I C subfragments, map position 0.393, and proceeding leftward; Table 2) was determined by using the method of

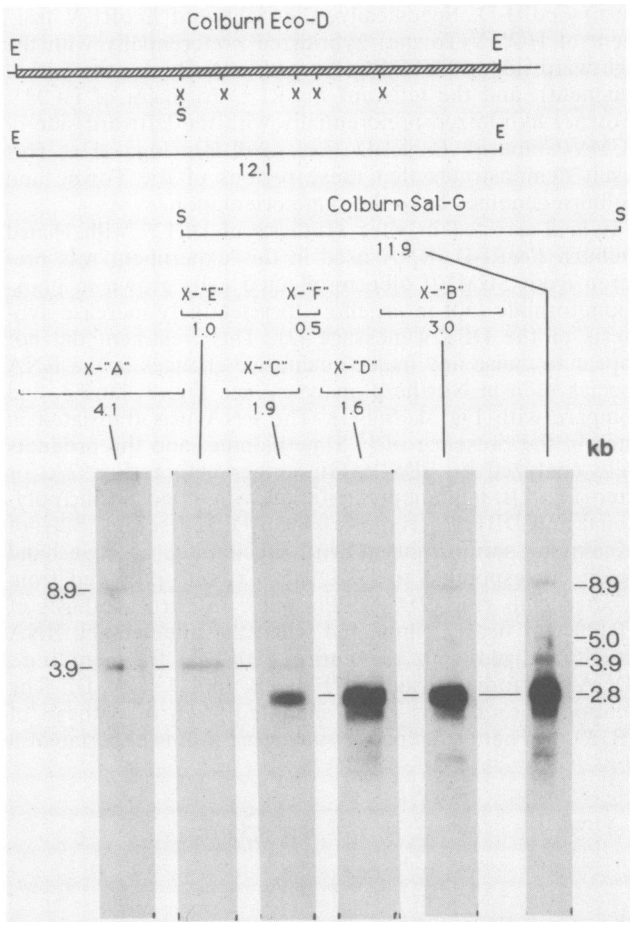

FIG. 4. Northern analysis with *Xba*I fragments spanning Colburn *Eco*RI-D. Total SCMV(Colburn)-infected-cell RNA was prepared as described in the legend to Fig. 2. Equivalent portions were subjected to electrophoresis through a formaldehyde-containing 1% agarose gel and transferred to a nitrocellulose sheet which was then cut into six strips. After being baked, each strip was probed as described in Materials and Methods with one of five *Xba*I subfragments which span Colburn *Eco*RI-D (the 0.5-kb *Xba*I F fragment was not used as a probe) or with the *Sal*I G fragment, as indicated above. Positions of the respective fragments within *Eco*RI-D and their approximate sizes are indicated at the top. The gel-purified fragments used as probes were ^{32}P-labeled by the random oligomer-primed synthesis method (8, 9). Bound probe was visualized by autoradiography. Positions of the 2.8- , 3.9- , 5.0- , and 8.9-kb RNA species are indicated.

TABLE 2. Results of Northern hybridizations

Fragment[a]	Estimated map coordinates[b]	Fragment size (kb)	RNA species detected (kb)[c]
*Xba*I-A	0.370–0.388	4.1	3.9, 8.9, 10
*Xba*I-B	0.411–0.425	3.0	2.2, 2.8, 5.0, 8.9
*Xba*I-C	0.393–0.415	1.9	2.2, 2.8, 3.9, 5.0, 8.9, 10
*Xba*I-D	0.404–0.411	1.6	2.2, 2.8, 5.0, 8.9
*Xba*I-E	0.388–0.393	1.0	3.9, 8.9, 10
pSalG161	0.388–0.442	11.9	2.2, 2.8, 3.9, 5.0, 8.9, 10

[a] Gel-purified fragments excised from pDGA8 (*Eco*RI-D) are named in descending order of size.

[b] On the basis of a 220,000-bp Colburn genome with the left-hand side of the *Eco*RI D fragment taken to be at coordinate 0.37.

[c] Sizes estimated by comparison with the Bethesda Research Laboratories RNA ladder.

Sanger et al. ([48] Fig. 5). The amino acid sequence predicted for each open reading frame of the determined sequence was compared with that predicted from the HSV-1 ICP8 gene sequence (43) and with that of the homologous B95-8 Epstein-Barr virus (EBV) open reading frame BALF2 (3, 43). Clear similarities with the sequence of EBV BALF2 were identified by a homology matrix analysis (Fig. 6A). With numbering beginning at the right end of EBV *Bam*HI-A, the amino acids of BALF2 stretch leftward from position 617 through position 1745; sequence similarity was observed roughly between residues 820 and 910. An alignment of the two sequences yielded greater than 30% identity over a 96-amino acid stretch (Fig. 6B). Compared with EBV, similarities between this short portion of the DB129 gene sequence and the HSV ICP8 gene were minimal, although some common features were observed. Most notable of these features was the sequence Ala-Leu-Arg at positions 33 to 35 in the Colburn sequence, which also is found in EBV BALF2. Optimal alignment, which yielded 21% identity over amino acids 10 to 105 of the Colburn sequence, required introduction of two single-residue gaps in the DB129 sequence and one in the HSV-1 sequence (not shown).

DISCUSSION

This report identifies the coding region for an SCMV(Colburn) early DBP DB129 and by extension, its HCMV (Towne) counterpart DB140 (2, 2a, 11, 13). A clone containing over 90% of the HSV-1 ICP8 gene was initially used to locate cross-hybridizing sequences in the HCMV genome.

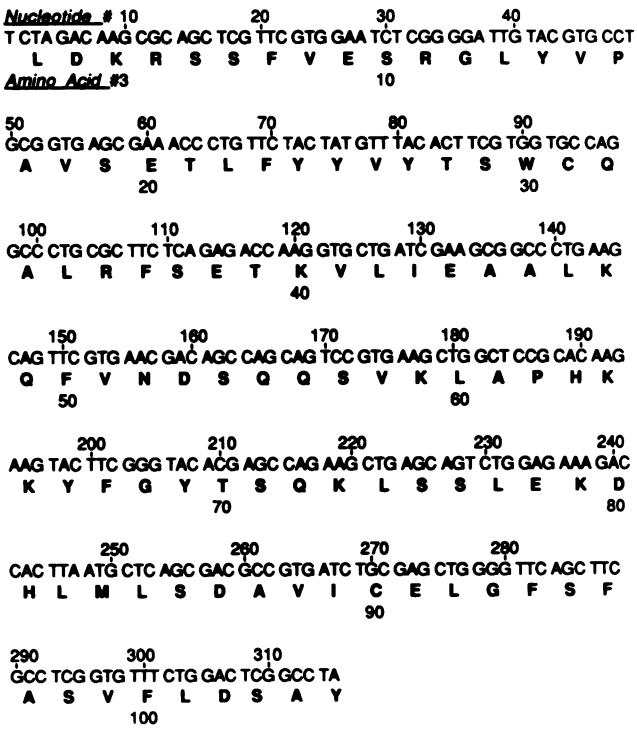

FIG. 5. Partial nucleotide sequence of the *Xba*I E subfragment from Colburn *Eco*RI-D. Nucleotide sequence was determined as described in Materials and Methods. Nucleotides are numbered from the right end of *Xba*I-E as drawn in Fig. 1 and 4, with the first base assigned as 1. Amino acids are indicated by standard single-letter abbreviations. The reading frame giving the deduced amino acid sequence shown was selected for its amino acid sequence similarity to EBV BALF2 (see Fig. 6).

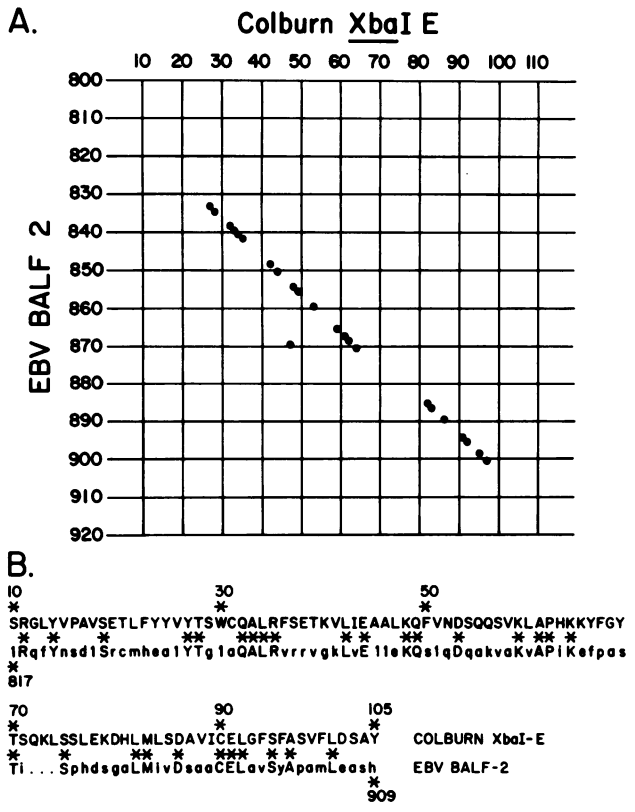

FIG. 6. (A) Homology matrix comparison of deduced amino acid sequences from the Colburn *Xba*I E subfragment of *Eco*RI-D and B95-8 EBV BALF2 (3). Comparison was done by using Pustell programs (International Biotechnologies, Inc., New Haven, Conn. [41]). The Colburn sequence is plotted on the horizontal axis, and the EBV BALF2 sequence is plotted on the vertical axis. A range of 30 amino acids was used, and a point was plotted only when 25% or more of the residues in a given window matched. Other parameters were set to the following numbers: hash level, 1; scale factor, 0.95; step, 1; and jump, 1. The figure shown is redrawn from the original plot. (B) An alignment of the Colburn and EBV BALF2 sequences. A three-amino-acid gap was introduced into the EBV sequence to maximize matches. Equivalent residues are marked by an asterisk. The numbering system is explained in the text.

An HCMV restriction fragment containing the identified region was then used to locate similar sequences in the SCMV(Colburn) genome to take advantage of previously characterized antisera that had been raised against the Colburn DB129 protein (2a). Hybrid-arrested translation experiments and Northern analyses were then used to correlate the protein with its coding sequence and to identify possible coding transcripts. Finally, a partial nucleotide sequence obtained from the Colburn DB129 gene established its similarity to the gene for the EBV DBP that is homologous to HSV ICP8.

The first series of experiments demonstrated that the cloned HSV ICP8 gene cross-hybridized with DNA sequences around 0.4 on the physical map of the HCMV genome. Hybridization was detected only under conditions of moderate stringency (equivalent to T_m −34°C or lower), indicating limited nucleotide sequence similarity. A clone containing the identified HCMV sequence was constructed and used in experiments which established that SCMV(Colburn) contains a related, colinear (i.e., around 0.4 on its physical map) sequence which is present in the same relative

orientation. These CMV sequences are themselves approximately colinear with the position of the ICP8 gene in the HSV-1 genome (43) and correspond to a region of the HCMV(AD169) genome with sequence similarity to the HSV ICP8 gene (T. Kouzarides, A. Bankier, K. Weston, S. C. Satchwell, S. Beck, and B. G. Barrell, Abstr. 11th Int. Herpesvirus Workshop, 1986, p. 60; B. Barrell, personal communication).

Hybrid-arrested in vitro translation experiments which were done by using subclones of the cross-hybridizing Colburn region established that DB129 is coded for within the left half of the *Eco*RI D fragment. Positive and negative controls demonstrated the specificity of the hybrid-arrested translation assay. This approach was used instead of hybrid selection for two reasons. (i) It is more specific for sequences within the coding body of the transcript (35). (ii) The anticipated large size and low abundance of DB129 mRNA diminished the chances for its successful selection. The availability of a monospecific antiserum (2a) provided a means to identify the small amounts of translation product in these experiments.

Northern analyses identified six RNA species that were recognized by the set of subclones used for hybrid-arrested translation experiments (Fig. 4 and 7). Of these, the 3.9- and 10.0-kb RNAs are the most likely to be transcripts for DB129, since they hybridized with only those *Eco*RI subfragments that hybrid arrested DB129 synthesis (*Xba*I-A, *Xba*I-E, and *Xba*I-C). The 2.8- and 5.0-kb species arise from sequences located to the right (as shown in Fig. 7) of the DB129 gene, and the predicted maximum size of a protein encoded by the abundant 2.8-kb RNA (i.e., 100 kilodaltons [kDa]) would be somewhat smaller than the size estimated by SDS-PAGE for DB129 (i.e., 129 kDa). An arrangement of the RNA species detected with *Eco*RI-D that is consistent with the results presented in this report is shown in Fig. 7. Results of the sequencing experiments indicated that the DB129 gene is transcribed from right to left, as drawn in this figure. Since the 3.9-kb transcript is large enough to encode DB129, it is suspected to be the primary coding species. We note that the HSV ICP8 gene is transcribed in the same direction as is the DB129 gene and gives rise to two apparently 5′ coterminal RNAs of 4.5 and 10.2 kb (20, 44) that are close in size to the Colburn 3.9- and 10.0-kb RNAs and may be counterparts. The report by Kemble et al. (24) that the mRNA for the homologous HCMV(Towne) gene is 10 to 12 kb in size suggests that processing of the transcript for the HCMV gene may differ from that of the SCMV(Colburn) and HSV-1 genes.

The partial nucleotide sequence obtained for the Colburn DB129 gene, and its similarity to the BALF2 open reading frame of EBV, aided in positioning the DB129 gene more precisely, as follows. Given that the DB129 and EBV BALF2 proteins are about the same size (i.e., observed, 129 kDa, and predicted, 123 kDa, respectively) and assuming that no major rearrangements exist between these CMV and EBV genes, the DB129 amino acid sequence predicted here begins about 200 residues downstream from its amino terminus. Therefore, the 5′ end of the DB129 coding region is predicted to begin about 0.6 kb to the right of the *Xba*I E-*Xba*I C junction and end about 2 kb to the left of the *Sal*I site in the *Eco*RI D fragment. This position for the 3′ end is consistent with the finding that cDNA clones to DB129 extend leftward from the *Xba*I E subfragment about 2 kb into the *Xba*I A fragment (L. Robson, D. Anders, and W. Gibson, unpublished results).

We have previously identified the HCMV counterpart to

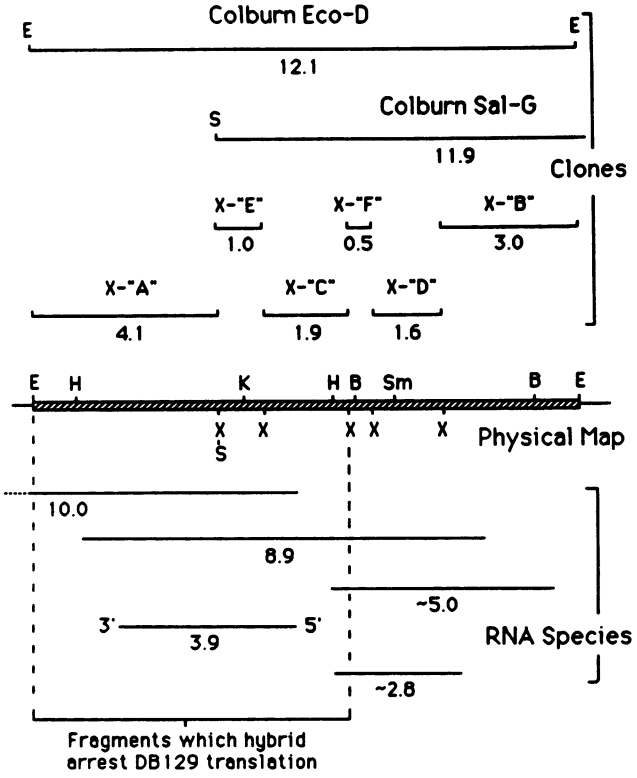

FIG. 7. Summary relating hybrid-arrested in vitro translation and Northern blot results. The approximate positions of transcripts are shown, and the fragments which hybrid arrested in vitro synthesis of DB129 are indicated. The transcriptional orientation of the 3.9-kb RNA, suspected to encode DB129, is predicted from the nucleotide sequence data. Restriction sites are as follows: Sm, *Sma*I; H, *Hin*dIII; K, *Kpn*I; others are as indicated in the legend to Fig. 1. Sizes given are in kilobases.

DB129 (2, 13), DB140, and have shown it to be immunologically reactive with monospecific antisera prepared against purified DB129 (2a). Evidence presented here indicates that the gene for DB140 is encoded within the *Bam*HI K fragment of HCMV(Towne), around 0.4 on the physical map. Cross-hybridization experiments showed that the SCMV(Colburn) region containing the DB129 gene corresponds to HCMV (Towne) sequences spanning the *Eco*RI Q-*Eco*RI V junction (Fig. 1). Taken together with the Colburn nucleotide sequence data, these results predict that the DB140 gene extends leftward from Towne *Eco*RI-V into the adjacent *Eco*RI Q fragment. This location corresponds to the identified ICP8-like sequences of HCMV(AD169) (unpublished data; Kouzarides et al., Abstr. 11th Int. Herpesvirus Workshop, 1986; B. Barrell, personal communication) and the HCMV region which was detected by expression cloning and determined to encode an ICP8-like protein with properties of DB140 (24). Furthermore, the partial nucleotide sequence that we obtained from the Colburn DB129 gene showed striking (i.e., 86%) similarity with the HCMV(AD169) ICP8-like gene identified in nucleotide sequence analysis of the HCMV(AD169) genome (D. Anders, W. Gibson, and B. Barrell, unpublished results). The observation that the ICP8 gene probe did not hybridize as well with the HCMV(Towne) *Eco*RI Q fragment as it did with the *Eco*RI V fragment suggests that the 5′ end of the gene is more conserved between HSV and CMV than the 3′ end or that the 5′ end may contain a small conserved region.

Three lines of evidence support the conclusion that CMV proteins DB129 and DB140 are homologs of HSV ICP8, as suggested previously (2). First, they exhibit similar biochemical properties, including an early time of synthesis and comparatively low abundance, a nuclear location, absence from the mature virion, a size of approximately 130 kDa, and the capacity to bind to DNA (2, 2a, 24). Second, the genes for these CMV proteins cross-hybridize with the ICP8 gene (Fig. 1B), are approximately colinear on the physical map (Fig. 1A and C), are transcribed in the same relative direction (Fig. 7), and in the case of the DB129 gene, give rise to a putative coding transcript which is similar in size to that for ICP8 (Fig. 3, 4, and 7). And third, a partial nucleotide sequence of the gene that encodes DB129 shows clear similarity to the EBV open reading frame BALF2 (Fig. 6), a putative homolog of the HSV ICP8 gene (43). Similarities such as these in both individual gene sequences, and in the relative arrangements of blocks of genes (27), add to evidence for an evolutionary link among the herpes group viruses. It is noteworthy that the EBV BALF2 sequence from around amino acids 840 through 900 (as numbered here), like the corresponding Colburn sequence (Fig. 6), shows only weak similarity to the corresponding ICP8 sequence, although flanking sequences show a clear similarity (43). In this regard, the greater similarity of the CMV and EBV genes in this region, by comparison with HSV-1, may be of functional or evolutionary significance.

In summary, we have identified the coding region for DB129 and provided evidence by hybridization experiments and sequence comparison that it and HCMV DB140 represent CMV homologs of this herpesvirus group of common DBP species. The identification and cloning of these CMV genes will facilitate their biochemical and genetic analysis and help establish their role during infection.

ACKNOWLEDGMENTS

We thank David Knipe for clone pSG18:*Sal*A, Gary Hayward for clones p*Sal*G161 and pTJ148, Lori Robson for the oligo(dT)-selected RNA, and Marie Hardwick for providing the use of her computer for sequence comparisons and Ivan Auger for performing some additional comparisons by using GENALIGN on the BIONET computer resource and FASTP. We also thank Bart Barrell for helpful information regarding the HCMV(AD169) nucleotide sequence and Bob LaFemina and Gary Hayward and members of his laboratory for helpful discussions. We appreciate the excellent technical assistance of Donna Woods and Jennifer Kidd.

This work was aided by Public Health Service grants AI 13718, AI 22711, and AI 19373 from the National Institute of Allergy and Infectious Diseases and by American Cancer Society Institutional grant IN-11X to W.G. D.G.A. was supported during the initial period of this work by Public Health Service Individual National Research Service Award AI 07293-01 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Alestrom, P., A. Stenlund, P. Li, A. Bellett, and U. Pettersson.** 1982. Sequence homology between avian and human adenoviruses. J. Virol. **42**:306–310.
2. **Anders, D. G., A. Irmiere, and W. Gibson.** 1986. Identification and characterization of a major early cytomegalovirus DNA-binding protein. J. Virol. **58**:253–262.
2a. **Anders, D. G., J. R. Kidd, and W. Gibson.** 1987. Immunological characterization of an early cytomegalovirus single-strand DNA binding with similarities to the HSV major DNA-binding protein. Virology **161**:579–588.
3. **Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
4. **Bayliss, G. J., H. S. Marsden, and J. Hay.** 1975. Herpes simplex virus proteins: DNA-binding proteins in infected cells and in the virus structure. Virology **68**:124–134.
5. **Chiou, H. C., S. K. Weller, and D. M. Coen.** 1985. Mutations in the herpes simplex virus major DNA-binding protein gene leading to altered sensitivity to DNA polymerase inhibitors. Virology **145**:213–226.
6. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
7. **Conley, A. J., D. M. Knipe, P. C. Jones, and B. Roizman.** 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of γ polypeptides. J. Virol. **37**:191–206.
8. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
9. **Feinberg, A. P., and B. Vogelstein.** 1984. Addendum. A technique for radiolabeling fragments to high specific activity. Anal. Biochem. **137**:266–267.
10. **Fenwick, M. L., M. J. Walker, and J. M. Petkevich.** 1978. On the association of virus proteins with the nuclei of cells infected with herpes simplex virus. J. Gen. Virol. **39**:519–529.
11. **Gibson, W.** 1981. Structural and nonstructural proteins of strain Colburn cytomegalovirus. Virology **111**:516–537.
12. **Gibson, W.** 1981. Immediate-early proteins of human cytomegalovirus strains AD169, Davis, and Towne differ in electrophoretic mobility. Virology **112**:350–354.
13. **Gibson, W.** 1983. Protein counterparts of human and simian cytomegaloviruses. Virology **128**:391–406.
14. **Gibson, W., T. Murphy, and C. Roby.** 1981. Cytomegalovirus-infected cells contain a DNA-binding protein. Virology **111**:251–262.
15. **Godowski, P. J., and D. M. Knipe.** 1983. Mutations in the major DNA-binding protein gene of herpes simplex virus type 1 result in increased levels of viral gene expression. J. Virol. **47**:478–486.
16. **Godowski, P. J., and D. M. Knipe.** 1985. Identification of a herpes simplex virus function that represses late gene expression from parental viral genomes. J. Virol. **55**:357–365.
17. **Godowski, P. J., and D. M. Knipe.** 1986. Transcriptional control of herpesvirus gene expression: gene functions required for positive and negative regulations. Proc. Natl. Acad. Sci. USA **83**:256–260.
18. **Goldin, A. L., R. M. Sandri-Goldin, M. Levine, and J. C. Glorioso.** 1981. Cloning of herpes simplex virus type 1 sequences representing the whole genome. J. Virol. **38**:50–58.
19. **Heilbronn, R., G. Jahn, A. Bürkle, U.-K. Freese, B. Fleckenstein, and H. zur Hausen.** 1987. Genomic localization, sequence analysis, and transcription of the putative human cytomegalovirus DNA polymerase gene. J. Virol. **61**:119–124.
20. **Holland, L. E., R. M. Sandri-Goldin, A. L. Goldin, J. C. Glorioso, and M. Levine.** 1984. Transcriptional and genetic analyses of the herpes simplex virus type 1 genome: coordinates 0.29 to 0.45. J. Virol. **49**:947–959.
21. **Howley, P. M., M. A. Israel, M.-F. Law, and M. A. Martin.** 1979. A rapid method for detecting and mapping homology between heterologous DNAs. J. Biol. Chem. **254**:4876–4883.
22. **Jeang, K.-T., M.-S. Cho, and G. S. Hayward.** 1984. Abundant constitutive expression of the immediate-early 94K protein from cytomegalovirus (Colburn) in a DNA-transfected mouse cell line. Mol. Cell. Biol. **4**:2214–2223.
23. **Jeang, K.-T., G. S. Hayward.** 1983. A cytomegalovirus DNA sequence containing tracts of tandemly repeated CA dinucleotides hybridizes to highly repetitive dispersed elements in mammalian cell genomes. Mol. Cell. Biol. **3**:1389–1402.
24. **Kemble, G. W., A. L. McCormick, L. Pereira, and E. S. Mocarski.** 1987. A cytomegalovirus protein with properties of herpes simplex virus ICP8: partial purification of the polypep-

tide and map position of the gene. J. Virol. **61**:3143–3151.
25. **Knipe, D. M., M. P. Quinlan, and A. E. Spang.** 1982. Characterization of two conformational forms of the major DNA-binding protein encoded by herpes simplex virus 1. J. Virol. **44**: 736–741.
26. **Knipe, D. M., and A. E. Spang.** 1982. Definition of a series of stages in the association of two herpesviral proteins with the cell nucleus. J. Virol. **43**:314–324.
27. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, K. Weston, P. Tomlinson, and B. G. Barrell.** 1987. Large scale rearrangement of homologous regions in the genomes of HCMV and EBV. Virology **157**:397–413.
28. **LaFemina, R. L., and G. S. Hayward.** 1980. Structural organization of the DNA molecules from human cytomegalovirus, p. 39–55. *In* B. N. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics, vol. 28. Academic Press, Inc., New York.
29. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56**:335–341.
30. **Laskey, R. A., and A. D. Mills.** 1977. Enhanced autoradiographic detection of ^{32}P and ^{125}I using intensifying screens and hypersensitized film. FEBS Lett. **82**:314–316.
31. **Lee, C. K., and D. M. Knipe.** 1983. Thermolabile in vivo DNA-binding activity associated with a protein encoded by mutants of herpes simplex virus type 1. J. Virol. **46**:909–919.
32. **Leinbach, S. S., J. F. Casto, and T. K. Pickett.** 1984. Deoxyribonucleoprotein complexes and DNA synthesis of herpes simplex virus type 1. Virology **137**:287–296.
33. **Littler, E., D. Purifoy, A. Minson, and K. L. Powell.** 1983. Herpes simplex virus non-structural proteins. III. Function of the major DNA-binding protein. J. Gen. Virol. **64**:983–995.
34. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Miller, J. S., B. M. Paterson, R. P. Ricciardi, L. Cohen, and B. E. Roberts.** 1983. Methods utilizing cell-free protein-synthesizing systems for the identification of recombinant DNA molecules. Methods Enzymol. **101**:650–674.
36. **O'Donnell, M. E., P. Elias, B. E. Funnell, and I. R. Lehman.** 1987. Interaction between the DNA polymerase and single-stranded DNA-binding protein (infected cell protein 8) of herpes simplex virus 1. J. Biol. Chem. **262**:4260–4266.
37. **O'Donnell, M. E., P. Elias, and I. R. Lehman.** 1987. Processive replication of single-stranded DNA templates by the herpes simplex virus-induced DNA polymerase. J. Biol. Chem. **262**: 4252–4259.
38. **Paterson, B. M., B. E. Roberts, and E. L. Kuff.** 1977. Structural gene identification and mapping by DNA · mRNA hybrid-arrested cell-free translation. Proc. Natl. Acad. Sci. USA **74**: 4370–4374.
39. **Powell, K. L., E. Littler, and D. J. M. Purifoy.** 1981. Nonstructural proteins of herpes simplex virus. II. Major virus-specific DNA-binding protein. J. Virol. **39**:894–902.
40. **Purifoy, D. J. M., and K. L. Powell.** 1976. DNA-binding proteins induced by herpes simplex virus type 2 in HEp-2 cells. J. Virol. **19**:717–731.
41. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. Nucleic Acids Res. **12**:643–655.
42. **Quinlan, M. P., L. B. Chen, and D. M. Knipe.** 1984. The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. Cell **36**:857–868.
43. **Quinn, J. P., and D. J. McGeoch.** 1985. DNA sequence of the region in the genome of herpes simplex virus type 1 containing the genes for DNA polymerase and the major DNA-binding protein. Nucleic Acids Res. **13**:8143–8163.
44. **Rafield, L. F., and D. M. Knipe.** 1984. Characterization of the major mRNAs transcribed from the genes for glycoprotein B and DNA-binding protein ICP8 of herpes simplex virus type 1. J. Virol. **49**:960–969.
45. **Rapp, F.** 1983. The biology of cytomegaloviruses, p. 1–53. *In* B. Roizman (ed.), The herpesvirus, vol. 2. Plenum Publishing Corp., New York.
46. **Rigby, P. W. J., M. Dieckman, G. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**: 237–251.
47. **Ruyechan, W. T., and A. C. Weir.** 1984. Interaction with nucleic acids and stimulation of the viral DNA polymerase by the herpes simplex virus type 1 major DNA-binding protein. J. Virol. **52**:727–733.
48. **Sanger, R., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
49. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
50. **Wahren, B., and B. Oberg.** 1979. Inhibition of cytomegalovirus late antigens by phosphonoformate. Intervirology **12**:335–339.