JOURNAL OF VIROLOGY, Mar. 1996, p. 2044–2048 Vol. 70, No. 3
0022-538X/96/$04.00+0

Defined Large-Scale Alteration of the Human Cytomegalovirus Genome Constructed by Cotransfection of Overlapping Cosmids

GEORGE KEMBLE,* GREG DUKE, RUTH WINTER, AND RICHARD SPAETE

Aviron, Mountain View, California 94043

Received 22 September 1995/Accepted 14 December 1995

We have constructed defined human cytomegalovirus (CMV) mutants by cotransfecting overlapping cosmid clones spanning the 230-kbp genome. Using this strategy, we have introduced a 13-kbp region of DNA from a virulent strain of CMV into a defined position within the avirulent CMV(Towne) genome. Although more than 80% of the genome of these recombinant viruses was derived from Towne DNA, their plaque morphology more closely resembled that of Toledo. To date, CMV is the largest virus and requires the greatest number of cosmids to be regenerated via overlapping cosmid cotransfection.

The 230-kbp genome of human cytomegalovirus (CMV) ranks among the largest of all DNA virus genomes and contains more than 200 open reading frames (3). The overall genome structure, depicted as *ab*-U_L-*b′a′c′*-U_S-*ca* (Fig. 1A) consists of two separate regions of unique long and short sequences (U_L and U_S, respectively), each of which is bracketed by large inverted repeats. The viral progeny from an infected cell contains four genomic isomers, each differing in the orientation of the U_L and U_S components to one another (17). While molecular and biochemical techniques have elucidated the complex structure of the CMV genome, the inability to efficiently construct defined genetic mutants of CMV has hindered structural and functional studies of many viral genes and their products.

Several biological properties of CMV contribute to the difficulties in generating defined viral recombinants. CMV has a narrow host range in vitro and replicates slowly, reaching peak levels of DNA synthesis approximately 72 h postinfection (17). The generation of the small number of CMV mutants to date has relied on integration of a marker into the genome. The integration event occurs in a permissive host cell via homologous recombination between the parental virus genome and a plasmid containing the marker flanked by sequences homologous to the desired insertion site (14). However, the progeny of such an event predominantly consists of parental virus. The recombinant CMV must be extensively purified by assaying for the marker. Depending on the type of marker used, recombinant virus can be purified by screening or applying selective conditions to the progeny (8, 11, 14, 20). However, all of these procedures can take an extended period of time to purify the small proportion of recombinant virus from the population because of CMV's slow replication, restricted host range, and relatively labile nature in culture.

Cotransfection of four to five overlapping, genome-spanning cosmid clones of either herpes simplex virus, pseudorabies virus, varicella-zoster virus, or Epstein-Barr virus regenerates an infectious virus (4, 5, 18, 19). Introduction of a mutation into one of the cosmids results in the progeny being homogeneous with respect to the mutation; there is no requirement for further purification of the recombinant virus. This strategy led to the generation of recombinant varicella-zoster virus and Epstein-Barr virus, two human viruses that were largely refractory to recombinant techniques because of their highly cell-associated growth in culture (4, 18). Here, we describe the first use of overlapping cosmids to generate a defined, large-scale insertion into the CMV(Towne) genome. Recently, a 13-kbp region of DNA has been identified in the low-passage, virulent strain CMV(Toledo) and other low-passage clinical isolates which has no counterpart in the highly passaged, avirulent Towne strain (2, 13). We used cosmids to insert this region into the Towne genome to begin our assessment of the contribution of these sequences to the virulence of CMV.

Cosmid subclones of the avirulent CMV(Towne) and virulent CMV(Toledo) genomes were constructed in a similar manner. Human foreskin fibroblast (HF) cells were infected with either Towne or Toledo, and after the development of extensive cytopathic effect, DNA was isolated from nucleocapsids by a procedure similar to that used for the preparation of herpes simplex virus nucleocapsids (6). The DNA was partially digested with *Sau*3AI, fractionated by agarose gel electrophoresis, and ligated to the *Bam*HI site of *Bam*HI-*Xba*I-digested arms of the SuperCos · A1 cosmid vector. SuperCos · A1 was derived from SuperCos-1 (Stratagene, San Diego, Calif.) by the insertion of an oligonucleotide incorporating *Srf*I and *Pac*I recognition sequences flanking the unique *Bam*HI site. The majority of cosmids derived from this strategy hybridized to CMV DNA. Fifty-six cosmid clones of Towne were mapped relative to the viral genome, and one set of eight overlapping clones is represented in (Fig. 1A).

The eight cosmids constituting the Towne genome were digested with *Pac*I to release the intact viral insert from the cosmid vector, combined, extracted with phenol-chloroform, and precipitated with ethanol. A $CaPO_4$ precipitate was formed from approximately 16 μg of this mixture and transfected onto 10^6 low-passage (<15 passes) HF cells that had been trypsinized and reseeded 3 to 5 h prior to the addition of the DNA. At 2 to 3 weeks following transfection, plaques with characteristic CMV cytopathic effect were observed, and the virus derived from this transfection had a restriction pattern characteristic of Towne (data not shown).

To generate the desired recombinant virus, we searched for a cosmid derived from Toledo that satisfied two criteria. First,

* Corresponding author. Mailing address: Aviron, 297 N. Bernardo Ave., Mountain View, CA 94043. Phone: (415) 919-6553. Fax: (415) 919-6610. Electronic mail address: gkemble@aviron.com.

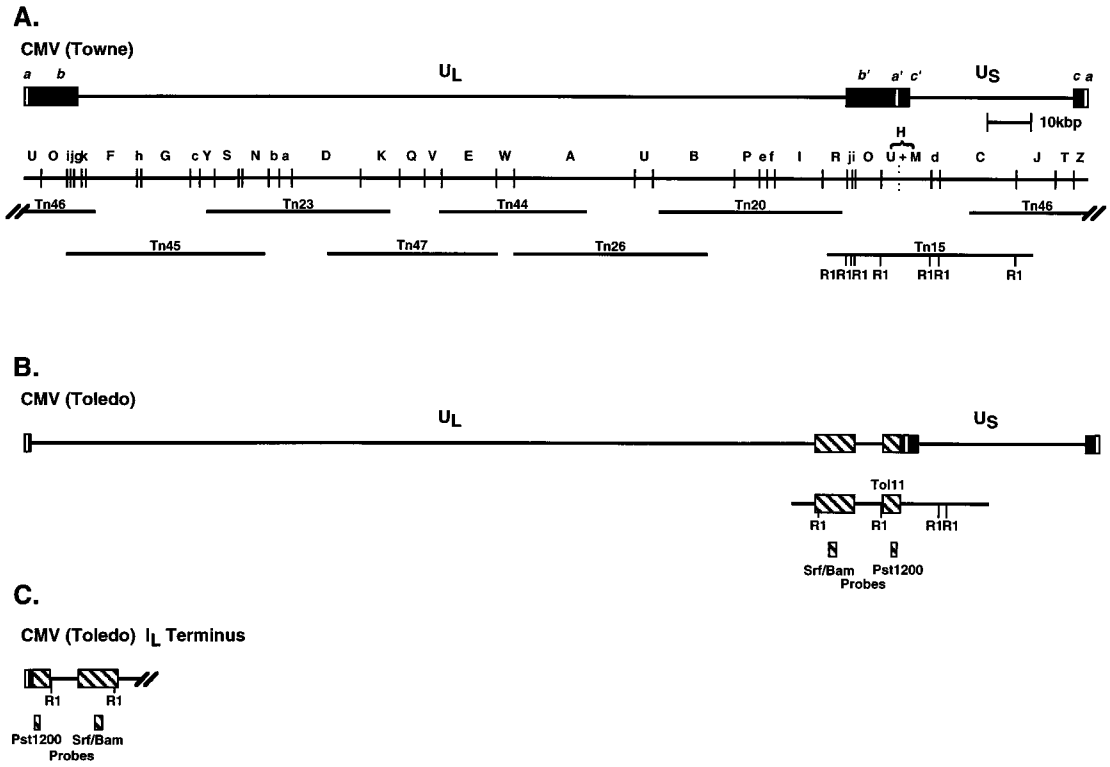

FIG. 1. Schematic representation of the cosmids derived from CMV. (A) Structure of the CMV(Towne) genome (top line) including the long and short unique components (U_L and U_S, respectively) (▬) and the repeated *b*, *c*, (■) and *a* (□) sequences in their noninverted and inverted (′) orientation. The second line depicts the *Eco*RI map of CMV(Towne) adapted from Kemble et al. (10). The positions of the cosmids relative to the genome were determined by *Eco*RI digestion of the cosmid DNA and Southern analysis with probes derived from specific regions of the genome. (B) Depiction of the prototype isomer of Toledo. The *b* repeats are much reduced in size compared to Towne (GenBank no. U33331) (2) (unpublished data). The two large regions of Toledo sequence that are not present in the Towne genome are indicated (▧) (2). The Pst1200 and Srf/Bam probes derived from Toledo-specific DNA were used to determine the structures of the recombinants. (C) Structure of the terminus of a genomic isomer (I_L) in which the U_L segment has been inverted. All diagrams are drawn to scale.

the appropriate cosmid must have contained the 13-kbp region present in Toledo and lacking in Towne (2). Second, this cosmid must have been able to replace Tn15 by overlapping with Tn20 and Tn46. Restriction enzyme digests and Southern blot analyses demonstrated that the cosmid Tol11 contained the Toledo-specific DNA, located near the U_L/*b*′ junction, and overlapped with approximately 3 kbp of Tn20 and 1.5 kbp of Tn46 (Fig. 1B and data not shown). The one Toledo cosmid, Tol11, and the seven Towne cosmids (Tn46, Tn45, Tn23, Tn47, Tn44, Tn26, and Tn20) were digested, mixed, and transfected into HF cells. Approximately 3 weeks posttransfection, plaques were observed. This set of eight cosmids regenerated infectious virus in all four independent transfections attempted.

Because seven of the eight cosmids were derived from Towne, we expected that the majority of the genome would have a restriction enzyme pattern reminiscent of Towne. DNA was extracted from two independent cotransfections, designated Towne/Tol11 V-7 and Towne/Tol11 VI-5, and was analyzed by both restriction enzyme and Southern blot analyses (Fig. 2). As judged from the ethidium bromide-stained agarose gel, the *Eco*RI digestion patterns of the two recombinants closely resembled that of Towne (Fig. 2A). Restriction fragment length polymorphism analysis of a segment of U_L demonstrated the closer relationship of the recombinants to Towne than to Toledo. The Tn45 probe hybridized to *Eco*RI fragments F, G, N, S, Y, and *c* of Towne and *Eco*RI fragments of similar mobilities in both of the recombinants. Because of the similar sizes of the F and G fragments, only five bands were resolved in this experiment. In contrast, several *Eco*RI frag-

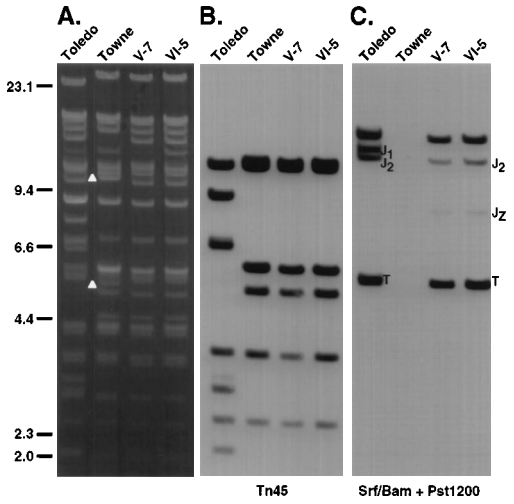

FIG. 2. Structural analysis of recombinant virus genomes by restriction enzyme and Southern blot analyses. (A) Ethidium bromide-stained agarose gel of *Eco*RI-digested viral DNAs. △, positions of the Towne *Eco*RI H and R fragments. (B) Immobilized *Eco*RI-digested DNA hybridized to ECL-labeled Tn45 (Amersham, Indianapolis, Ind.). (C) *Eco*RI-digested DNA hybridized to the Srf/Bam and Pst1200 ECL-labeled probes to detect Toledo-specific sequences. Fragments representing genomic termini (T) and internal junction (J_1, J_2, and J_Z) fragments are indicated. Positions of the lambda/*Hin*dIII sizes are indicated to the left. V-7, Towne/Tol11 V-7; VI-5, Towne/Tol11 VI-5.

ments of Toledo did not comigrate with either Towne or the recombinants (Fig. 2B).

The omission of Tn15 from the cosmid set was predicted to yield recombinants which lacked certain restriction fragments. The *Eco*RI H (U-plus-M junction) and R fragments, portions of which were present in Tn15, had no counterparts of similar sizes in the Toledo genome. The ethidium bromide-stained *Eco*RI digest clearly demonstrated the absence of these two fragments in the recombinants (Fig. 2A). The joining of Towne and Toledo sequences also predicted alterations in other restriction fragments. The right edge of Tn20 overlapped with the left edge of Tol11 within the *Eco*RI I fragment. The DNA sequence at the right edge of fragment I was not conserved between these two strains. An *Eco*RI site within this heterologous region in Toledo produced an I fragment smaller than that of Towne. Crossover between Tn20 and Tol11 within the homologous region of fragment I would predict the incorporation of the right edge of Toledo fragment I, generating a smaller Toledo-size I fragment in the recombinants. As shown in Fig. 2A, the I fragment in Towne comigrated with the H fragment as expected, whereas Toledo and both recombinants had smaller I fragments.

We determined whether the recombinant viruses contained the desired 13-kbp region of Toledo DNA. DNA hybridization probes, which were predicted to hybridize only to Toledo DNA derived from the U_L/b' region (designated Pst1200 and Srf/Bam [Fig. 1]) were isolated. These probes hybridized to the predicted 13-kbp U_L fragment, the 11.5-and 11-kbp junction fragments, and the 5.5-kbp terminal fragment (Fig. 2C). The terminal fragment was generated by inversion of the U_L component (Fig. 1C). None of the Towne fragments hybridized to these probes. The recombinants had a pattern similar to that of Toledo (Fig. 2C). These probes detected fragments of 13, 11, 8, and 5.5 kbp. The slight discrepancy in size between the 13-kbp fragments reflected a gel artifact; other samples showed no such differences in mobility. The 8-kbp fragment (J_Z in Fig. 2C) arose from an inversion of U_S, joining the Toledo U_L/b' region in the recombinants to the Towne *Eco*RI Z fragment. Since Toledo has a much larger fragment than the Towne Z fragment at the analogous position in its genome, the junction formed by inversion of U_S in Toledo was much larger (J_1 in Fig. 2C).

The recombinant viruses described here replace most of the 10-kbp b' repeat with 13 kbp of sequence present in CMV(Toledo). This results in an overall 3-kbp expansion of the Towne genome. The recombinants demonstrated the fidelity of this system in reproducing several features of the CMV replication cycle. The unique components isomerized, and both the Towne and Toledo *a* sequences were used as cleavage/packaging signals in the recombinants. This result was expected, given that even herpes simplex virus can use the CMV *a* sequence as a cleavage/packaging signal (15). The recombinant viruses appear to be genetically stable on the basis of limited passage in cell culture to date. It will be of interest to document the identity of the *a* sequences after extended passage in vitro to determine whether a genetic difference is retained or whether the *a* sequence from one strain becomes predominant.

During analysis and mapping of cosmid clones, the *Eco*RI E fragment of Towne was demonstrated to have significant heterogeneity (Fig. 3). The Toledo E fragment did not exhibit this heterogeneity. The *Eco*RI E region of the Towne genome borders the origin of lytic-phase DNA replication (*ori*Lyt) (1, 7, 12). The cosmid clones spanning this region harbored either one (Tn47) or two (Tn44) multimeric units of this heterogeneity (data not shown). The recombinant viruses predominantly had either the smallest (V-7) or second smallest (VI-5) forms of *Eco*RI E, most likely representing the cosmid from which this region of the genome was derived (Fig. 2A and 3). Regardless of the predominant form, Fig. 3 demonstrates that the recombinants had at least three heterogeneous forms of the E fragment; upon longer exposure, additional slowly migrating species were detected for Towne and both of the recombinants.

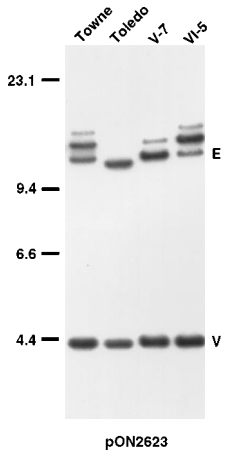

FIG. 3. Heterogeneities in the *Eco*RI E region of the cosmid-derived viruses. *Eco*RI-digested viral DNAs were subjected to Southern blot analysis utilizing a minimal *cis ori*Lyt probe, ^{32}P-labeled pON2623. This probe is a subclone of the *Pvu*I-*Kpn*I fragment of pON2400 and spans the junction of the *Eco*RI V and E fragments (12). Positions of the lambda/*Hin*dII sizes are indicated to the left.

This heterogeneous *Eco*RI E fragment in Towne was regenerated in the two cosmid-derived recombinants, even though only one or two copies of the multimeric unit was used to regenerate the virus. This heterogeneity may be influenced by the proximity of *Eco*RI E DNA to *ori*Lyt. Although Hamzeh et al. have mapped the initiation of DNA synthesis to the Towne *Eco*RI V fragment, the minimal *cis* region that contains a functional *ori*Lyt maps to two corresponding fragments of AD169 *Eco*RI fragments V and I (corresponding to the *Eco*RI V and E fragments of Towne, respectively) (1, 7, 12). This region of the CMV genome has a very high density of repeat and palindromic sequences which may influence the recombinogenic character of this region. This situation is reminiscent of the human herpesvirus type 6 origin of DNA replication. This origin sequence amplifies during passage of the virus in cell culture. The human herpesvirus type 6 *ori*Lyt region also has a high density of repeat structures (16). Although no direct evidence yet exists, the viruses with multimeric origins may have a growth advantage in culture.

During the propagation of the recombinant viruses, we observed that their plaque morphology was different than that of Towne. Plaques formed from Towne-infected cells appeared to have more space between infected cells, and the individual cells remained more elongated and fibroblastic in shape (Fig. 4). However, plaques formed by Toledo-infected cells appeared as clusters of rounded cells. The Toledo-infected cells themselves appeared more refractile and spherical than those infected by Towne (Fig. 4). Infection of HF cells with recombinant virus generated plaques and infected cells that more closely resemble those of Toledo (Fig. 4). Since more than 80% of the genome of the recombinants was derived from Towne DNA, the plaque morphology must result from the replacement of the Towne U_L/b' region with Toledo DNA. This replacement inserted approximately 13 kbp of unique Toledo DNA and removed approximately 2 kbp of unique

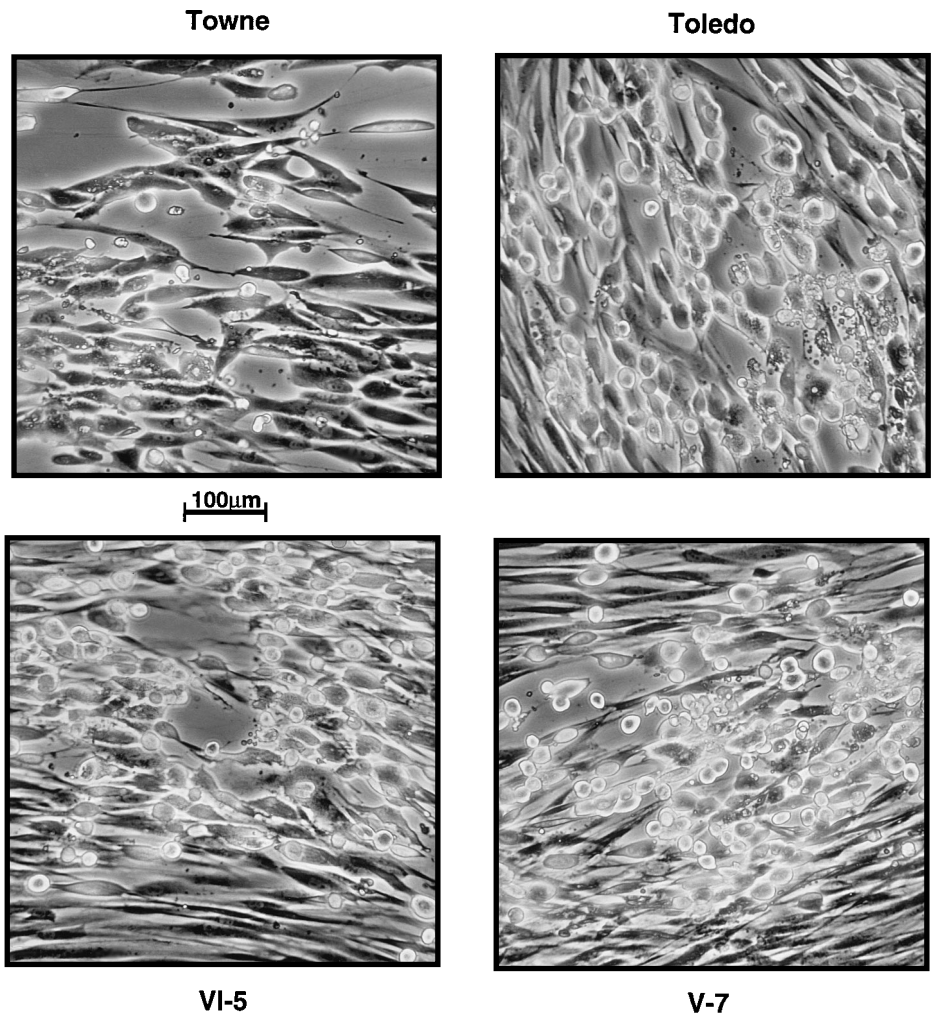

FIG. 4. Plaque morphologies of recombinant viruses in comparison with Towne and Toledo. HF cells were infected at a multiplicity of infection of approximately 0.001 with the indicated viruses and were incubated in the presence of 0.1% pooled human gamma globulin for 9 days. After 9 days postinfection, the cells were photographed. Towne and Towne regenerated by cosmid cotransfection had the same plaque morphologies (9).

Towne DNA (Fig. 1) (2). Although we cannot rule out the possibility that deletion of this small region of Towne resulted in the generation of a Toledo-like plaque morphology, we believe that incorporation of the Toledo DNA determined the plaque morphology of these recombinants.

We have utilized cotransfection of overlapping CMV DNA fragments to generate recombinant CMV. This technique yielded a pure population of virus harboring the desired mutation. This strategy eliminates the time-consuming and difficult steps involved in purifying the recombinant virus from the overwhelming majority of nonmutant virus present in stocks derived by other techniques. We have generated a large-scale alteration of the CMV(Towne) genome utilizing overlapping cosmids; however, deletions and insertions as well as single- or multiple-point mutations can be introduced in the same manner.

The technique presented here for generating recombinant CMV genomes will enable rapid analysis of the viral genome, which has heretofore been difficult. Eliminating the need to purify the recombinant from an excess of parental virus will enhance the efficiency of mutant virus generation. Furthermore, avoiding the use of marker insertion will ease analysis of the phenotype, since markers carried by the virus could influence the viral replication cycle. In addition, viruses which are destined for vaccine trials in humans may have altered immunogenicity because of expression of the marker. We believe this advance will accelerate research of the molecular biology and pathogenicity of CMV.

REFERENCES

1. **Anders, D. G., M. A. Kacica, G. Pari, and S. M. Punturieri.** 1992. Boundaries and structure of human cytomegalovirus *ori*Lyt, a complex origin for lytic-phase DNA replication. J. Virol. **66:**3373–3384.
2. **Cha, T.-A., E. Tom, G. W. Kemble, G. M. Duke, E. S. Mocarski, and R. R. Spaete.** 1996. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. J. Virol. **70:**78–83.
3. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchinson III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell.** 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. **154:**125–170.
4. **Cohen, J. I., and K. E. Seidel.** 1993. Generation of varicella-zoster virus (VZV) and viral mutants from cosmid DNAs: VZV thymidylate synthetase is not essential for replication in vitro. Proc. Natl. Acad. Sci. USA **90:** 7376–7380.
5. **Cunningham, C., and A. Davison.** 1993. A cosmid-based system for constructing mutants of herpes simplex virus type 1. Virology **197:**116–124.
6. **Denniston, K. J., M. J. Madden, L. W. Enquist, and G. Vande Woude.** 1981.

Characterization of coliphage lambda hybrids carrying DNA fragments from herpes simplex virus type 1 defective interfering particles. Gene **15:**365–378.
7. **Hamzeh, F. M., P. S. Lietman, W. Gibson, and G. S. Hayward.** 1990. Identification of the lytic origin of DNA replication in human cytomegalovirus by a novel approach utilizing ganciclovir-induced chain termination. J. Virol. **64:**6184–6195.
8. **Jones, T. R., V. P. Muzithras, and Y. Gluzman.** 1991. Replacement mutagenesis of the human cytomegalovirus genome: US10 and US11 gene products are nonessential. J. Virol. **65:**5860–5872.
9. **Kemble, G., and R. Winter.** Unpublished data.
10. **Kemble, G. W., A. L. McCormick, L. Periera, and E. S. Mocarski.** 1987. A cytomegalovirus protein with properties of herpes simplex virus ICP8: partial purification of the polypeptide and map position of the gene. J. Virol. **61:** 3143–3151.
11. **Kemble, G. W., and E. S. Mocarski.** 1989. A host cell protein binds to a highly conserved sequence element (pac-2) within the cytomegalovirus *a* sequence. J. Virol. **63:**4715–4728.
12. **Masse, M. J. O., S. Karlin, G. A. Schachtel, and E. S. Mocarski.** 1992. Human cytomegalovirus origin of DNA replication (oriLyt) resides within a highly repetitive region. Proc. Natl. Acad. Sci. USA **89:**5246–5250.
13. **Quinnan, G. V., M. Delery, A. H. Rook, W. R. Frederick, J. S. Epstein, F. Manischewitz, L. Jackson, K. M. Ramsey, K. Mittal, S. A. Plotkin, and M. R. Hilleman.** 1984. Comparative virulence and immunogenicity of the Towne strain and a nonattenuated strain of cytomegalovirus. Ann. Intern. Med. **101:** 478–483.
14. **Spaete, R. R., and E. S. Mocarski.** 1987. Insertion and deletion mutagenesis of the human cytomegalovirus genome. Proc. Natl. Acad. Sci. USA **84:** 7213–7217.
15. **Spaete, R. R., and E. S. Mocarski.** 1985. The *a* sequence of the cytomegalovirus genome functions as a cleavage packaging signal for herpes simplex virus defective genomes. J. Virol. **54:**817–824.
16. **Stamey, F. R., G. Dominguez, J. B. Black, T. R. Dambaugh, and P. E. Pellett.** 1995. Intragenomic linear amplification of human herpesvirus 6B *ori*Lyt suggests acquisition of *ori*Lyt by transposition. J. Virol. **69:**589–596.
17. **Stinski, M. F.** 1990. Cytomegalovirus and its replication, p. 1959–1980. *In* B. N. Fields and D. M. Knipe (ed.), Virology, 2nd ed., vol. 2. Raven Press, New York.
18. **Tomkinson, B., E. Robertson, R. Yalamanchili, R. Longnecker, and E. Kieff.** 1993. Epstein-Barr virus recombinants from overlapping cosmid fragments. J. Virol. **67:**7298–7306.
19. **van Zijl, M., W. Quint, J. Briaire, T. de Rover, A. Gielkens, and A. Berns.** 1988. Regeneration of herpesviruses from molecularly cloned subgenomic fragments. J. Virol. **62:**2191–2195.
20. **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 a *lacZ*/*gpt* cassette. J. Virol. **68:**4837–4846.