JOURNAL OF VIROLOGY, June 1997, p. 4209–4217
0022-538X/97/$04.00+0
Copyright © 1997, American Society for Microbiology
Vol. 71, No. 6

Circularization and Cleavage of Guinea Pig Cytomegalovirus Genomes

MICHAEL A. McVOY,* DANIEL E. NIXON, AND STUART P. ADLER

Department of Pediatrics, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298-0163

Received 20 November 1996/Accepted 5 February 1997

The mechanisms by which herpesvirus genome ends are fused to form circles after infection and are re-formed by cleavage from concatemeric DNA are unknown. We used the simple structure of guinea pig cytomegalovirus genomes, which have either one repeated DNA sequence at each end or one repeat at one end and no repeat at the other, to study these mechanisms. In circular DNA, two restriction fragments contained fused terminal sequences and had sizes consistent with the presence of single or double terminal repeats. This result implies a simple ligation of genomic ends and shows that circularization does not occur by annealing of single-stranded terminal repeats formed by exonuclease digestion. Cleavage to form the two genome types occurred at two sites, and homologies between these sites identified two potential *cis* elements that may be necessary for cleavage. One element coincided with the A-rich region of a *pac2* sequence and had 9 of 11 bases identical between the two sites. The second element had six bases identical at both sites, in each case 7 bp from the termini. To confirm the presence of *cis* cleavage elements, a recombinant virus in which foreign sequences displaced the 6- and 11-bp elements 1 kb from the cleavage point was constructed. Cleavage at the disrupted site did not occur. In a second recombinant virus, restoration of 64 bases containing the 6- and 11-bp elements to the disrupted cleavage site restored cleavage. Therefore, *cis* cleavage elements exist within this 64-base region, and sequence conservation suggests that they are the 6- and 11-bp elements.

Herpesviruses contain large (130- to 230-kb) double-stranded linear genomes that circularize shortly after infection (8, 15). Subsequent DNA synthesis produces large concatemers of head-to-tail-linked genomes that are inserted into viral capsids and cleaved at specific sequences to form unit-length linear genomes (21).

The mechanisms of concatemer cleavage and genome circularization are not known. The ends of herpes simplex virus (HSV) genomes have complementary 3′ single-base overhangs that suggest that cleavage occurs by a simple staggered cut (18). For several other herpesviruses, 3′ single-base terminal overhangs are implied by the loss of a single base during cloning of terminal fragments (3, 10, 14, 23, 28, 33), and the sequences of terminal fusions in concatemeric DNA are consistent with circularization by annealing and ligation of these complementary single-base overhangs (3, 10, 14, 18, 23, 28).

Many herpesvirus genomes, however, have directly repeated sequences at each end. In vivo, genomic ends can become single stranded after infection (12). In vitro, if made single stranded by addition of exonuclease, terminal repeats can anneal to form circles (9, 13, 34), suggesting that genome circularization occurs by annealing of complementary single-stranded ends formed by exonucleolytic digestion of the terminal repeats (2). A third mechanism, proposed for circularization of Epstein-Barr virus (EBV) genomes, involves recombination between repeated terminal sequences (35). The latter two mechanisms require direct terminal repeats which are not present in some herpesvirus genomes (3, 10, 11, 23). Thus, the mechanism of genome circularization remains unresolved.

Circularization brings together terminal sequences, creating junctions which later act as cleavage sites between genomes in concatemeric DNA. These sites contain *cis* DNA elements that are required for cleavage, but these elements are not accurately defined. Sequence comparisons identified two DNA elements, *pac1* and *pac2*, that are strongly conserved at cleavage sites from different herpesviruses (4, 27). A *pac1* element consists of a 5- to 9-bp A/T-rich region bilaterally flanked by 5- to 10-bp C runs 30 to 35 bases from the terminus. A *pac2* element consists of a conserved CGCGGCG motif 25 to 30 bases from a 5- to 10-bp A-rich region, again 30 to 35 bases from the terminus (4). Both *pac1* and *pac2* are present in the HSV terminal repeat, or *a* sequence (4), which is able to mediate cleavage and packaging when present in HSV amplicons (concatemeric plasmid DNA replicated by helper virus superinfection) (1, 4, 6, 25, 26, 31) or when inserted at an ectopic site in the HSV genome (16, 17, 18, 24, 31). Evidence from deletional mutagenesis of the *a* sequence using both of these systems supports the hypothesis that *pac1* and *pac2* mediate cleavage (5, 20, 24, 31); however, *pac1* and *pac2* have not as yet been shown to be *cis* cleavage elements by site-specific mutagenesis.

Guinea pig cytomegalovirus (GPCMV) genomes have one copy of a direct repeat at the left R terminus and one copy at the right M terminus. In approximately equimolar amounts, a second genome structure which lacks the repeat at the right O terminus (1A) occurs (7). For studying circularization and cleavage, this genome structure provides three advantages. First, because the GPCMV genome lacks the internal repeats and multiply reiterated terminal repeats, novel restriction fragments formed by circularization of GPCMV termini can easily be detected. Second, the two genomes of GPCMV are formed by cleavage at two different cleavage sites, and sequence homology between these two sites may indicate potential *cis* cleavage elements. Third, the presence of two different cleavage sites permits mutagenesis of one site while propagating the virus using the other. These advantages led us to study genome circularization and cleavage of GPCMV.

* Corresponding author. Mailing address: Department of Pediatrics, Medical College of Virginia, Virginia Commonwealth University, P.O. Box 980163, Richmond, VA 23298-0163. Phone: (804) 828-0132. Fax: (804) 828-6455. E-mail: MMCVOY@GEMS.VCU.EDU.

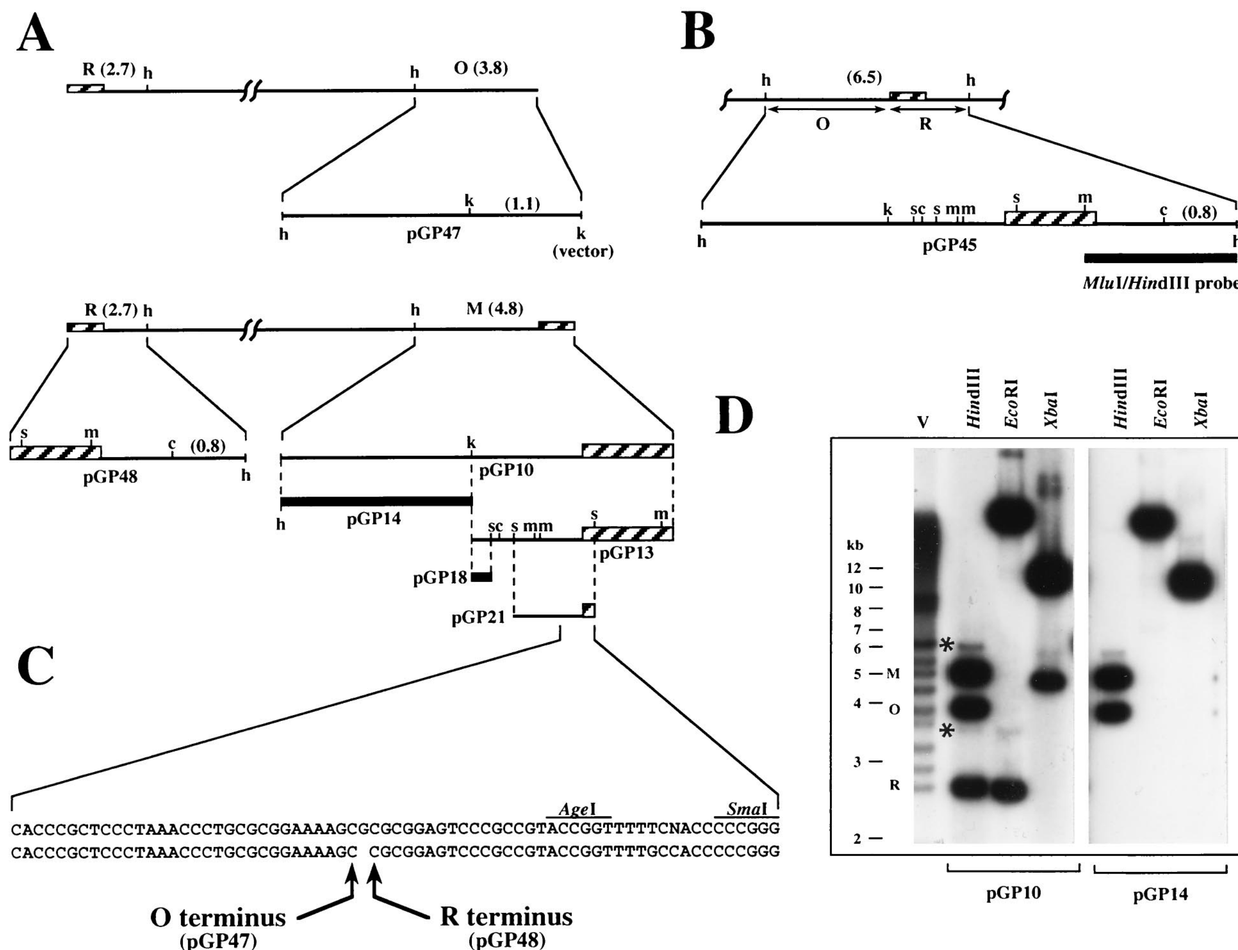

FIG. 1. GPCMV terminal sequence clones. (A) Cloned terminal fragments and subclones are shown expanded beneath representations of the two GPCMV genome types. Termini are referred to by their respective *Hin*dIII restriction fragments. The upper genome has one terminal repeat (hatched box) at the R terminus and lacks a repeat at the O terminus. The lower genome has one repeat at both R and M termini. Heavy lines represent sequences used as hybridization probes in other experiments. Relevant restriction sites are shown: h, *Hin*dIII; k, *Kpn*I; c, *Cla*I; m, *Mlu*I; and s, *Sma*I. Fragment sizes are given in parentheses in kb. (B) A *Hin*dIII fragment containing a single repeat cloned from GPCMV concatemeric DNA is illustrated. The heavy line represents an *Mlu*I/*Hin*dIII fragment used as a hybridization probe to screen for R-fragment clones. (C) Sequence from pGP21 was aligned with terminal sequences from pGP47 (O) and pGP48 (R). (D) Hybridizations were performed to confirm the sequences cloned in pGP10 and pGP14. GPCMV virion DNA was digested with *Hin*dIII, *Eco*RI, or *Xba*I and then electrophoretically separated, transferred, and hybridized with pGP10 or pGP14 as indicated. Lane V contains *Hin*dIII-digested GPCMV virion DNA ^{32}P labeled in situ. The positions of DNA size markers and GPCMV *Hin*dIII M, O, and R fragments are indicated. Faint *Hin*dIII fragments labeled with asterisks may be derived from rare M and R termini with two repeats (see text).

MATERIALS AND METHODS

Viral culture and drug treatments. GPCMV strain 22122 (American Type Culture Collection) was propagated by using 104C1 cells, an immortal guinea pig cell line (American Type Culture Collection), or guinea pig embryo fibroblasts (GEFs) prepared as described by Tenser and Hsiung (29). All cells and viruses were propagated by using Eagle's minimal essential medium supplemented with 10% fetal calf serum and 50 U of penicillin (MA Bioproducts) per ml. In some experiments, the culture medium contained 50 μg of cycloheximide (Sigma) per ml.

Preparation of viral DNAs. Virion DNA was prepared from cell culture medium after removal of cell debris by low-speed centrifugation (800 × *g* for 5 min). Viral particles were pelleted from the supernatants by ultracentrifugation at 20,000 × *g* for 30 min. The viral pellet was resuspended in TE buffer (10 mM Tris, 1 mM EDTA [pH 8.0]), mixed with an equal volume of lysis buffer (300 mM NaCl, 1% sodium dodecyl sulfate, 30 μg of proteinase K per ml), and incubated overnight at 65°C. The DNA was extracted once with an equal volume of phenol-chloroform and once with chloroform and precipitated by addition of 2 volumes of ethanol. After incubation at −20°C for 1 h, precipitated DNA was pelleted by centrifugation at 13,000 × *g* for 10 min. The DNA pellet was washed with 70% ethanol, centrifuged again for 5 min at 13,000 × *g*, air dried, and finally dissolved in TE.

To prepare infected cell DNA, infected cells were removed from culture dishes by trypsin digestion, pelleted by low-speed centrifugation, resuspended in TE, mixed with lysis buffer, and extracted as described for virion DNA.

Cloning of GPCMV terminal sequences. Plasmid pGP10, which contains the GPCMV M terminal *Hin*dIII fragment, was made by ligation of *Hin*dIII-digested GPCMV virion DNA fragments into *Hin*dIII-digested pUC18 DNA. Two fragments were subcloned from pGP10 to make plasmids pGP13 and pGP14, using a central *Kpn*I site (Fig. 1A). Hybridization studies confirmed that pGP10 contains the M fragment and that pGP14 contains the terminal repeat (see Results). Two *Sma*I fragments were subcloned from pGP13 to make pGP18 and pGP21. The order and orientation of these fragments (Fig. 1A) were determined by sequencing and mapping.

Plasmids pGP47 (containing *Hin*dIII O) and pGP48 (containing *Hin*dIII R) were isolated by library screening. The library was constructed by blunt ending GPCMV virion DNA with T4 DNA polymerase and then digesting with *Hin*dIII. The resulting fragments were ligated into *Hin*cII/*Hin*dIII-digested pGEM3Zf+ vector DNA. The ligation mixture was electroporated into *Escherichia coli* Top10 (Invitrogen) and spread on ampicillin plates. Colonies were transferred to Protran nitrocellulose disks (Schleicher & Schuell) and screened by colony lift hybridization with ^{32}P-labeled probes according to the manufacturer's instructions.

The library was screened for O-fragment-containing clones by using a 200-bp *Sma*I fragment gel purified from pGP18 and containing unique sequences adjacent to the repeat (Fig. 1A). Hybridization-positive clones were further screened by *Kpn*I and *Sma*I digestion. pGP47 contained a 1.1-kb *Kpn*I fragment predicted for an O-fragment insertion (Fig. 1A) and two *Sma*I fragments identical in size to *Sma*I fragments derived from pGP10 (data not shown).

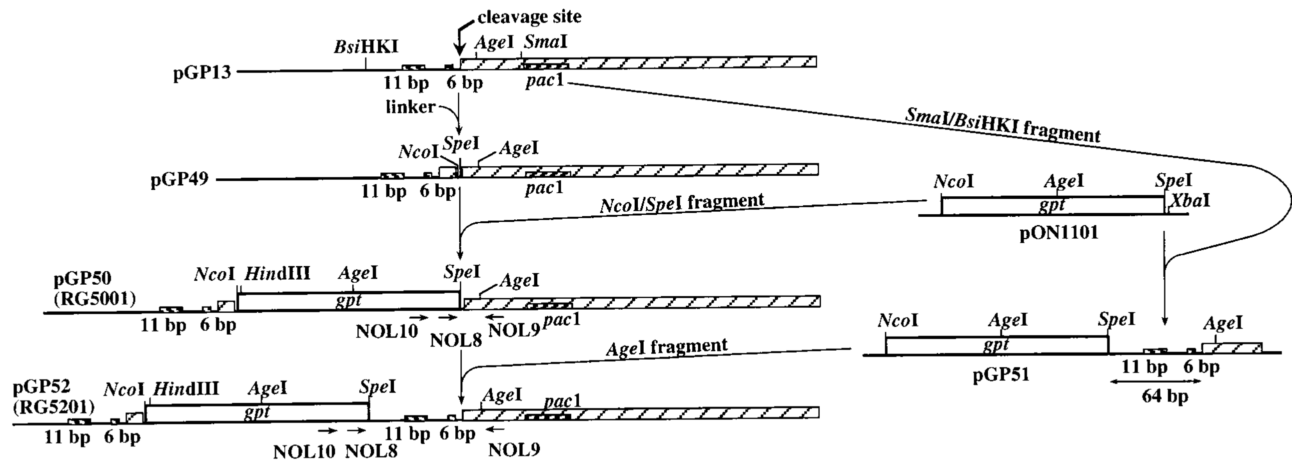

FIG. 2. Construction of recombinant plasmids and viruses. To make pGP50, a synthetic linker inserted at the *Age*I site of pGP13 reconstructed 18 bp of the terminal repeat (large hatched box) and introduced *Nco*I and *Spe*I sites, which were used to insert a *gpt* expression cassette (open box) where O and R sequences meet. To make pGP52, a 100-bp fragment from pGP13 was inserted adjacent to *gpt* and then exchanged for an *Age*I fragment in pGP50. See Materials and Methods for details. Homologous recombination and selection for *gpt* were used to make recombinant virus RG5001 from pGP50 and recombinant virus RG5201 from pGP52. These viruses therefore contain the same sequence arrangements as the plasmids from which they were derived. Putative *cis* elements identified by sequencing are indicated by small hatched boxes. The positions of primers used for PCR (NOL9 and NOL10) or for sequencing (NOL8) are shown).

To obtain a hybridization probe specific for the R fragment, we cloned a *Hin*dIII fragment containing a fusion of the O and R termini from concatemeric DNA (Fig. 1B). *Hin*dIII fragments in the 6- to 7-kb range were agarose gel purified from infected cell DNA and ligated to *Hin*dIII-digested pGEM3Zf+ vector DNA. Plasmid pGP45 had a 6.5-kb *Hin*dIII insert which contained two *Sma*I and four *Mlu*I fragments identical in size to *Sma*I or *Mlu*I fragments in pGP10 and contained M-terminal repeat sequences as determined by sequencing (data not shown). A 1.6-kb *Mlu*I/*Hin*dIII fragment from pGP45 containing sequences mostly unique to the R fragment (Fig. 1B) was gel purified and used to screen the library for R-fragment-containing clones. Hybridization-positive clones were screened by *Hin*dIII/*Mlu*I and *Hin*dIII/*Cla*I digestion. Plasmid pGP48 contained both *Mlu*I and *Cla*I sites in predicted locations (Fig. 1A).

FIGE. Infected cells were washed with phosphate-buffered saline–EDTA, trypsinized, resuspended in 10 ml of phosphate-buffered saline–EDTA, pelleted, resuspended in 50 μl of molten (52°C) 1% SeaPlaque agarose (FMC) in TE, and cast into molds. After cooling, agarose plugs containing cells were suspended in SE (0.5 M EDTA, 1% Sarkosyl) with 1 mg of proteinase K per ml, incubated at 52°C for 48 h, dialyzed (three times, 2 h each) with TE, and stored at 4°C. DNA plugs were inserted into the loading wells of 11- by 14-cm 1% SeaKem agarose (FMC) gels and separated by field-inversion gel electrophoresis (FIGE) in 0.5× Tris-borate electrophoresis buffer for 36 h at 120 V and 14°C. The pulse times began at 5 s and increased to 60 s with a 3:1 forward-to-backward ratio.

Southern hybridization. FIGE-separated GPCMV DNAs were cut from FIGE gels and extracted from the agarose by using a Qiaex kit (Qiagen) prior to restriction enzyme digestion. Intact viral DNAs separated on FIGE gels or restriction-digested DNAs separated on 0.6% agarose gels were transferred to Nytran nylon membranes (Schleicher & Schuell) and hybridized as described previously (15). DNA probes were ^{32}P labeled by the random hexamer method as described previously (15).

In situ labeling of viral DNA. GPCMV virion DNA was metabolically labeled by addition of $^{32}P_i$ to low-phosphate culture medium as previously described (15).

Recombinant plasmid construction. To make recombinant viruses, plasmids pGP50 and pGP52 were first constructed as shown in Fig. 2. Sequencing and mapping revealed a unique *Age*I site within the terminal repeat 18 bp from the point at which O termini are cleaved (Fig. 1C). Complementary 5′-phosphorylated synthetic oligonucleotides were annealed to form a double-stranded oligonucleotide with *Age*I cohesive ends and then ligated to *Age*I-digested pGP13 DNA. Insertion of this oligonucleotide in the correct orientation reconstructed 18 bp of terminal repeat and 2 bp of the O terminus and then altered the viral sequence to create novel *Nco*I and *Spe*I sites. The resulting plasmid, pGP49, was then digested with *Nco*I and *Spe*I, and a 1.1-kb *Nco*I/*Spe*I fragment from plasmid pON1101 (kindly provided by Edward Mocarski) containing an expression cassette for the *E. coli* xanthine-guanine phosphoribosyltransferase (*gpt*) gene was inserted to create pGP50.

To construct pGP52, a 100-bp *Sma*I/*Bsi*HKAI fragment from pGP13 was made blunt ended with T4 DNA polymerase, polyacrylamide gel purified, ligated into pON1101 that had been digested with *Xba*I, and then blunt ended with Klenow. The resulting plasmid, pGP51, contains 64 bp of the O terminus, the cleavage site, and 36 bp of the R terminus adjacent to the *gpt* cassette. An *Age*I fragment from pGP51 was polyacrylamide gel purified and ligated to *Age*I-digested pGP50. In the resulting plasmid, pGP52, the *Age*I fragment in pGP50 was replaced with the *Age*I fragment from pGP51. The sequences of both plasmids at their modified cleavage sites were confirmed by sequencing from primer NOL8 (5′-ACTCATCAATGTATCTTATC) within the *gpt* gene (Fig. 2).

Recombinant virus construction. Two recombinant GPCMVs were constructed by the method of Vieira et al. (32), using *gpt* as a selectable marker. Thirty micrograms of plasmid DNA (pGP50 for virus RG5001 and pGP52 for virus RG5201) was linearized by *Mlu*I digestion, extracted once with phenol-chloroform and once with chloroform, ethanol precipitated, and resuspended in 20 μl of TE. GEFs were passaged 1:2 into two 225-cm^2 tissue culture flasks. Twenty-four hours later (just prior to electroporation), the cells were removed by trypsin digestion, centrifuged, washed once with Hanks' balanced salt solution, resuspended in 500 μl of Hanks' balanced salt solution, mixed with linearized plasmid DNA, transferred to a 0.4-cm cuvette (Invitrogen), and chilled on ice. Electroporation was performed with an Electroporator II apparatus (Invitrogen) set to 300 V, 1,000 μF, and infinite resistance. Cells were kept on ice in the cuvette for 10 min after electroporation, then diluted into 10 ml of culture medium, and placed in a 25-cm^2 flask overnight. They were then infected with GPCMV strain 22122 at a multiplicity of infection of 3. Five days postinfection (p.i.), progeny virus in the culture supernatant was used to infect fresh GEFs. Three hours p.i., the cells were washed and fresh medium containing 200 μM mycophenolic acid (Gibco/BRL), and 10 μM xanthine (Sigma) was added to select for *gpt*+ recombinants. Five days p.i., the culture supernatants were again used to infect fresh GEFs, and selective medium was added as before. When large foci were noted, culture supernatants were again used to infect fresh GEFs, and selective medium was added as before. Clonal viruses were isolated from supernatants of the third round of selection by limited-dilution cloning on 96-well plates containing confluent GEFs. Candidate clones were screened for the *gpt* insertion by PCR, and recombinant virus mutations were confirmed by sequencing of PCR products as described below.

PCR. PCR was used to screen viral isolates for insertion of *gpt* sequences by amplification between a primer within the *gpt* sequence (NOL10; 5′-GACTAT GTTGTTGATATC) and a second primer within the R terminal sequence (NOL9; 5′-CGAAACCGCTCGGAAGCG) (Fig. 2). Infected cells were pelleted, resuspended in culture medium, and boiled for 5 min. Five microliters of boiled cell extract was added to the PCR containing 1× PCR reaction buffer (Promega), 2 μM $MgCl_2$, 200 μM each dATP, dCTP, dTTP, and dGTP (Promega), and 1 μM oligonucleotides NOL9 and NOL10. The mixture was brought to 95°C before addition of 2 U of *taq* polymerase (Promega). The mixture was overlaid with 50 μl of mineral oil and cycled 35 times through the following temperatures: 95°C for 1 min; 55°C for 1 min; and 95°C for 2 min. PCR products were separated by electrophoresis on 0.7% agarose gels and visualized with ethidium bromide staining and UV light.

Sequencing. Plasmid clones of GPCMV DNA were sequenced by the Sanger dideoxy method by using either an ABI automated sequencer or an *fmol* or TaqTrack sequencing kit (Promega) with ^{32}P-labeled oligonucleotide primers. PCR products from recombinant viruses were extracted once with phenol-chloroform and once with chloroform and then washed three times with TE, using a Centricon 30 filter (Amicon). PCR products were then sequenced from primer NOL8 (5′-ACTCATCAATGTATCTTATCA [Fig. 2]) by using a TaqTrack kit.

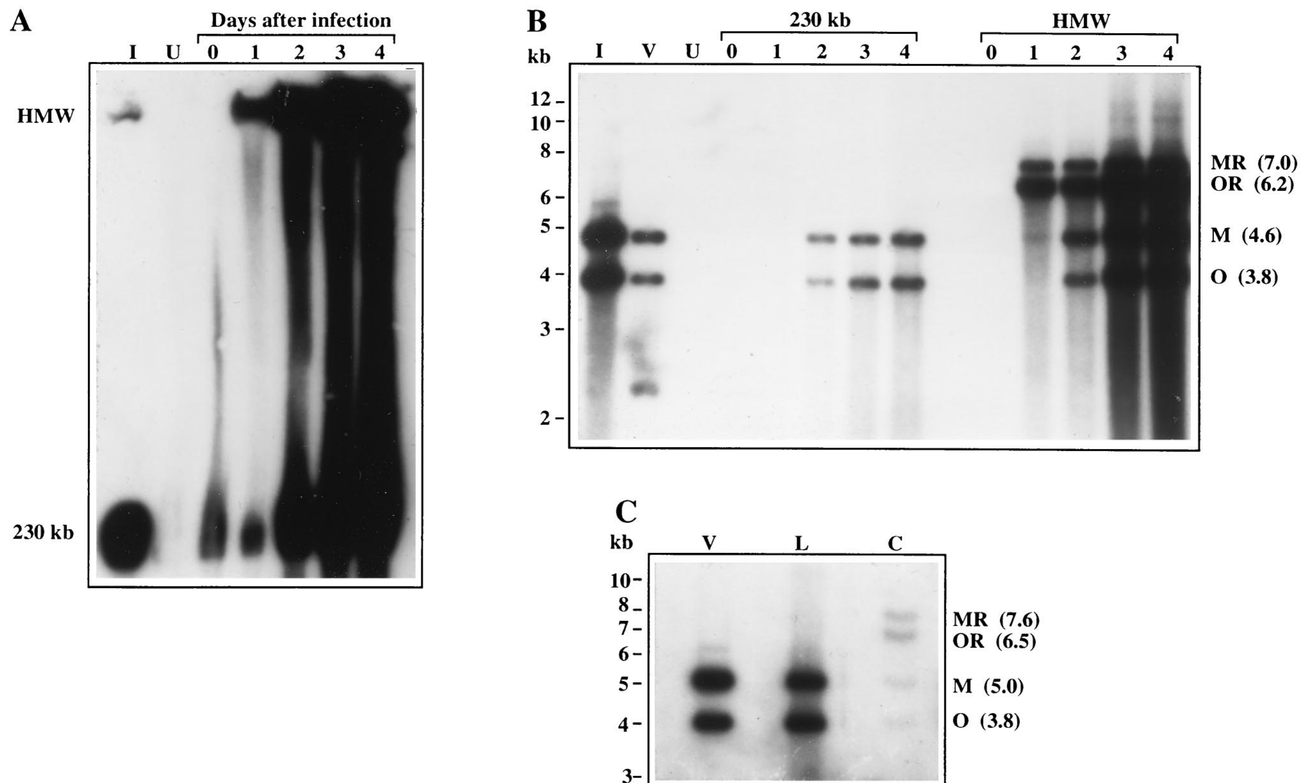

FIG. 3. Terminal repeats in replicative forms of GPCMV DNA. (A) Total GPCMV infected cell DNA was prepared 3 h p.i. (day 0) and daily thereafter as indicated, separated by FIGE, transferred, and hybridized with pGP14 DNA to detect intact GPCMV DNA forms. DNAs prepared from the inoculum (lane I) and from uninfected cells (lane U) were also included. (B) HMW and 230-kb DNAs were cut from a duplicate of the gel shown in panel A and digested with *Hin*dIII, separated electrophoretically, transferred, and hybridized with pGP14 to detect GPCMV terminal sequences. DNAs from the inoculum (lane I), the culture supernatant 4 days p.i. (lane V), and uninfected cells (lane U) were also included without prior FIGE separation. The positions of DNA size markers are shown to the left. Junction and terminal fragments and their measured sizes in kilobases (in parentheses) are indicated. (C) DNA was prepared 24 h p.i. from cells infected in the presence of cycloheximide to prevent concatemer synthesis and separated by FIGE. Circular HMW DNA (lane C) and 230-kb linear DNA (lane L) were cut from the gel and analyzed as for panel B. *Hin*dIII-digested DNA from extracellular virions (lane V) was also included without prior FIGE separation.

RESULTS

Cloning the GPCMV M terminus. To obtain a hybridization probe to detect GPCMV terminal sequences, random clones of GPCMV *Hin*dIII fragments were screened for insert size. Plasmid pGP10 contained a 4.8-kb insertion and only one *Hin*dIII site, both consistent with insertion of the M terminal fragment. To confirm that pGP10 contains the M fragment, GPCMV virion DNA was digested with *Hin*dIII, *Xba*I, and *Eco*RI, separated by agarose electrophoresis, transferred to a nylon membrane, and hybridized with ^{32}P-labeled pGP10 DNA. Three GPCMV *Hin*dIII fragments hybridized to pGP10 DNA, and their measured sizes (2.7, 3.8, and 4.8 kb) were consistent with sizes previously reported for terminal *Hin*dIII fragments R, O, and M (7), respectively (Fig. 1D). In addition, the measured sizes of the *Eco*RI and *Xba*I fragments which hybridized to pGP10 were also consistent with sizes reported for the GPCMV *Eco*RI and *Xba*I terminal fragments (7). Therefore, pGP10 contains the terminal M fragment.

Two fragments were subcloned from pGP10 to make plasmids pGP13 and pGP14, using a central *Kpn*I site (Fig. 1A). DNA from pGP13 hybridized to the same fragments as pGP10 (not shown) and therefore contains the terminal repeat. DNA from pGP14 hybridized to M and O, but not to R (Fig. 1D), confirming that pGP14 contains unique sequences common to M and O but lacks the terminal repeat (Fig. 1A).

Hybridization of GPCMV *Hin*dIII fragments with pGP10 also detected faint 3.7- and 5.8-kb fragments (Fig. 1D). These sizes are 1.0 kb (one repeat length) larger than the 2.7-kb R and 4.8-kb M terminal fragments and thus may represent rare termini with two terminal repeats.

Terminal repeats in GPCMV replicative intermediates. GPCMV circular genomes should contain single repeats if formed by recombination or by annealing of single-stranded terminal repeats. Circular genomes formed by end-to-end ligation of genomic ends should contain a mixture of single and double repeats. Therefore, we sought to resolve the mechanism of genome circularization by determining the number of terminal repeats in circular GPCMV DNA.

Using FIGE, we have shown that the predominant replicative intermediates of human cytomegalovirus (HCMV) migrate either as 230-kb DNA that contains unit-length linear viral genomes or as high-molecular-weight (HMW) DNAs, which fail to migrate from the sample plug and include circular genomes shortly after infection or concatemeric genomes late in infection (15). Similar findings have been reported for HSV DNAs (8). We used FIGE to separate DNAs from GPCMV-infected cells at various times p.i., transferred the DNAs to a membrane, and detected GPCMV DNAs by hybridization with pGP14 DNA (Fig. 3A). At 3 h p.i. (day 0), we detected only 230-kb DNA, but by day 1 p.i., a small amount of HMW DNA was formed. The amount of HMW DNA increased significantly by day 2, as did the amount of 230-kb DNA. Based on our studies with HCMV (15), the 230-kb DNA at 3 h represents linear inoculum genomes and the HMW DNA at day 1

represents inoculum genomes that have circularized. The dramatic increase in HMW DNA on day 2 represents synthesis of concatemeric DNA, and the increase in 230-kb DNA represents progeny genomes that have been cleaved from the concatemers.

We next sought to detect the restriction fragments that contain terminal sequences for each of these replicative forms. Duplicates of the DNA samples shown in Fig. 3A were separated on a duplicate FIGE gel. For each time point, 230-kb and HMW DNAs were excised from the gel, digested with *Hin*dIII, separated electrophoretically, transferred, and hybridized with pGP14 DNA to detect terminal sequences. For comparison, extracellular virion DNAs prepared from the inoculum, and the culture supernatants 4 days p.i. were included. Consistent with a linear structure, intracellular 230-kb DNAs were qualitatively indistinguishable from virion DNAs and contained terminal O and M fragments which measured 3.8 and 4.6 kb, respectively (Fig. 3B). Terminal fragments were nearly absent from HMW DNA on day 1, consistent with a circular structure, but were detected in HMW DNAs from days 2 to 4, consistent with initiation of concatemer synthesis and cleavage since cleavage leaves terminal fragments on concatemer ends (15). In addition, both circular and concatemeric HMW DNAs contained two larger fragments not present in virion or 230-kb DNAs (Fig. 3B). The measured sizes of these fragments, 7.0 and 6.2 kb, were consistent with molecular weights predicted for end-to-end ligation of M with R (4.6 + 2.7 = 7.3), producing MR junction fragments containing two repeats, and for ligation of O with R (3.8 + 2.7 = 6.5), producing OR junction fragments containing one repeat.

Since concatemer synthesis may have been initiated by day 1 in this experiment, HMW DNA from day 1 may have contained some concatemeric DNA. To analyze strictly circular GPCMV DNA, cells were infected in the presence of cycloheximide to block concatemer synthesis without blocking circularization (8, 15). DNA was prepared 22 h p.i. and separated by FIGE. HMW DNA, which in this case contains only circular genomes, and 230-kb linear DNAs were excised from the FIGE gel and analyzed by *Hin*dIII digestion and Southern hybridization (Fig. 3C).

As before, both extracellular virion and intracellular 230-kb linear DNAs contained terminal M and O fragments but lacked the larger MR and OR fragments. Circular DNA contained MR and OR fragments with trace amounts of terminal fragments that may have resulted from incomplete separation from linear genomes. In this experiment, the terminal M and O fragments measured 5.0 and 3.8 kb, and the junction fragments measured 7.6 and 6.5 kb, again consistent with molecular weights predicted for MR (5.0 + 2.7 = 7.7) and OR (3.8 + 2.7 = 6.5) if these fragments are formed by end-to-end ligation.

Sequences of the GPCMV cleavage sites. A staggered cut mechanism (18) predicts that cleavage to generate M termini must occur between double repeats, since R termini lacking a repeat are not observed (see below and reference 7). We refer to this as the M-R cleavage site (Fig. 4A). To generate O termini, cleavage of concatemeric DNA must occur where unique O terminal sequences end and repeat sequences begin. As this end of the repeat sequence is identical to the R terminus, we refer to this as the O-R cleavage site (Fig. 4B). To reconstruct the sequences at these sites, we cloned the GPCMV *Hin*dIII O and R terminal fragments.

As described above, plasmid pGP47, containing O, and pGP48, containing R, were isolated by screening a library of GPCMV virion DNA fragment clones with probes containing sequences unique to O or R and confirmed by restriction mapping. Final evidence that pGP47 contains O and pGP48 contains R came when sequences derived from ends of the pGP47 and pGP48 insertions exactly matched sequences from pGP21. The pGP48 sequence began one base pair after the pGP47 sequence ended (Fig. 1C). The missing base pair is consistent with T4 DNA polymerase removal of 3′ single-base overhangs at each terminus prior to cloning as has been reported for other herpesviruses (3, 10, 14, 23, 28, 33).

The terminal sequences from these clones were used to reconstruct the M-R and O-R cleavage sites. The M terminal side of the M-R cleavage site contained a typical *pac2* element with an A-rich region and a CGCGGCG motif. The R terminal side contained a typical *pac1* element with three A's flanked by G/C-rich sequences (Fig. 4A). The O-R cleavage site shared the *pac1* containing R terminal sequences on one side but on the other side contained unique sequences from the O terminus (Fig. 4B). Sequences at the O terminus analogous to the M terminal *pac2* were not immediately apparent; however, when O and M sequences were aligned, the O sequence contained 9 bases identical to the M sequence within an 11-bp region coinciding with the A rich region of *pac2* (Fig. 4C). Furthermore, the homologous bases were precisely the same distance from their respective termini. A second region seven bases from the termini contained six identical bases. No O terminal homolog of the *pac2* CGCGGCG motif was present.

Mutagenic analysis of the O-R cleavage site. To confirm that the O terminus contains *cis* sequences required for cleavage, two GPCMV recombinant viruses were constructed. In virus RG5001, a *gpt* cassette was inserted 2 bp from the point at which O and R sequences meet. With the exception of these 2 bp, all O terminal sequences, including the 6- and 11-bp homologous regions, were displaced by over 1 kb of foreign DNA (Fig. 2). In RG5201, 64 bp of O terminal sequence, including the 6- and 11-bp regions, were restored to the cleavage site adjacent to the *gpt* cassette (Fig. 2).

To assay for cleavage at the recombinant O-R cleavage sites, virion DNAs from RG5001 and RG5201 were digested with *Hin*dIII, separated electrophoretically, transferred, and hybridized with pON1101 DNA, which contains *gpt* sequences. Because the *gpt* insertion introduced a new *Hin*dIII site (Fig. 2), a 2.2-kb terminal fragment analogous to the M fragment and designated M* was predicted from cleavage at the M-R cleavage site, and a 1.2-kb fragment analogous to the O fragment and designated O* was predicted from cleavage at the recombinant O-R cleavage site.

In RG5201, both O* and M* fragments were observed, indicating that cleavage occurred at both the unaltered M-R site and the recombinant O-R site. In RG5001, however, the M* fragment was observed but the O* fragment was absent, indicating that cleavage did not occur at the recombinant O-R cleavage site which lacked the 64 bp of O terminal sequence (Fig. 5A). As these viruses differ only by the additional 64 bp in RG5201, we conclude that the 64 bases of the O terminus contain *cis* sequence elements required for cleavage.

The role of R terminal sequences in cleavage. At the right end of the repeat within R, M terminal sequences containing *pac2* join unique sequences (Fig. 4). The unique region adjacent to the M terminal *pac2* was sequenced and contained no sequences with homology to *pac1* (data not shown). If M terminal sequences alone can mediate cleavage, then a staggered cut at this site should produce R terminal fragments which lack a repeat. Initial characterization of the GPCMV genome did not detect these termini (7). We examined wild-type GPCMV and our two recombinant GPCMVs for the presence of R terminal fragments. Virion DNAs were digested with *Hin*dIII, separated, and transferred as described above and then hybridized with the *Mlu*I/*Hin*dIII fragment of pGP45 (Fig. 1B) to specifically detect unique sequences proximal to the repeat in the R

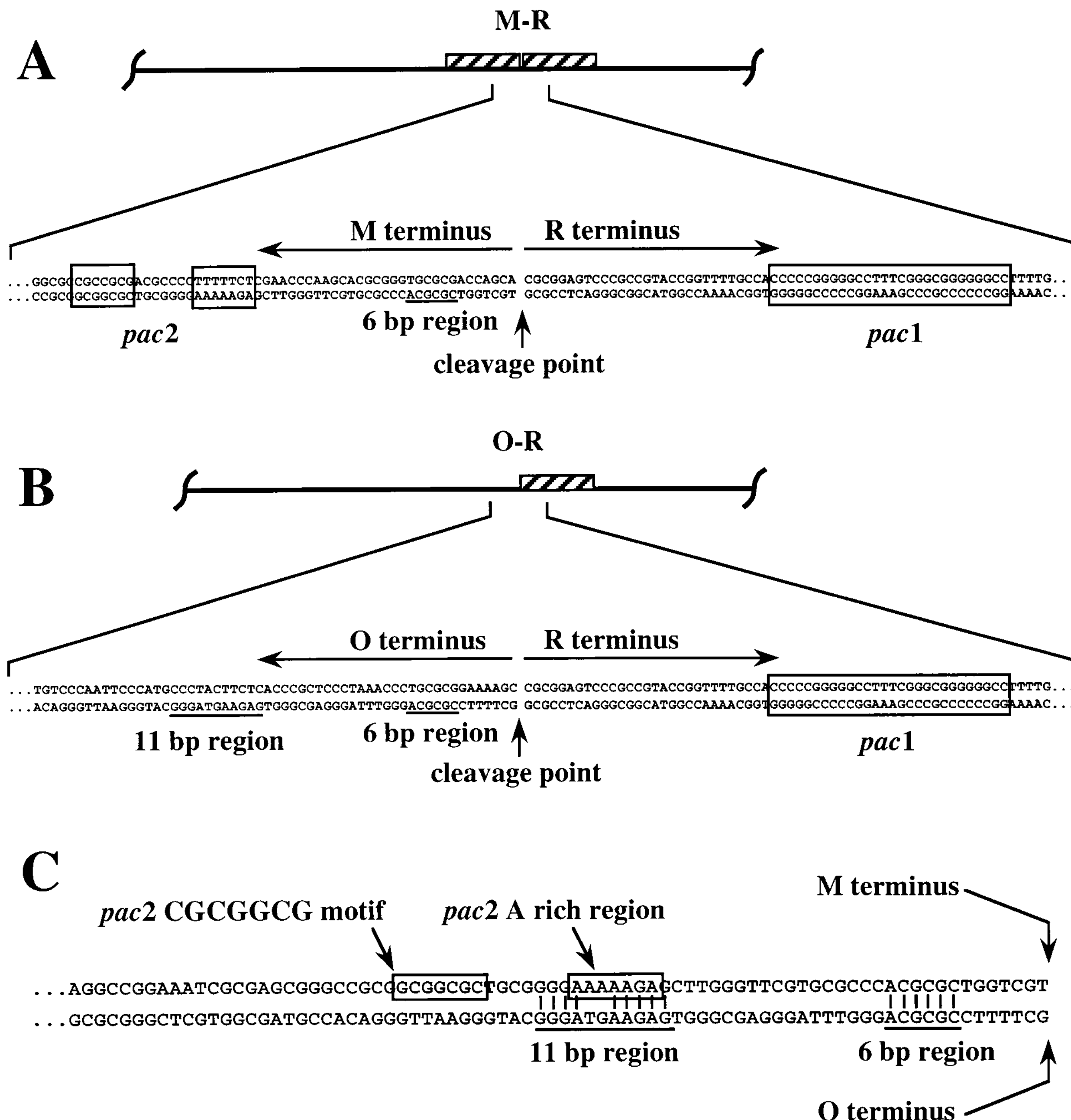

FIG. 4. Conserved sequences at GPCMV cleavage sites. (A) An M-R cleavage site consisting of double repeats (hatched boxes) is shown within a GPCMV concatemer. R and M terminal sequences are aligned beneath to reconstruct the sequences flanking the M-R site. A gap between the two sequences represents a base pair presumably lost during cloning. The *pac1* and *pac2* elements are boxed. (B) An O-R cleavage site consisting of a single repeat is shown with the R and O terminal sequences aligned beneath. The 6- and 11-bp regions of homology are underlined. (C) Comparison of M and O terminal sequences aligned from the point of cleavage. Identical bases are connected by vertical lines. The sequences shown are in 3′-to-5′ orientation.

terminal fragment. Although 2.7-kb R fragments were observed in all three viruses, 1.7-kb fragments lacking the terminal repeat were not detected (Fig. 5B), indicating that staggered cuts did not occur at this site. Therefore, sequences within the R terminus are essential for cleavage by a staggered cut mechanism.

DISCUSSION

The genomic structure of GPCMV provided three unique opportunities. First, the presence of no more than one repeat at each end enabled us to resolve the mechanism of circularization by determining the number of repeats in circular GPCMV DNA. We found that two novel junction fragments are formed following circularization of GPCMV genomes. Their measured molecular weights match the predicted values for junctions containing single and double repeats. Because both the exonuclease and recombination mechanisms predict a loss of one terminal repeat, neither can account for the formation of double-repeat-containing junctions unless a significant proportion of viral genomes contain two repeats at one end and one repeat at the other. Although virion DNA did reveal evidence

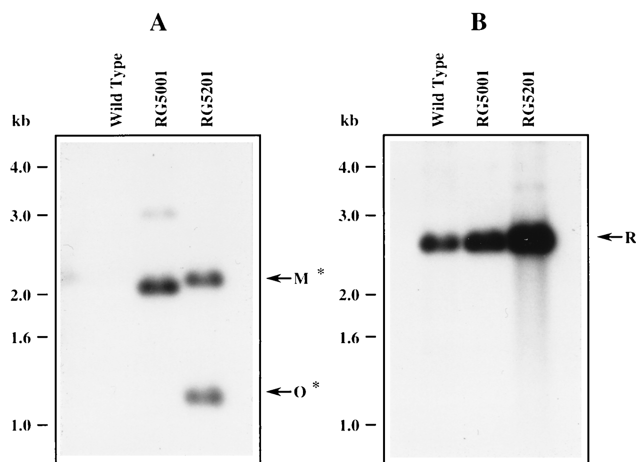

FIG. 5. Cleavage at GPCMV cleavage sites. (A) Virion DNAs from recombinant viruses RG5001 and RG5201 were digested with *Hin*dIII, electrophoretically separated, transferred, and hybridized with pON1101 DNA to detect terminal fragments containing *gpt* sequences. The M* fragment is slightly larger in RG52001 due to the additional 64 bp of sequence. (B) The membrane shown in panel A was rehybridized with an *Mlu*I/*Hin*dIII fragment from pGP45 (Fig. 1B) to detect R terminal fragments. The positions of molecular weight markers are indicated.

for a small fraction of genomes with two repeats at either terminus (Fig. 1D), the quantity of these genomes is insufficient to account for the large proportion of double-repeat-containing junctions in circular genomes.

The presence of single and double repeats in GPCMV circular DNA is consistent with circularization by an end-to-end ligation mechanism, since double repeats would result from circularization of genomes with one repeat at M and R termini and single repeats would result from circularization of genomes with one repeat at R termini and no repeat at O termini. Like GPCMV, a significant proportion of HCMV and murine cytomegalovirus (MCMV) genomes have been reported to lack terminal repeats (14, 28), but the ability of these genomes to circularize and enter replication has not been previously established.

The sequences of R and O termini also support a ligation mechanism. No repeated regions exist between R and O termini, and with the exception of one missing base, the sequence predicted from ligation of R and O termini matches the sequence where R and O sequences meet within the M fragment. Analogous results were observed when comparing terminal sequences with concatemeric junctions from HSV (18), HCMV (28), MCMV (14), varicella-zoster virus (3), equine herpesvirus 1 (23), bovine herpesvirus 1 (10), and pseudorabies virus (PRV) (11). Direct sequencing of HSV virion DNA revealed a 3′ single-base overhang at each terminus (18), and evidence based on the absence of single bases from terminal fragment clones (presumably removed by DNA polymerase treatment prior to cloning) suggests 3′ single-base terminal overhangs for GPCMV (this study), MCMV (14), rat cytomegalovirus (RCMV) (33), HCMV (28), varicella-zoster virus (3), equine herpesvirus 1 (23), and bovine herpesvirus 1 (10). In conjunction with our findings for GPCMV, these observations indicate that the majority of herpesviruses circularize by annealing and ligation of 3′ single-base overhangs.

There are, however, two notable exceptions. In PRV, one end is blunt and the other has a 2-base 3′ overhang, yet the sequence of concatemeric junctions is consistent with ligation of these termini and repair of the resulting 2-bp single-strand gap (11). In EBV, one end is blunt and the other has a single-base 3′ overhang. The sequence of circular episomal EBV, however, cannot be explained by ligation of virion termini and may involve recombination (35).

The sequence comparisons described above also suggest that cleavage for most herpesviruses occurs by a single-base staggered cut (3, 10, 14, 18, 23, 28). Although strongly supportive of this mechanism, certain aspects of our data suggest additional levels of complexity. For GPCMV, the staggered cut mechanism implies that M termini arise from cleavage between double repeats and that O termini arise from cleavage adjacent to single repeats. Since M and O terminal fragments are equally prevalent in virion DNA, MR and OR fragments should be equally prevalent in concatemeric DNA. Examination of HMW DNA in Fig. 3B, however, reveals that OR fragments are three- to fourfold more prevalent than MR fragments. We also found that genomes having O termini produce circular molecules which contain one repeat, yet concatemeric DNA from stocks derived from single viral plaques contains both single and double repeats (data not shown), suggesting that at some point the terminal repeat undergoes

```
GPCMV(M)  tgctggtcgcgcacccgcgtgcttgggttcg  AGAAAAA  ggggcgt  CGCGGCG...
GPCMV(O)  gcttttccgcgcagggtttagggagcgggtg  AGAAGTA  gggcatgggaattgggacaccgtagc...
MCMV      gggcccgcgcgcactcagacggccgggggg  ATAAAAA  gtggtactgagctaggtcccctatg aaaaaaagaagtatctgc  CGCGGCG..
RCMV      ccaacggcgcgcacccgggaggagaggggg  AAAAAA  gcgcaagcgggatgagcgggatgagcgggag...
HCMV   ...gttgcgtgtggtgcggaggacggcgacggcg  AATAAAA  gcgacgtgcggcgcgcacggcgaaaagaaga...
HHV-6     cacgcgcctcttgcagtggcgcgcgtggtggtg  AAAAACA  cggcgcctcccatagaggcgcgcgcgcgcg
HHV-7     tgtttttcactgagtgctgcgcgcggctcgc  AGAAAAA  gtgccttgaaggcattacaagacattggctg
HSV-1     gcagcccgggcccccgcgggcggggcg gcgcgc  AAAAAA  ggcgggcggcggtccgggcggcgtgcgcgcg  CGCGGCG...
EBV       ggggcatgggggccgcgcattcctgg  AAAAA  gtggagggggcgtggctggcccc  CGCGGC c...
```

FIG. 6. Conservation of the 6-bp region at herpesvirus termini. The GPCMV O and M terminal sequences are aligned with terminal *pac2* containing sequences from MCMV (14), RCMV (33), HCMV (19), HHV-6 (30), HHV-7 (22), HSV (type 1) (18), and EBV (35). The left ends correspond to genomic termini, except for HCMV, which has an additional 94 bp between the left end of the sequence and the terminus. Sequences from the GPCMV 6-bp region are underlined. The 8-bp sequence conserved between RCMV, MCMV, and HCMV is boxed. A 7-bp sequence conserved between HHV-6 and HHV-7 (dashed box) contains 5 bp from the GPCMV 6-bp region (underlined). The *pac2* A-rich regions and CGCGGCG motifs (if present) are offset by spaces and capitalized. Terminal 3′ single-base overhangs are not included.

limited duplication. One explanation proposed for HSV *a* sequence amplification is that terminal repeats are acquired from neighboring genomes in the concatemer (31). This mechanism is unlikely for GPCMV, as it would result in genomes lacking R terminal repeats, and none were detected by hybridization. A second explanation is that the *a* sequence is duplicated during the cleavage process (31). This clearly is not the case for all cleavage events in GPCMV, since O termini have no repeat; however, it is possible that some portion of O-R sites are cleaved by a mechanism that duplicates the repeat to produce M termini. This could also explain how equal quantities of M and O termini arise from disproportionate quantities of M-R and O-R sites, since some M termini would be derived from O-R sites.

Second, the existence of two cleavage sites within one virus allowed identification of two regions that are conserved between the two sites. The 11-bp region coincides with the A-rich region of *pac2*, and the 6-bp region occurs 7 bases from the M and O termini. The 6-bp region does not coincide with either *pac1* or *pac2* but is found at *pac2*-containing termini of several herpesviruses (Fig. 6). Because the 6-bp region is small and G/C rich, it may occur randomly in herpesvirus terminal regions; however, its specific association with *pac2*-containing termini and its conserved spacing relative to the termini, particularly within the rodent cytomegaloviruses (Fig. 6), suggest that the placement of the 6-bp region is not random.

The *pac2* CGCGGCG motif was not conserved between the two GPCMV cleavage sites, and its absence from one site clearly indicates that it is not essential for cleavage. This is consistent with the fact that CGCGGCG motifs do not exist near the *pac2* A-rich regions of RCMV (33), human herpesvirus 6 (HHV-6) (30), or HHV-7 (22). However, its presence at the other GPCMV site implies that it does serve a purpose, possibly to increase cleavage efficiency at this site. Because a double repeat is essentially an O-R site adjacent to a second repeat (Fig. 4), cleavage might be expected adjacent to double repeats, producing R termini with two repeats. Hybridization, however, detected only a very faint 3.7-kb fragment resulting from such termini (Fig. 1D), indicating that cleavage preferentially occurs between double repeats. The presence of the CGCGGCG motif may be responsible for this preferential cleavage.

Third, the presence of two cleavage sites enabled us to mutagenize one cleavage site in a recombinant virus while maintaining cleavage at the second site. This confirmed the existence of *cis* cleavage elements within a 64-base sequence. In addition, our failure to detect R termini lacking a repeat indicates a need for *cis* cleavage elements within the *pac1*-containing R terminal region. This is consistent with data showing that regions of the HSV *a* sequence which contain *pac1* are necessary for serial passage of HSV amplicons (5, 20, 31) and for cleavage at ectopic sites inserted into the HSV genome (24, 31).

As the 6- and 11-bp regions of homology are within the 64-base region, they are strong candidates for *cis* cleavage elements. Additional data are needed to define the function of the 6-bp region. The 11-bp region, however, includes the *pac2* A-rich region, an element conserved at all herpesvirus cleavage sites that have been sequenced. Deletions that remove the HSV *pac2* from the *a* sequence block serial passage of HSV amplicons (5, 31) and are inefficiently cleaved when inserted at an ectopic site in the HSV genome (24, 31). Because these deletions remove sequences other than *pac2*, they do not specifically show that *pac2* or sequences within *pac2* are required for cleavage. Our data show that an 11-bp region coincident with the A-rich region of *pac2* is conserved between the two GPCMV cleavage sites and lies within a 64-bp region that contains sequences essential for cleavage. These results strongly imply a role for the *pac2* A-rich region in GPCMV cleavage and, because it is found at all herpesvirus cleavage sites, suggest that the *pac2* A-rich region is a *cis* cleavage element common to all herpesviruses.

ACKNOWLEDGMENTS

We thank Mark Prichard for technical advice regarding selection of recombinant viruses and Edward Mocarski for providing pON1101.

D.N. was supported by a fellowship from the National Foundation for Infectious Diseases and Astra Pharmaceuticals. This work was supported in part by the A. D. Williams Fund of the Medical College of Virginia, Virginia Commonwealth University.

REFERENCES

1. **Barnett, J. W., D. A. Epstein, and H. W. Chan.** 1983. Class I defective herpes simplex virus DNA as a molecular cloning vehicle in eucaryotic cells. J. Virol. **48:**384–395.
2. **Ben-Porat, T.** 1983. Replication of herpesvirus DNA, p. 81–106. *In* B. Roizman (ed.), The herpesviruses, vol. I. Plenum Press, New York, N.Y.
3. **Davison, A. J.** 1984. Structure of the genome termini of varicella-zoster virus. J. Gen. Virol. **65:**1969–1977.
4. **Deiss, L. P., and N. Frenkel.** 1986. Herpes simplex virus amplicon: cleavage of concatemeric DNA is linked to packaging and involves amplification of the terminally reiterated *a* sequence. J. Virol. **57:**933–941.
5. **Deiss, L. P., J. Chou, and N. Frenkel.** 1986. Functional domains within the *a* sequence involved in the cleavage-packaging of herpes simplex virus DNA. J. Virol. **59:**605–618.
6. **Frenkel, N.** 1981. Defective interfering herpesviruses, p. 91–120. *In* A. H. Hahmias, W. R. Dowdle, and R. S. Schinazy (ed.), The human herpesviruses—an interdisciplinary perspective. Elsevier Science Publishing, Inc., New York, N.Y.
7. **Gao, M., and H. C. Isom.** 1984. Characterization of the guinea pig cytomegalovirus genome by molecular cloning and physical mapping. J. Virol. **52:** 436–447.
8. **Garber, D., S. Beverly, and D. Coen.** 1993. Demonstration of circularization of herpes simplex virus DNA following infection using pulsed field gel electrophoresis. Virology **197:**459–462.
9. **Grafstrom, R. H., J. C. Alwine, W. L. Steinhart, and C. W. Hill.** 1975. Terminal repetition of herpes simplex virus DNA. Virology **67:**144–157.
10. **Hammerschmidt, W., H. Ludwig, and H. Buhk.** 1988. Specificity of cleavage in replicative-form DNA of bovine herpesvirus 1. J. Virol. **62:**1355–1363.
11. **Harper, L., J. Demarchi, and T. Ben-Porat.** 1986. Sequence of the genome ends and the junction between the ends in concatemeric DNA of pseudorabies virus. J. Virol. **60:**1183–1185.
12. **Jean, J., and T. Ben-Porat.** 1976. Appearance *in vivo* of single-stranded complementary ends on parental herpesvirus DNA. Proc. Natl. Acad. Sci. USA **73:**2674–2678.
13. **Kudler, L., and R. W. Hyman.** 1979. Exonuclease III digestion of herpes simplex virus DNA. Virology **92:**68–91.
14. **Marks, J. R., and D. H. Spector.** 1988. Replication of the murine cytomegalovirus genome: structure and role of the termini in generation and cleavage of concatenates. Virology **162:**98–107.
15. **McVoy, M., and S. P. Adler.** 1994. Human Cytomegalovirus DNA replicates after early circularization and inversion occurs within the concatemer. J. Virol. **68:**1040–1051.
16. **Mocarski, E. S., L. E. Post, and B. Roizman.** 1980. Molecular engineering of the herpes simplex virus genome: insertion of a second L-S junction into the genome causes additional genome inversions. Cell **22:**243–255.
17. **Mocarski, E. S., and B. Roizman.** 1981. Site-specific inversion sequence of the herpes simplex virus genome: domain and structural features. Proc. Natl. Acad. Sci. USA **78:**7047–7051.
18. **Mocarski, E. S., and B. Roizman.** 1982. Structure and role of the herpes simplex virus DNA termini in inversion, circularization, and generation of virion DNA. Cell **31:**89–97.
19. **Mocarski, E. S., A. C. Liu, and R. R. Spaete.** 1987. Structure and variability of the *a* sequence in the genome of human cytomegalovirus (Towne strain). J. Gen. Virol. **68:**2223–2230.
20. **Nasseri, M., and E. S. Mocarski.** 1988. The cleavage recognition signal is contained within sequences surrounding an *a-a* junction in herpes simplex virus DNA. Virology **167:**25–30.
21. **Roizman, B., and A. E. Sears.** 1996. Herpes simplex viruses and their replication, p. 1048–1066. *In* B. N. Fields (ed.), Fundamental virology. Raven Press, New York, N.Y.
22. **Secchiero, P., J. Nicholas, H. Deng, T. Xiaopeng, N. VanLoon, V. Ruvolo, Z. N. Berneman, M. S. Reitz, and S. Dewhurst.** 1995. Identification of human telomeric repeat motifs at the genome termini of human herpesvirus 7:

structural analysis and heterogeneity. J. Virol. **69:**8041–8045.

23. **Shafiqul, I. C., H. Buhk, H. Ludwig, and W. Hammerschmidt.** 1990. Genomic termini of equine herpesvirus 1. J. Virol. **64:**873–880.
24. **Smiley, J. R., J. Duncan, and M. Howes.** 1990. Sequence requirements for DNA rearrangements induced by the terminal repeat of herpes simplex virus 1 KOS DNA. J. Virol. **64:**5036–5050.
25. **Spaete, R. R., and N. Frenkel.** 1982. The herpes simplex virus amplicon: a new eucaryotic defective-virus cloning-amplifying vector. Cell **30:**295–304.
26. **Stow, N., E. McMonagle, and A. Davison.** 1983. Fragments from both termini of the herpes simplex virus type 1 genome contain signals required for the encapsidation of viral DNA. Nucleic Acids Res. **11:**8205–8820.
27. **Tamashiro, J. C., D. Filpula, R. Friedmann, and D. H. Spector.** 1984. Structure of the heterogeneous L-S junction region of human cytomegalovirus strain AD169 DNA. J. Virol. **52:**541–548.
28. **Tamashiro, J. C., and D. H. Spector.** 1986. Terminal structure and heterogeneity in human cytomegalovirus strain AD169. J. Virol. **59:**591–604.
29. **Tenser, R. B., and G. D. Hsiung.** 1976. Comparison of guinea pig cytomegalovirus and guinea pig herpes-like virus: pathogenesis and persistence in experimentally infected animals. Infect. Immun. **13:**934–940.
30. **Thomson, B. J., S. Dewhurst, and D. Gray.** 1994. Structure and heterogeneity of the *a* sequences of human herpesvirus 6 strain variants U1102 and Z29 and identification of human telomeric repeat sequences at the genomic termini. J. Virol. **68:**3007–3014.
31. **Varmuza, S. L., and J. R. Smiley.** 1985. Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. Cell **41:**793–802.
32. **Vieira, J., H. E. Farrell, W. D. Rawlinson, and E. S. Mocarski.** 1994. Genes in the *Hin*dIII J fragment of the murine cytomegalovirus genome are dispensable for growth in cultured cells: insertion mutagenesis with the *lacZ/gpt* cassette. J. Virol. **68:**4037–4046.
33. **Vink, C., E. Beuken, and C. A. Bruggeman.** 1996. Structure of the rat cytomegalovirus genome termini. J. Virol. **70:**5221–5229.
34. **Wadsworth, S., G. S. Hayward, and B. Roizman.** 1976. Anatomy of herpes simplex virus DNA. V. Terminally repetitive sequences. J. Virol. **17:**503–512.
35. **Zimmermann, J., and W. Hammerschmidt.** 1995. Structure and role of the terminal repeats of Epstein-Barr virus. J. Virol. **69:**3147–3155.