JOURNAL OF VIROLOGY, May 1982, p. 547-557
0022-538X/82/050547-11$02.00/0

Vol. 42, No. 2

Construction of a Cloned Library of the *Eco*RI Fragments from the Human Cytomegalovirus Genome (Strain AD169)

JOYCE C. TAMASHIRO, LISA J. HOCK, AND DEBORAH H. SPECTOR*

Department of Biology, University of California, San Diego, La Jolla, California 92093

Received 6 October 1981/Accepted 11 January 1982

The DNA genome of human cytomegalovirus (HCMV) strain AD169 is 158×10^6 M_r. Cleavage of the HCMV DNA with the restriction endonuclease *Eco*RI yields 35 major fragments ranging in size from 0.54×10^6 to 11.4×10^6 M_r. We have constructed a cloned library of the *Eco*RI fragments of this strain of HCMV, using the plasmid pACYC184 and the recipient bacterium *Escherichia coli* strain HB101 $RecA^-$. The viral origin of the cloned inserts was determined by hybridization to viral DNA. The fragments were characterized further by digestion with other restriction enzymes. Several clones were obtained which contained sequences spanning the junction between the long (L) and short (S) components of the viral DNA sequences. These clones differed in molecular weight by multiples of 0.3×10^6 to 0.4×10^6 M_r. The variability found in the clones was also reflected in the genome. Each clone containing a junction sequence hybridized to a series of bands on Southern filters of *Eco*RI-digested HCMV DNA. This "ladder effect" provided evidence for a region of heterogeneity within the L-S junction.

Human cytomegalovirus (HCMV), a member of the herpesvirus family, has gained particular recognition in recent years as its association with a large variety of diseases has been appreciated. HCMV is medically significant both as a cause of birth defects and source of problems in immunosuppressed individuals. In light of the multiple pathogenic manifestations of HCMV, it has become essential to define thoroughly the molecular biology of this virus. A major problem in studying the molecular biology of HCMV is the large size and complexity of the genome. According to several reports, the molecular weight of the HCMV genome is 150×10^6 to 160×10^6 M_r (2, 3, 9, 10, 15). The genome structure, similar to that of herpes simplex virus (16), consists of long (L) and short (S) unique regions bounded by inverted repeats (9). The point at which the repeats flanking the L and S segments join is called the L-S junction. Furthermore, the unique segments may be inverted with respect to one another, establishing the potential for four sequence orientations. The studies of LaFemina and Hayward (9) as well as the results of our experiments (15) indicate that the virus exists in at least two sequence orientations. The cosmid mapping data of B. Fleckenstein (personal communication) and the partial denaturation mapping studies of Kilpatrick and Huang (7) supply evidence for three sequence orientations. It is probable that all four orientations exist. In addition to the variable orientations of the genome segments, herpes simplex (18) and the Towne strain of HCMV (9) also contain regions of heterogeneity within the repeat sequences of the L and S regions. We present evidence in this paper which suggests that the repeat sequences in the AD169 strain of HCMV contain regions of heterogeneity.

To circumvent these problems and to provide large quantities of subgenomic fragments of the HCMV genome, we have cloned *Eco*RI fragments of HCMV strain AD169, using the plasmid vector pACYC184 and the recipient *Escherichia coli* strain HB101 $RecA^-$. The preliminary results of the cloning were presented at the 1980 ICN-UCLA Symposium on Animal Virus Genetics (17). The procedures used for cloning and the characterization of the resulting clones are described below. In the study presented in the accompanying paper (15), we describe the use of these cloned DNA fragments in the construction of restriction endonuclease maps of the genome for the enzymes *Eco*RI, *Hin*dIII, and *Bgl*II.

MATERIALS AND METHODS

Virus and cells. HCMV strain AD169 was obtained from the American Type Culture Collection. Virus titers were determined by plaque assay (19). Human embryonic lung cells were a gift from Stephen Spector and were grown in Dulbecco modified Eagle medium containing 10% calf serum, L-glutamine (0.292 mg/ml),

penicillin-streptomycin (0.2 mg/ml; Irvine Scientific), Amphotericin B (3 μg/ml; Irvine), and gentamicin (50 μg/ml; Schering Corp.). Infection was carried out at a multiplicity of infection of 0.05 to minimize the production of defective virus. When 80% of the cells showed cytopathic effect, [^{3}H]thymidine (5 μCi/ml; Amersham Corp.) was added, and the calf serum reduced to 3%.

Preparation of the viral DNA. Four to five days after the addition of label, extracellular virus was harvested from the medium. The cell debris was removed by low-speed centrifugation, and the virus was precipitated with polyethylene glycol as described by Hamelin and Lussier (5). The virus was pelleted by centrifugation for 1 h at 19,000 rpm in a Beckman 19 rotor at 4°C. The pellets, suspended in DNA buffer (0.1 M NaCl, 0.01 M Tris, and 0.01 M EDTA, pH 8) containing 10% sorbitol (wt/vol), were centrifuged through a sorbitol step gradient (75, 48, and 20% sorbitol) for 1 h at 20°C and 26,500 rpm in a Beckman SW27 rotor. Fractions comprising the viral peak (at the interface of the 48 and 75% sorbitol layers) were pooled, diluted with 0.15 M NaCl–0.05 M Tris, pH 7.2, and centrifuged for 3 h at 26,500 rpm, 20°C, in a Beckman SW27 rotor. The viral pellet was suspended in DNA buffer containing 1% sodium dodecyl sulfate (SDS), treated with RNase A (50 μg/ml; Boeringer Mannheim Corp.) and pronase (1 mg/ml; Calbiochem) at 37°C for 1 h, and extracted twice with 2 volumes of phenol-chloroform-isoamyl alcohol (50:48:2) and twice with 2 volumes of chloroform-isoamyl alcohol (96:4). Two equilibrium cesium chloride centrifugations were performed, using an initial cesium chloride density of 1.72 g/ml. The gradients were centrifuged at 38,000 rpm for 60 h in a Beckman Ti60 rotor. Fractions containing the viral DNA (banding at a density of 1.716 g/ml) were combined and dialyzed against 5 mM Tris–0.1 mM EDTA, pH 7.2. The purity of the isolated DNA was assessed as previously described (17).

Construction, transfection, screening, and isolation of recombinant plasmids. The plasmid pACYC184, which contains tetracycline and chloramphenicol resistance markers (1), was extracted from bacteria and purified by equilibrium cesium chloride-ethidium bromide centrifugation by the method of Kahn et al. (6). Plasmid and HCMV DNA were cleaved with the restriction endonuclease *Eco*RI (generous gift from Pat Green) in 0.1 M Tris, 0.05 M NaCl, 0.005 M $MgCl_2$, and 0.05% Nonidet P-40, pH 7.5, at 37°C for 1 h followed by inactivation of the enzyme at 65°C for 10 min. *Eco*RI-cleaved plasmid was treated with bacterial alkaline phosphatase (Millipore Corp.) at a concentration of 2 U per μg of DNA in 0.01 M Tris–0.1% SDS, pH 9.5, for 1 h at 65°C to prevent recircularization of the plasmid alone. The reaction mixture was extracted twice with 2 volumes of phenol and twice with 2 volumes of ether. The DNA solution was adjusted to 0.2 M NaCl and precipitated by the addition of 2 volumes of 95% ethanol. Viral *Eco*RI restriction fragments were ligated to the cloning vehicle by incubating 1 μg of HCMV DNA, 0.25 μg of pACYC184, and 2.5 U of DNA ligase (Bethesda Research Laboratories) at 4°C in ligase buffer (0.02 M Tris, 0.01 M $MgCl_2$, 0.01 M dithiothreitol, and 0.03 M NaCl, pH 7.6) containing 0.001 M ATP for 36 h (total volume of 25 μl). In some experiments, the viral *Eco*RI fragments were fractionated by size before ligation to plasmid. In these cases, the *Eco*RI fragments were subjected to electrophoresis through 0.8% Seaplaque agarose (Marine Colloids) in buffer containing 40 mM Tris, 20 mM sodium acetate, 18 mM NaCl, and 2 mM EDTA, pH 8, and the bands were visualized by transillumination after staining with ethidium bromide. The gel slice containing the desired band was then melted at 65°C in 10 volumes of 0.3 M NaCl, 10 mM EDTA, and 10 mM Tris, pH 7.2, and loaded onto a benzoyl-naphthol-DEAE (BND)-cellulose column (Serva, Heidelberg) which was maintained at 45°C. The column was washed extensively with loading buffer to remove the agarose, and the DNA was eluted with buffer containing 1 M NaCl, 1% caffeine, 10 mM EDTA, and 10 mM Tris, pH 7.2.

Transfection of the ligated plasmid into *E. coli* strain HB101 $RecA^-$ was done by the $CaCl_2$-RbCl method of Kushner (8). Bacterial colonies containing pACYC184 were selected on agar plates containing 5 to 20 μg of tetracycline per ml and were tested for sensitivity to 25 μg of chloramphenicol per ml. Since *Eco*RI cleaves within the chloramphenicol resistance marker of the plasmid, bacteria containing recombinant plasmid should be sensitive to chloramphenicol.

The selected clones were grown overnight with shaking at 37°C in 5-ml cultures of D medium [15 mM KH_2PO_4, 40 mM K_2HPO_4, 8 mM $(NH_4)_2SO_4$, 0.4 mM $MgSO_4$, 2 mM sodium citrate, 0.5% Casamino Acids, 5 mg of glucose per ml, 10 μg of thiamine per ml, and 5 μg of tetracycline per ml]. A portion from these cultures was mixed with an equal volume of glycerol and placed at −20°C for storage. An inoculum from the 5-ml cultures was diluted 1:50 into 35 ml of D medium and incubated at 37°C, with shaking, until the cultures reached an optical density at 350 nm of 0.3 to 0.35. Chloramphenicol (4.5 ml of a 1-mg/ml solution in D medium) was then added to amplify the plasmid DNA, and the cultures were incubated at 37°C, with shaking, overnight. The recombinant plasmids were isolated by the rapid purification method of Kahn et al. (6), which involves treatment of the bacteria with lysozyme and detergent and a high-speed centrifugation to separate cell debris and chromosomal DNA from the plasmid.

Methods used in characterization of cloned inserts. The following restriction enzymes were used in characterizing the cloned inserts: *Eco*RI, *Hin*dIII, *Bam*HI, *Bgl*II, and *Pvu*II. All restriction endonucleases with the exception of *Eco*RI were obtained from New England Biolabs and were used in the buffer recommended by the supplier. All incubations were carried out at 37°C for 1 h followed by inactivation at 65°C for 10 min. Restriction fragments were separated on 0.8% agarose gels, using the same buffer system as the Seaplaque agarose gels. DNA was transferred from gels to nitrocellulose filters by the method of Southern (14).

DNA was ^{32}P labeled by nick translation (13) with DNA polymerase I (Bethesda Research Laboratories) and [α-^{32}P]dCTP or end labeled by treatment with polynucleotide kinase (Bethesda Research Laboratories) in the presence of [γ-^{32}P]ATP (ICN). In preparation for end labeling, purified HCMV DNA was cleaved with *Eco*RI, extracted twice with phenol and twice with ether, and precipitated with ethanol as described above. The viral fragments were treated with bacterial alkaline phosphatase, and the reaction mixture was extracted with phenol and ether and subjected to ethanol precipitation. The kinasing reac-

tion was performed by incubating the DNA with 12 U of T4 polynucleotide kinase in 50 mM glycine, pH 9.5, 10 mM $MgCl_2$, 5 mM dithiothreitol, 25% glycerol, and 0.2 mM spermidine at 37°C for 1 h. The reaction was inactivated at 65°C for 10 min, diluted with 4 volumes of DNA buffer, and loaded onto a Sephadex G-75 column to separate the labeled viral fragments from free label. Fractions containing the labeled DNA were combined and precipitated in ethanol, using salmon sperm DNA as carrier. The sample was suspended in 5 mM Tris-0.1 mM EDTA, pH 7.2, and stored at 4°C. Hybridization of nick-translated DNA to Southern filters was accomplished by incubating the filters at 37°C for 16 h in 50% formamide, 50 mM *N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid (HEPES), pH 7.0, 1× Denhardt (0.02% each of polyvinylpyrrolidone, Ficoll, and bovine serum albumin), 3× SSC (1× SSC = 0.15 M NaCl plus 0.015 M sodium citrate), 0.1 mg of salmon sperm DNA per ml, and 10 mM EDTA followed by hybridization with labeled DNA in the same buffer at 37°C for 4 to 5 days. After hybridization, filters were washed in 2× SSC for 1 h at room temperature; 0.1× SSC-0.1% SDS at 50°C for 1 h; and several rinses with 0.1× SSC at room temperature.

Colony hybridization. A modification of the method of Grunstein and Hogness (4) was used in the colony hybridizations. Colonies were grown on agar plates containing tetracycline (5 μg/ml) and rubbed onto Whatman 541 filters. The filters were placed sequentially, for the period of time indicated, on Whatman 3MM paper saturated with each of the following buffers: 0.5 N NaOH, 7 min; 1.0 M Tris, pH 7.4, 1 min; 1.0 M Tris, pH 7.4, 1 min; and 1.5 M NaCl-0.5 M Tris, pH 7.4, 5 min minimum. Each filter was then placed on the vacuum apparatus described by Grunstein and Hogness (4). A vacuum was applied for 2 min, 95% ethanol (1 ml per cm^2 of filter) was passed through the filter by vacuum filtration, and the filter was air dried on Whatman 1MM paper. The filters were baked at 80°C for 2 h in a vacuum oven and hybridized with appropriate probes (see Results) in the same manner as Southern filters.

TABLE 1. Cloned *Eco*RI fragments of HCMV DNA

Band nomenclature		Size of *Eco*RI fragment ($M_r \times 10^6$)	No. of clones analyzed
Revised	Original[a]		
A	1a	11.4	7
B	1b	11.2	8
C		11.2	2
D		11.2	3
E	2	10.3	2
F		8.1–9.2[b]	5
G	3	8.3	7
H		7.0–8.2[b]	5
I	4	7.4	4
J	5	6.8	4
K	6	6.3	14
L		5.8	Terminus
M	7	5.0	1
N		5.0	Terminus
O	8c	4.55	5
P	8a,b	4.45	1
Q	9	4.35	5
R	10	3.9	8
S	11a,d,f,g	3.45	9
T	11b,e	3.45	2
U	12	2.85	1
V	13	2.8	8
W		2.8	Terminus
X	14a,b	2.6	2
Y		2.5	1
Z	15	2.4	1
a	16	2.3	6
b	17	1.9	4
c	18	1.65	6
d	19	1.1	1
e	20	1.0	1
f	21	0.95	1
g		0.6	1
h		0.56	1
i		0.54	1

[a] As reported at the ICN-UCLA Symposium on Animal Virus Genetics (17). At the time of our preliminary report, we were not aware of some of the bands.
[b] Heterogeneous joint fragments.

RESULTS

An *Eco*RI restriction digest of HCMV strain AD169 produced 35 major fragments, ranging in size from 0.54 × 10^6 to 11.4 × 10^6 M_r (Fig. 1). Table 1 lists the *Eco*RI restriction fragments and their molecular weights. To be consistent with other workers in the field, we have revised the band numbering system published previously (17) to one using letters. Both the old and revised systems are presented in Table 1.

General methods used in the cloning of the HCMV (strain AD169) genome. The viral DNA used in the cloning was determined to be greater than 99% pure on the basis of solution hybridization to viral DNA. The cloning was first attempted with unfractionated HCMV DNA. In the initial transfection, 0.27 μg of HCMV DNA yielded approximately 9,000 tetracycline-resistant, chloramphenicol-sensitive colonies, which corresponds to a transformation frequency of 3.7 × 10^4 transformants per μg of HCMV DNA. We obtained clones representing 60% of the genome with this method. To facilitate the cloning of the remainder of the genome, gel-purified restriction fragments were used. *Eco*RI-cleaved viral DNA was subjected to electrophoresis on a 0.8% Seaplaque agarose gel, the desired restriction bands were cut from the gel, and the DNA was separated from the agarose through the use of BND-cellulose columns. The isolated viral restriction bands were then ligated to plasmid. The yield of transformed colonies was lower with this procedure, but we were able to obtain clones for all but one viral *Eco*RI band (*Eco*RI fragment F).

*Eco*RI-F and -H are fragments which span the junction of the L and S unique regions of the

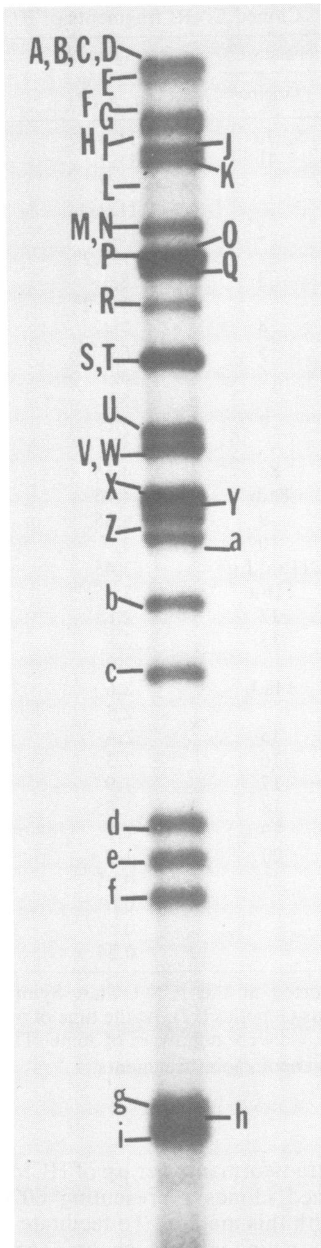

FIG. 1. *Eco*RI restriction digest of HCMV DNA. Purified viral DNA was cleaved with *Eco*RI and end labeled with ^{32}P. The restriction fragments were separated by electrophoresis on a 0.8% agarose gel, and the gel was dried before autoradiography.

genome (see accompanying paper [15]). The data used to construct an *Eco*RI map of the HCMV genome indicate that *Eco*RI cleaves within the repeat sequences bounding the unique L segment of AD169 and within the unique sequence of the S segment. Thus, assuming that L and S segments can invert relative to each other, cleavage with *Eco*RI will produce three termini (labeled fragments L, N, and W) and two junction sequences (fragments F and H); junction F equals L plus W, and junction H equals N plus W. Both fragments spanning the junction should contain the repeat bounding the S segment as well as a portion of the L repeat (identical in sequence to *Eco*RI-W). They should contain different sequences from the S unique region. Clones containing the recombinant plasmid with *Eco*RI-F were identified by the method of Grunstein and Hogness (4). Subfragments of *Eco*RI-H were used as hybridization reagents. The subfragments were obtained by subjecting *Eco*RI-H to double restriction digests with the endonucleases *Pvu*II and *Eco*RI or *Bam*HI and *Eco*RI. The resulting fragments were separated by electrophoresis on a 0.8% Seaplaque agarose gel, isolated by BND-cellulose column chromatography, and ^{32}P labeled by nick translation. Fragments representing either the S unique region (a sequence unique to *Eco*RI-H) or the repeat (common to both *Eco*RI-H and -F) were hybridized to identical Whatman 541 filters containing the transformed colonies. The clones containing *Eco*RI-F were identified by hybridization with the common sequence and the absence of hybridization to the unique fragment.

Characterization of cloned fragments. All 35 major *Eco*RI fragments were represented among the 260 clones characterized, with the exception of the fragments corresponding to the termini of the genome. Table 1 shows the number of clones obtained for each restriction band.

The cloned inserts were first identified on the basis of electrophoretic mobilities. Recombinant plasmids were cleaved with *Eco*RI and subjected to electrophoresis on 0.8% agarose gels. The gels were stained with ethidium bromide to visualize the DNA (Fig. 2). The fragments obtained ranged in size from 11.4×10^6 to 0.54×10^6 M_r. The 2.85×10^6-M_r fragment common to every lane was the cloning vehicle pACYC184. *Hin*dIII-digested lambda DNA served as the molecular weight marker with bands of 18.5×10^6, 15.6×10^6, 6.2×10^6, 4.4×10^6, 2.9×10^6, 1.5×10^6, and 1.3×10^6 M_r.

To verify that the cloned inserts were of viral origin, the *Eco*RI fragments which had been separated by agarose gel electrophoresis were transferred to nitrocellulose filters and hybridized with ^{32}P-labeled HCMV DNA (data not shown). Each of the cloned inserts comigrated with a viral *Eco*RI restriction fragment and hybridized with HCMV DNA. In addition, ^{32}P-labeled cloned fragments were hybridized to Southern filters containing *Eco*RI-digested virion DNA. Each cloned insert hybridized uniquely to the *Eco*RI restriction fragment it was

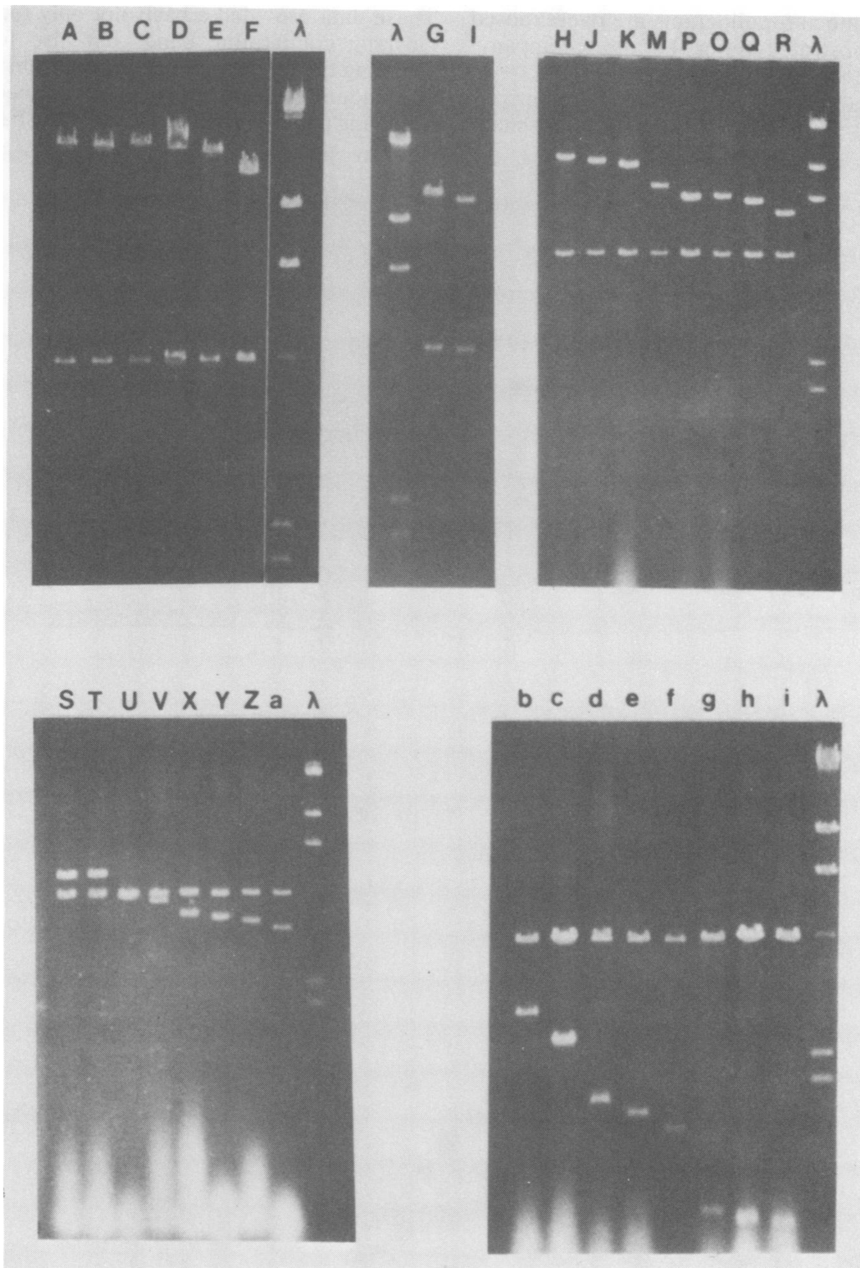

FIG. 2. Representative set of HCMV *Eco*RI clones. Recombinant plasmids were cleaved with *Eco*RI. The resulting fragments were subjected to electrophoresis on 0.8% agarose gels and were visualized with an ethidium bromide stain. *Hin*dIII-cleaved λ DNA served as molecular weight markers ($\times 10^6$): 18.5, 15.6, 6.2, 4.4, 2.9, 1.5, and 1.3. The 15.6×10^6- and 2.9×10^6-M_r bands are not always visible. The plasmid pACYC184 has a molecular weight of 2.85×10^6 M_r.

purported to represent, with the exception of clones containing sequences from the inverted repeats (Fig. 3). The clones representing the junction sequences, *Eco*RI-F and -H, hybridized to each other and to the terminal fragments (further discussion in a later section). *Eco*RI-a and -h mapped partially within the terminal repeats of the L segment of the genome (15) and were observed to cross-hybridize. (The cross-hybridization cannot be seen in Fig. 3 but can be

seen when the autoradiograms are overexposed and is demonstrated in Fig. 4 of the accompanying paper [15]).

The cloned inserts were further characterized by analysis with other restriction enzymes. These data provided a basis not only for distinguishing comigrating bands but also for constructing the linkage groups for different restriction endonuclease fragments as described in the following paper (15). Data for all of the bands are

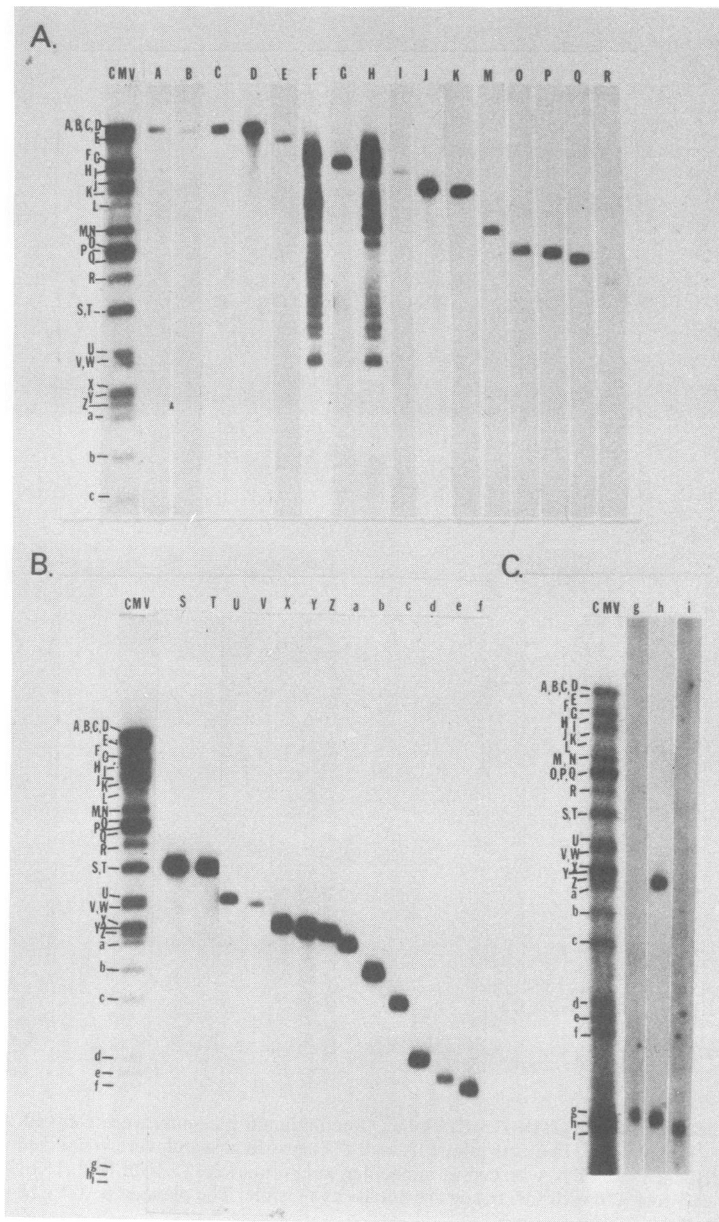

FIG. 3. Hybridization of ^{32}P-labeled cloned HCMV DNA fragments to Southern blots of *Eco*RI-cleaved HCMV DNA. Identical strips from preparative Southern blots containing the *Eco*RI fragments of HCMV DNA were hybridized with the labeled cloned HCMV *Eco*RI fragment indicated at the top of each strip (as described in reference 15). The filter strip labeled "CMV" in panels A and B represents ^{32}P-labeled HCMV DNA cleaved with *Eco*RI and subjected to electrophoresis in a parallel lane of the gel. In panel C, *Eco*RI-digested HCMV DNA end labeled with ^{32}P was included for comparison. Panels A, B, and C represent the autoradiograms of filter strips from separate gels.

presented, but only those *Eco*RI fragments with similar or identical mobilities will be discussed in detail. *Hin*dIII-*Eco*RI and *Bgl*II-*Eco*RI double digestions were performed on all the clones, and the fragments were separated on 0.8% agarose gels. Table 2 presents the molecular weights for the viral fragments.

Analysis of comigrating *Eco*RI fragments. The electrophoretic mobilities of the cloned fragments were usually sufficient for assignment to a particular *Eco*RI restriction band. In a few instances, however, analysis with other restriction enzymes was necessary to distinguish among fragments with similar or identical mobilities.

The very heavy *Eco*RI band of 11.2×10^6 to 11.4×10^6 M_r was composed of four comigrating fragments, A through D. The vastly different restriction patterns obtained with *Hin*dIII-*Eco*RI double digestions of the clones representing these bands suggested that each of these fragments was unique (Table 2). Band A yielded viral fragments of 8.0×10^6 and 3.3×10^6 M_r; double digestion of band B resulted in a 2 M fragment of 3.8×10^6 M_r and 1 M fragments of 2.4×10^6 and 1.1×10^6 M_r; band C produced viral fragments of 4.6×10^6, 4.2×10^6, and 2.45×10^6 M_r; and band D yielded major fragments of 4.4×10^6, 3.9×10^6, 1.4×10^6, and 0.88×10^6 M_r. *Bgl*II-*Eco*RI double digestions also indicated that *Eco*RI fragments A through D were unique. The *Bgl*II-*Eco*RI double-digestion products of these fragments were 5.3×10^6, 4.8×10^6, 0.79×10^6, and 0.54×10^6 M_r for band A; 6.6×10^6, 4.0×10^6, and 0.63×10^6 M_r for band B; 7.5×10^6, 3.4×10^6, and 0.3×10^6 M_r for band C; and 11.0×10^6 and 0.36×10^6 M_r for band D (Table 2). The 0.3×10^6-M_r fragment in band C and the 0.36×10^6-M_r fragment for band D were not visible on our gels, but are supported by our mapping data (15).

Bands O and P were only slightly different in size (4.55×10^6 M_r for O; 4.45×10^6 M_r for P).

TABLE 2. Molecular weights ($\times 10^6$) of viral fragments generated after cleavage of cloned *Eco*RI HCMV DNA fragments with *Hin*dIII and *Bgl*II

*Eco*RI fragment	Mol wt	*Eco*RI + *Hin*dIII digestion products	*Eco*RI + *Bgl*II digestion products
A	11.4	8.0, 3.3	5.3, 4.8, 0.79, 0.54
B	11.2	3.8, 3.8, 2.4, 1.1	6.6, 4.0, 0.63
C	11.2	4.6, 4.2, 2.45	7.5, 3.4, (0.3)[a]
D	11.2	4.4, 3.9, 1.4, 0.88, (<0.5)	~11, (0.36)
E	10.3	6.9, 3.4	4.3, 4.2, 1.35, 0.43
F	8.1	7.0, 1.1	8.1
G	8.3	5.6, 2.7	8.3
H	7.0	7.0	7.0
I	7.4	7.4	2.55, 2.45, 1.9, 0.52
J	6.8	6.8	4.1, 2.7
K	6.3	4.7, 1.6	3.7, 1.9, 0.72
M	5.0	5.0	3.8, 1.15
O	4.55	4.55	4.55
P	4.45	2.3, 2.1	1.35, 1.1, 0.82, 0.55, (0.4), (0.15)
Q	4.35	2.85, 1.45	4.35
R	3.9	3.9	2.4, 1.45
S	3.45	3.45	3.25, (0.2)
T	3.45	3.45	2.2, 1.2
U	2.85	2.85	2.85
V	2.8	1.95, 0.8	2.5, (0.3)
X	2.6	2.6	2.6
Y	2.5	2.5	2.5
Z	2.4	2.4	2.35, (0.05)
a	2.3	1.35, 0.98	2.3
b	1.9	1.8, (0.1)	1.9
c	1.65	1.2, (0.4)	1.65
d	1.1	1.1	1.1
e	1.0	1.0	1.0
f	0.95	0.95	0.95
g	0.6	0.6	0.6
h	0.56	0.56	0.56
i	0.54	0.54	0.54

[a] Numbers in parentheses represent fragments which presumably exist, based on the sum of molecular weights of the double-digestion products, but could not be seen in the gels due to the large amount of RNA in those regions.

Upon digestion with the restriction endonucleases *Hin*dIII and *Eco*RI (Table 2), the two viral fragments produced by band P had molecular weights of 2.3×10^6 and 2.1×10^6 M_r; band O was not cleaved with *Hin*dIII. Band O was also not cleaved with *Bgl*II, whereas band P yielded discernible fragments of 1.35×10^6, 1.1×10^6, 0.82×10^6, and 0.55×10^6 M_r. The results of our mapping data suggest that two additional *Bgl*II fragments of 0.4×10^6 and 0.15×10^6 M_r were also contained in *Eco*RI fragment P (15).

Bands S and T both had a molecular weight of 3.45×10^6 M_r. *Hin*dIII did not cleave within either of these viral sequences, but a *Bgl*II-*Eco*RI double digest revealed differences in cleavage patterns (Table 2). Band S produced a 3.25×10^6-M_r fragment, and there was probably one small fragment of 0.2×10^6 M_r that was not detected. Band T was cleaved into two fragments of 1.2×10^6 and 2.2×10^6 M_r.

Termini and L-S junction fragments. The end-labeling experiments presented in the accompanying paper (15) indicate that the *Eco*RI cleavage of the HCMV strain AD169 produces three major terminal fragments: bands L (5.8×10^6 M_r), N (5.0×10^6 M_r), and W (2.8×10^6 M_r). The results of our mapping data suggest that *Eco*RI cleaves within the inverted repeat bounding the L unique segment resulting in terminal fragment W (15). At the opposite end of the genome, there are no *Eco*RI sites within the inverted repeat bounding the S unique segment, so that the terminal fragments L and N reflect the two possible orientations of the S unique segment. Although the termini cannot be cloned without the use of a linker joined to the terminus, these sequences are represented in the two *Eco*RI fragments, F and H, which span the junction of the L and S segments of the genome (*Eco*RI fragment F contained the L and W sequences; fragment H contained the N and W sequences).

We obtained several clones representing each of the junction sequences. The *Eco*RI-F clones ranged in size from approximately 8.1×10^6 to 9.2×10^6 M_r, whereas the *Eco*RI-H clones ranged from 7.0×10^6 to 8.2×10^6 M_r. Each of the clones representing a particular junction sequence produced similar restriction patterns with another enzyme, *Pvu*II (Fig. 4). In all *Pvu*II-*Eco*RI double digestions, the plasmid produced bands of 2.55×10^6 and 0.3×10^6 M_r. All band F clones had fragments of 3.1×10^6, 1.1×10^6, and 0.74×10^6 M_r in common but differed in the size of one fragment which had molecular weights of 2.8×10^6, 3.2×10^6, 3.5×10^6, and 3.9×10^6 M_r (Table 3). Band H clones also had three common *Pvu*II-*Eco*RI fragments (1.65×10^6, 1.35×10^6, and 0.96×10^6 M_r) and differed in the size of one fragment which had molecular weights of 2.8×10^6, 3.2×10^6, and 4.0×10^6 M_r. When these *Pvu*II digestion products were ^{32}P labeled by nick translation and hybridized individually to nitrocellulose filters containing *Eco*RI-cleaved viral DNA, the *Eco*RI-*Pvu*II fragments containing the region of heterogeneity hybridized strongly to the *Eco*RI bands F, H, and W and weakly to L and N, indicating that these fragments contained the junction of the L and S repeats with only a minor portion of the S repeat represented. The other fragments hybridized mainly to either *Eco*RI fragments F and L

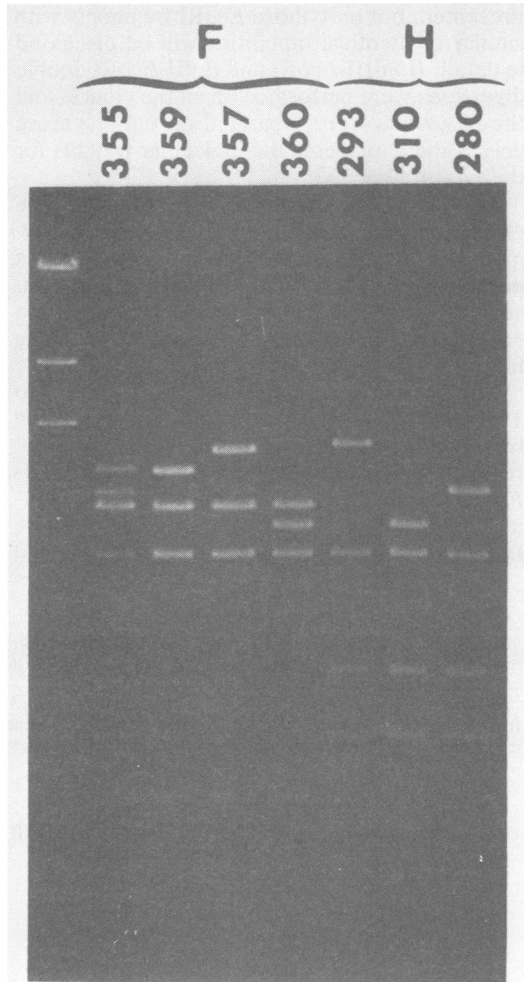

FIG. 4. Comparison of clones representing *Eco*RI bands F and H. Individual recombinant plasmids (number designating clone is indicated at the top of the lane) were cleaved with *Pvu*II plus *Eco*RI and subjected to electrophoresis on 0.8% agarose gels. The bands were visualized with ethidium bromide. *Hin*dIII-digested λ DNA served as molecular weight markers. The plasmid fragments have molecular weights of 2.55×10^6 and 0.3×10^6 M_r. Clone 355 contains two band F junction sequences.

TABLE 3. Molecular weights of restriction fragments obtained in an *EcoRI* or *PvuII* plus *EcoRI* digestion of cloned fragments representing the L-S junction[a]

Clone no.[b]	Mol wt (× 10^6) of *PvuII* + *EcoRI*[c]
355[d]	3.5, 3.2, 3.1, 1.1, 0.74
359	3.5, 3.1, 1.1, 0.74
357	3.9, 3.1, 1.1, 0.74
360	3.1, 2.8, 1.1, 0.74
293	4.0, 1.65, 1.35, 0.96
310	2.8, 1.65, 1.35, 0.96
280	3.2, 1.65, 1.35, 0.96

[a] The data were compiled from analysis of the gels shown in Fig. 4.

[b] Clones 355, 359, 357, and 360 are individual clones of *EcoRI* fragment F. Clones 293, 310, and 280 are individual clones of *EcoRI* fragment H.

[c] Only major digestion products are listed. Fragments less than 0.3 × 10^6 in molecular weight were not detected due to the presence of RNA in that region of the gel.

[d] Clone contains two junction sequences.

or H and N, but not to fragment W, which indicates that these fragments contain sequences from the S unique region (15).

The size heterogeneity seen in the cloned junction sequences was also observed in the viral genome DNA. Due to the inverted repeats bounding the L and S unique regions, we expected the junction fragments to hybridize to all three termini and both junction sequences in an *EcoRI* digest of viral DNA. Instead of the anticipated hybridization to five viral bands, however, the junction sequences hybridized to a series of fragments for each terminus and L-S junction producing a "ladder" effect (Fig. 3). The *EcoRI*-W series of bands appeared to differ by approximately 0.15 × 10^6 M_r; those of the L and N series were spaced by increments of approximately 0.4 × 10^6 M_r. As noted above, the *EcoRI*-F clones differed from one another in size by approximately 0.3 × 10^6 to 0.4 × 10^6 M_r; the *EcoRI*-H clones varied by 0.4 × 10^6 and 0.8 × 10^6 M_r. These differences are consistent with the insertion of one or more copies of a 0.15 × 10^6-M_r sequence in the inverted repeat bounding the L unique region of the genome and a 0.4 × 10^6-M_r sequence in the inverted repeat of the S region. Further studies on these sequences are in progress to determine the length of the repeats and the nature of the heterogeneity.

DISCUSSION

The HCMV genome (strain AD169) consists of linear, double-stranded DNA, with a molecular weight of 158 × 10^6 M_r. The slow growth and cell-associated nature of the virus makes it difficult to obtain large quantities of purified viral DNA. To provide adequate quantities of DNA for studies on the molecular biology of HCMV, we constructed a cloned library of the HCMV genome. We have obtained clones representing all of the HCMV strain AD169 genome by using *EcoRI* restriction fragments of HCMV DNA ligated to the plasmid vehicle pACYC184. The viral origin of the clones was determined by hybridization to purified HCMV DNA, using the method of Southern (14). The clones were initially identified by comparison of their electrophoretic mobilities on agarose gels with *EcoRI*-cleaved HCMV DNA; they were further characterized by double digestion with other restriction enzymes.

We also obtained many clones which contained only small inserts of approximately 0.1 × 10^6 to 0.2 × 10^6 M_r. These cloned inserts were not evident when the recombinant plasmids were cleaved with *EcoRI* because they migrated with the RNA present in our samples. We were able to detect these sequences in uncleaved recombinant plasmids as these had slightly altered migrations on agarose gels compared with pACYC184. Hybridization with HCMV DNA indicated the presence of viral sequences in these plasmids. They are currently undergoing further analysis.

At the time of our previous report (17), we believed there were additional comigrating fragments in the regions of *EcoRI* bands O-P, S-T, X, and a. We referred to these sequences as 8a-c, 11a-g, 14a-b, and 16a-b, respectively, based on the evidence provided by *HindIII* or *BamHI* single digests. Further studies revealed that some of these sequences appeared to be identical by *HindIII*-*EcoRI* and *BamHI*-*EcoRI* double digests as well as by their hybridization to HCMV DNA cleaved with *BglII* or *HindIII*. Cross-hybridization of clones with similar electrophoretic mobilities revealed that 11a, -d, -f and -g (band S) cross-hybridized, as did fragments 11b and -e (band T). Cloned inserts 8a and 8b both contained the sequence we now refer to as *EcoRI*-P, 14a and 14b both contained fragment X, and 16a and 16b both contained fragment a. The altered cleavage patterns in single restriction enzyme digests were caused by plasmids containing two inserted *EcoRI* fragments: a high-molecular-weight fragment and a small sequence of approximately 0.1 × 10^6 to 0.2 × 10^6 M_r. These small inserts are probably viral sequences fortuitously ligated to a plasmid along with another viral fragment, a phenomenon which has been observed to occur with the larger *EcoRI* fragments. Experiments are in progress to determine the origin of these small inserts and to further characterize them.

The HCMV genome contains two major re-

gions, the L and S regions, bounded by inverted repeats (9). The unique segments may be inverted with respect to one another, producing the potential for four sequence orientations. The point at which the repeats flanking the L and S segments join is called the L-S junction. We did not clone the termini of the genome, but identical sequences were contained in the clones representing the L-S junction, due to the presence of the inverted repeats in this region.

The clones representing the junction sequences of HCMV displayed marked size heterogeneity, with the molecular weights differing by multiples of 0.15×10^6 and 0.4×10^6 M_r. This heterogeneity was also present in the genome DNA, indicating that the heterogeneity in the cloned junction fragments was not due to deletions or duplications which had occurred during cloning. When the cloned junction sequences were hybridized to Southern blots of *Eco*RI-cleaved DNA, a series of bands could be resolved for each of the termini and L-S junctions. The series corresponding to the terminus of the L repeat (*Eco*RI fragment W) differed by increments of approximately 0.15×10^6 M_r (230 base pairs [bp]), whereas those representing the two termini of the S segments (*Eco*RI fragments L and N) varied by approximately 0.4×10^6 M_r (600 bp). This variation could be accounted for by a genome structure in which one or more copies of small sequences (0.15×10^6 and 0.4×10^6 M_r) are inserted within the inverted repeats bounding the unique regions.

Similar heterogeneity in the terminal and joint fragments has also been observed for herpes simplex virus type 1 (HSV-1) (11, 12, 18) and the Towne strain of HCMV (9). Wagner and Summers (18) described two types of heterogeneity within the L-S junction and the termini bordering the L segments of HSV-1: the first was the insertion of 280 bp (0.18×10^6 M_r) or multiples of 280 bp; and the second was a variable insertion or deletion of 10 to 50 bp occurring near the site of the 280-bp insertion. Subsequent studies have revealed that various strains of HSV-1 differ with respect to the number and size of the insertions. Locker and Frenkel (11) have suggested that the strain dependence of heterogeneity can be explained by variation in the number of copies of the terminally reiterated "a" sequence. In a preliminary report, LaFemina and Hayward (9) reported the occurrence of terminal heterogeneity in the short repeats of the Towne strain of HCMV. Both termini of the S region (from the two possible orientations) can be resolved into three subspecies which vary in size by multiples of 750 bp (0.5×10^6 M_r). Thus, the occurrence of tandem repeats within the inverted repeats of the genome appears to be a common feature of HSV-1 and HCMV. The Towne and AD169 strains of HCMV appear to be extremely similar with respect to the heterogeneity found in the S segment of the genome (in both cases, the inserted sequences have a molecular weight of 0.4×10^6 to 0.5×10^6 M_r); similarity of the L segment for the two strains could not be determined from the LaFemina and Hayward report. The origin of this heterogeneity in the termini and joints as well as its relationship to mechanisms of replication, recombination, and isomerization of the genome remain to be determined.

ACKNOWLEDGMENTS

We thank Judith Vacquier and Sherrol McDonough for helpful discussions.

This research was supported by Public Health Service grant AI15635 from the National Institutes of Health. J.C.T. is a predoctoral fellow of the National Science Foundation.

LITERATURE CITED

1. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:**1141–1156.
2. **DeMarchi, J. M., M. L. Blankenship, G. D. Brown, and A. S. Kaplan.** 1978. Size and complexity of human cytomegalovirus DNA. Virology **89:**643–646.
3. **Geelen, J. L. M. C., C. Walig, P. Wertheim, and J. van der Noordaa.** 1978. Human cytomegalovirus DNA. I. Molecular weight and infectivity. J. Virol. **26:**813–816.
4. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. **72:**3961–3965.
5. **Hamelin, C., and G. Lussier.** 1979. Concentration of human cytomegalovirus from large volumes of tissue culture fluids. J. Gen. Virol. **42:**193–197.
6. **Kahn, M., R. Kolter, C. Thomas, D. Figurski, R. Meyer, E. Remaut, and D. R. Helinski.** 1980. Plasmid cloning vehicles derived from plasmids ColE1, F, R6K and RK2. Methods Enzymol. **68:**268–280.
7. **Kilpatrick, B. A., and E. S. Huang.** 1977. Human cytomegalovirus genome: partial denaturation map and organization of genome sequences. J. Virol. **24:**261–276.
8. **Kushner, S. R.** 1978. An improved method for transformation of *Escherichia coli* with ColE1 derived plasmids, p. 17–23. *In* H. W. Boyer and S. Nicosia (ed.), Genetic engineering. Elsevier/North-Holland Biomedical Press, Amsterdam.
9. **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, ICN-UCLA Symposia on Molecular and Cellular Biology, vol. XVIII. Academic Press, Inc., New York.
10. **Lakeman, A. D., and J. E. Osborn.** 1979. Size of infectious DNA from human and murine cytomegaloviruses. J. Virol. **30:**414–416.
11. **Locker, H., and N. Frenkel.** 1979. *Bam*I, *Kpn*I, and *Sal*I restriction enzyme maps of the DNAs of herpes simplex virus strains Justin and F: occurrence of heterogeneities in defined regions of the viral DNA. J. Virol. **32:**429–441.
12. **Post, L. E., A. J. Conley, E. S. Mocarski, and B. Roizman.** 1980. Cloning of reiterated and non-reiterated herpes simplex virus 1 sequences as *Bam*HI fragments. Proc. Natl. Acad. Sci. U.S.A. **77:**4201–4205.
13. **Rigby, P. W. J., M. Dieckmann, C. 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.

14. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
15. **Spector, D. H., L. Hock, and J. C. Tamashiro.** 1982. Cleavage maps for human cytomegalovirus DNA strain AD169 for restriction endonucleases *Eco*RI, *Bgl*II, and *Hin*dIII. J. Virol. **42**:558–582.
16. **Spear, P. G., and B. Roizman.** 1980. Herpes simplex viruses, p. 615–745. *In* J. Tooze (ed.), DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Tamashiro, J. C., and D. H. Spector.** 1980. Molecular cloning of the human cytomegalovirus genome (strain AD169), p. 21–37. *In* B. N. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics, ICN-UCLA Symposia on Molecular and Cellular Biology, vol. XVIII. Academic Press, Inc., New York.
18. **Wagner, M. J., and W. C. Summers.** 1978. Structure of the joint region and the termini of the DNA of herpes simplex virus type 1. J. Virol. **27**:374–387.
19. **Wentworth, B. B., and L. French.** 1970. Plaque assay of cytomegalovirus strains of human origin. Proc. Soc. Exp. Biol. Med. **135**:253–258.