VIROLOGY 215, 61–72 (1996)
Article No. 0007

Nucleotide Sequence and Molecular Analysis of the Rhesus Cytomegalovirus Immediate-Early

PETER A. BARRY,[1] DONALD J. ALCENDOR,[2] MICHAEL D. POWER,* HEATHER KERR, and PAUL A. LUCIW

*Department of Medical Pathology, University of California, Davis, California 95616; and *Athena Neurosciences, Inc., South San Francisco, California 94080*

Received May 31, 1995; accepted October 20, 1995

Cytomegalovirus (CMV) has been isolated from many nonhuman primates, including rhesus macaques (*Macaca mulatta*). To better understand the molecular biology of rhesus CMV (RhCMV), a 9.2-kb DNA restriction fragment spanning the immediate-early (IE) gene has been molecularly cloned and sequenced. Open reading frames (ORF) have been identified and transcripts mapped for regions corresponding to exons 1, 2, 3, and 4 of the IE1 protein of human CMV (HCMV) and to exons 1, 2, 3, and 5 of IE2. The predicted RhCMV IE1 protein was 29 and 40% identical with the HCMV and African green monkey (AGM) CMV IE1 proteins, respectively, and the predicted RhCMV IE2 protein was 48 and 65% identical with the HCMV and AGM CMV IE2 proteins, respectively. Five additional ORF 3′ to the RhCMV IE gene were identified which contained significant homologies with the HCMV UL121–UL117 ORF. The predicted translation products ranged from 29 to 47% identical with, and 52 to 66% similarity to, the corresponding ORF of HCMV. Conservation of nucleic and amino acid sequences, and colinearity of genes, between primate CMV genomes contribute to a better understanding of primate CMV evolution, regulation, and pathogenesis.

INTRODUCTION

Human cytomegalovirus (HCMV) is ubiquitous in human populations. In the majority of cases CMV infections are subclinical, although in select circumstances (i.e., acquired immune deficiency, iatrogenically immunosuppressed individuals, and congenital transmission) CMV can cause debilitating or life-threatening disease (Ho, 1991). The mechanisms of CMV pathogenesis in these situations remain to be resolved as well as questions concerning latency or persistent infection in healthy individuals. Studies of a nonhuman primate CMV should be directly relevant to studies on the prevention, treatment, and understanding of CMV disease in humans because humans and nonhuman primates share close developmental and evolutionary relationships. Nonhuman primates have proven to be extremely useful for many studies of human biology (Hendrickx and Binkerd, 1990).

CMV was first recognized as an incidental, infectious agent in rhesus macaques over 60 years ago (Covell, 1932; Cowdry and Scott, 1935). Since that time, distinct CMV strains have been isolated from numerous species of nonhuman primates (Asher *et al.,* 1974; Eizuru *et al.,* 1989; Smith *et al.,* 1969; Swack *et al.,* 1971). Each viral isolate demonstrates fundamental differences from HCMV in host range, protein, and DNA content (Eizuru *et al.,* 1989; Gibson, 1983; Weiner and Gibson, 1981). While sharing antigenic and DNA relationships, the CMV isolates from different nonhuman primate species represent distinct strains (Asher *et al.,* 1974; Eizuru *et al.,* 1989; Smith *et al.,* 1969). The immediate-early (IE) gene of the African green monkey (AGM) CMV has been extensively analyzed; this gene shows about 40 to 75% homology to the HCMV IE gene (depending on the region analyzed) and shares a similar pattern of introns and exons (Chiou *et al.,* 1993; Jeang *et al.,* 1984, 1987, 1982; Jeang and Gibson, 1980; Chang *et al.,* 1995). Rhesus (Rh) CMV is a common infectious agent in captive rhesus macaque populations with rates of infection approaching 100% by the 1st year of life, well before sexual maturity (Kalter and Heberling, 1990; Kessler *et al.,* 1989; Vogel *et al.,* 1994). Virus is periodically shed in the urine from seropositive animals and can continue for the life of a healthy host (Asher *et al.,* 1974). As with humans, the majority of RhCMV infections are subclinical.

Similar to humans infected with human immunodeficiency virus, a proportion of rhesus macaques experimentally infected with simian immunodeficiency virus (SIV) suffer CMV-induced disease. Postmortem histological analysis of SIV-infected macaques has identified disseminated RhCMV infection in multiple tissues, including brain, lungs, gut, liver, and spleen (Baskin, 1987). A similar spectrum of RhCMV pathology has also been also

[1] To whom correspondence and reprint requests should be addressed. Fax: (916) 752-4548. E-mail: pabarry@ucdavis.edu.

[2] Current address: Molecular Virology Laboratories, Department of Pharmacology and Molecular Sciences, Johns Hopkins University, Baltimore, MD 21205-2185.

0042-6822/96 $12.00

observed in animals experimentally infected with the immunosuppressive simian retrovirus, a type D retrovirus (Osborn *et al.,* 1984). Experimental inoculation of fetal rhesus macaques with RhCMV has produced disease patterns, particularly of the brain, closely resembling CMV-induced sequelae of congenitally infected humans (London *et al.,* 1986) (A. F. Tarantal, S. Salamat, and P. A. Barry, manuscript in preparation). Thus, the biology of CMV in rhesus macaques mirrors the biology of HCMV in immunocompetent and immunodeficient individuals.

We have been investigating RhCMV both at the biological and at the molecular levels to use infection of rhesus macaques with RhCMV as a model for HCMV pathogenesis. To that end, a 9.2-kb fragment spanning the RhCMV IE gene has been cloned (Alcendor *et al.,* 1993). Analysis of the promoter region of the RhCMV IE gene (from −733 to +623, relative to the start site of transcription at +1) revealed that this region was 70 and 48% conserved with the nucleotide sequences of the AGM and human CMV IE promoters, respectively. In the present study, we have elucidated the nucleotide sequence and transcriptional map of the RhCMV IE gene. Also, additional open reading frames (ORF) corresponding to the UL121, 120, 119, 118, and 117 ORF of HCMV have been identified (Chee *et al.,* 1990).

MATERIALS AND METHODS

Plasmids

The plasmid pRhIE 9.4-2 has been previously described (Alcendor *et al.,* 1993). This plasmid contains a 9.2-kb *Sal*I–*Xho*I restriction fragment which spans the RhCMV IE gene cloned into the pSP72 vector (Promega, Madison, WI).

Cells and virus

All transfections and infections were done in U-373MG cells [human glioblastoma, American Type Culture Collection (ATCC), Rockville, MD]. Virus infections were done with the 68-1 strain of RhCMV (Asher *et al.,* 1974) (ATCC). U-373MG cells are fully permissive for RhCMV infection (data not shown).

RNA isolation and analysis

RNA was isolated as described by Chomczynski and Sacchi (1987). Reverse transcriptions (RT) were performed with a total of 1 μg of RNA from each sample in a reaction volume of 20 μl using 200 units (U) Superscript II RNase H^- reverse transcriptase (GIBCO/BRL, Gaithersburg, MD), RNasin (20 U) (Promega, Madison, WI), and random hexamers as primers (100 pmol) (Pharmacia, Piscataway, NJ); for amplification of full-length IE2 transcripts from cells transfected with p9.4-2, 5 μg of total RNA was used for RT. Reverse transcription reactions were performed in recommended buffers (GIBCO/BRL) at 42° for 60 min and terminated by heating at 75° for 15 min. For 3′ rapid amplification of cDNA ends (RACE) (Frohman *et al.,* 1988), 2 pmol of the $(dT)_{17}$-adapter primer (Frohman *et al.,* 1988) was used, instead of random hexamers. For RACE, the cDNAs were tailed and amplified with one round of PCR according to the protocol of Frohman *et al.* (1988).

PCR

Each cDNA (1–2 μl) was amplified in a reaction volume of 50 μl with a combination of *Taq* polymerase (0.5 U/reaction) (Promega) and *Taq* Extender (0.5 U/reaction) (Stratagene, La Jolla, CA) using the buffer recommendations for *Taq* Extender (Stratagene); for amplification of full-length IE2 transcripts from cells transfected with p9.4-2, 5 μl of RT reaction was used for RT. Primers for PCR are given in Table 1. The following cycle times were used for the first round of amplification on a Temp · Tronic thermocycler (Thermolyne, Dubuque, IA): 1 cycle of 4 min at 94°/2 min at 55°/5 min at 72°, 30 cycles of 1 min at 94°/1 min at 55°/1 min at 72°, and a 15-min dwell at 72°. For amplification of full-length IE1 and IE2 transcripts, 2-min extensions were used. Reaction products were analyzed by agarose gel electrophoresis and staining in ethidium bromide.

Sequencing

For sequencing, PCR products were cloned into the TA cloning vector (Invitrogen, San Diego, CA). The 9.2-kb genomic clone and all PCR products were sequenced by the dideoxy chain termination method (Sanger *et al.,* 1977). DNA sequences were analyzed using the Program Manual for the Wisconsin Package [V. 8; Genetics Computer Group (GCG), Madison, WI] and with MacDNAsis (V. 3.5; Hitachi, San Bruno, CA.)

GenBank Accession Number

The GenBank Accession No. for the sequence of the IE region of RhCMV is M93360.

RESULTS

Nucleotide composition of the RhCMV IE region

The nucleotide sequence for the remainder of the cloned fragment within p9.4-2 (Alcendor *et al.,* 1993) has been determined; the GenBank Accession No. is M93360. The G+C content of the entire sequence (−733 to +8471, relative to start site of transcription) was 46.9%; the GC content for a region that included just the IE gene, exons 1 through 5 (described below), was 47.2%, similar to that of the AGM CMV IE gene (49.1%) (Chang *et al.,* 1995). The G+C content of the IE gene of the Towne and AD169 strains of HCMV are 53.4 and 52.9%, respectively (Chee *et al.,* 1990; Stenberg *et al.,* 1984, 1985).

Within the RhCMV IE region, there was a striking deficiency in the expected frequency of CpG dinucleotide

TABLE 1

Primers Used for RT/PCR Analysis

No.		Position[a]	Exon
Upstream primers (5′ → 3′)			
175	GAG GAG CAC CAT AGA A	20–35	1
49	AGG GTA CCG GGA CCG ATC C	35–53	1
57	GGG TGC ATT GGA ACG C	71–86	1
60	GCT AAT GAT GAT CCA AT	420–436	3
55	TAG AGC AGA CTG AGG G	2775–2790	4
143	CAC CCT ATG CAA ACC AGA	2927–2944	4
58	CTT GTG CAG AGA TTC C	4747–4762	5
S7	AGA AAG CCA TCG AGG ACA	4910–4927	5
Downstream primers (5′ → 3′)			
53	CAG TTT GCC GGA TGA G	1885–1870	4
48	TCT GTC TTT GAT ATC ACC T	1939–1920	4
56	TCA GAC CCT ATC TGA G	2739–2724	4
176	AAT GCA GTT ATA GGG ACC	3010–2993	4
S2	AAT CCC TGC TTG ACT GAC	3444–3427	5
170	CTA TTG CTG CTG AGG GT	3634–3618	5
S3	TCA CAA TAC ATC CAG AGG	3883–3866	5
44	CAA GGATCC CAA GAA GTC CAC	3964–3944	5
S5	CAG TGC AAT GTT AAG TCC	4978–4961	5
S4	TAA GTC CAC TGT GCA GTT	4967–4950	5

[a] The numbering is relative to the start site of transcription (+1). The start site of transcription corresponds to position 734 of the GenBank sequence (M99360).

residues (Fig. 1A). The deficit in CpG dinucleotides began immediately downstream of the start site of transcription and extended well beyond the 3′ end of the IE2 polyadenylation site. (Mapping of transcripts is described below.) This was particularly evident for most of the IE coding region (+1078 to +4400). Between positions +982 and +3723, there were only four CpG dinucleotides, two within exon 3 and two within exon 4. In contrast, the frequency of GpC dinucleotides was generally very close to the expected frequency; there were no dramatic deviations from expected numbers for other dinucleotides (not shown).

Predicted translation products

The nucleotide sequence downstream from the start site of IE transcription (Alcendor *et al.,* 1993) was translated in all three reading frames to identify potential ORF. A number of larger ORF were identified that had homologies to the IE1 and IE2 proteins and the uncharacterized UL121, 120, 119, 118, and 117 ORF of HCMV (Fig. 1B) (described in detail below).

Mapping of IE transcripts and comparisons of exon/intron sequences

Based on the predicted translation products of these ORF, oligonucleotide primers were designed to map IE transcripts by a combination of RT/PCR and 5′ and 3′ RACE. Two sources of RNA were used for these studies. The plasmid pRhIE 9.4-2, containing the entire 9.2-kb insert, was transfected into U-373MG cells, and RNA was isolated 48 hr posttransfection. This was done because the 22659 strain of RhCMV, from which the 9.2-kb insert was cloned (Alcendor *et al.,* 1993), grew to extremely low titers in tissue culture. To determine that the transcription pattern of the transfected DNA was authentic to transcripts observed with viral infection, RNA was also isolated from U-373MG cells infected with the 68-1 strain of RhCMV in the presence of the translation inhibitor cycloheximide. The 68-1 strain, isolated in 1968 (Asher *et al.,* 1974), exhibits a strong conservation of restriction fragment lengths with the 22659 strain (Alcendor *et al.,* 1993) and grows to high titers.

IE transcripts were mapped using a variety of primer pairs (Table 1), and the results are presented in Fig. 2. For each primer pair, the amplification product of RNA isolated from cells transfected with pRhIE 9.4-2 was identical in size to that obtained with RNA from cells infected with RhCMV strain 68-1 (Figs. 2A–2C). To identify splice junctions for IE1 and IE2 transcripts, representative PCR products were cloned and sequenced. The IE1 and IE2 splice junctions of pRhIE 9.4-2 were sequenced using the PCR product from primer pairs 57/53 and 60/44, respectively (Table 1); splice sites for RhCMV strain 68-1 were mapped using primer pairs 49/176 and 49/S4. Only the PCR product for pRhIE IE1 is shown (lane 3, Fig. 2A). In each case, the splice sites for both pRhIE 9.4-2 and RhCMV 68-1 were identical.

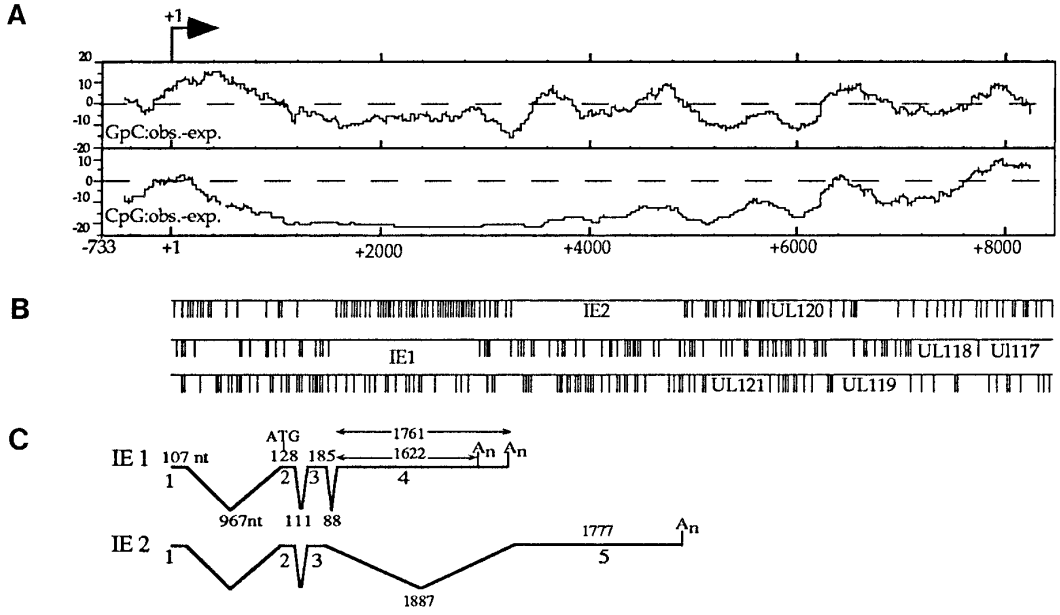

FIG. 1. Organization of the RhCMV IE gene and 3′ flanking region. (A) The frequencies of GpC and CpG dinucleotides for the entire sequenced region are plotted. The frequencies were calculated over a window of 300 bp and are presented as the observed frequency less the expected frequency. The expected frequency was calculated based on the frequency of C and G mononucleotides. The dashed line indicates the point at which the observed frequency equals the expected frequency. (B) All three forward reading frames were translated to identify extended ORF. The downward lines for each reading frame indicate translation stop codons. The translational analysis is aligned with the numbering in A. (C) The transcription maps for IE1 (exons 1–4) and IE2 (exons 1–3, 5) are presented with the lengths of the exons and introns (in nucleotides, nt) and sites of polyadenylation signal (A_n) are indicated. The transcription map is aligned with A.

The RhCMV IE1 transcript was identical in organization to those of HCMV (Stenberg *et al.,* 1984) and AGM CMV (Chang *et al.,* 1995) and was composed of four exons with lengths of 107 (exon 1), 128 (exon 2), and 185 nucleotides (nt) (exon 3); exon 4 was either 1622 or 1761 nt in length, depending on utilization of either of two IE1 polyadenylation signals (discussed below) (Fig. 1 and Table 1). Sizes for exons 1–4 of HCMV IE1 are 121, 88, 185, and 1315 nt, respectively (Stenberg *et al.,* 1984), and for AGM CMV IE they are 109, 153, 194, and 1513/1632 nt (depending on polyadenylation site for exon 4) (Chang *et al.,* 1995) (Table 1). The size of exon 5 for RhCMV was 1777 nt, compared to 1612 (AGM) and 1570 nt (H). Lengths of RhCMV IE introns 1–3 were 967, 111, and 88 nt, compared to intron sizes for HCMV of 827, 114, and 170 nt (Mach *et al.,* 1989; Stenberg *et al.,* 1984) and 895, 113, and 86 nt for AGM CMV (Chang *et al.,* 1995). All RhCMV introns were bounded by a 5′-terminal GT dinucleotide and a 3′-terminal AG dinucleotide. Aside from the terminal dinucleotides, there was very little sequence homology between these three primate CMV when intron sequences flanking splice junctions were compared (data not shown); very little homology was noted when just the RhCMV and AGM CMV intron termini were aligned.

For all amplifications, the predicted PCR products with primer pairs 57/53 (649 nt), 60/48 (431 nt), 60/170 (328 nt), 60/S3 (575 nt), 57/S2 (409 nt), and 175/S2 (460 nt) agreed with the observed amplification products (lanes 2–10, 13–16, Fig. 2A). Similarly, amplification with primer pairs 175/176 (IE1, 1825 nt) and 175/S5 (IE2, 1995 nt) (lanes 3–6, Fig. 2C) yielded nearly full-length products consistent with predicted sizes.

Several features emerged when the nucleic acid sequences for the different components of the IE region were compared. First, for all exons and introns, RhCMV and AGM CMV exhibited higher sequence identity to each other than either of these two had with HCMV (Table 2). Second, for most of the IE regions, the AGM CMV exhibited slightly higher homology to HCMV than did RhCMV (Table 2). The exceptions were introns 1 and 2. In each region, however, the degree of difference in homology between AGM CMV and HCMV, compared to RhCMV and HCMV, was small ($<$4.5%).

The regions with the highest level of nucleic acid sequence conservation among all three representatives were exons 1 and 5 (Table 2) (one exception to this is discussed below). Exon 5 exhibited a dramatically higher level of sequence conservation than did exon 4 for both nucleic acid and amino acid sequences (Table 2) (the amino acid alignments are discussed below). This pattern of homology was noted in comparisons of both nonhuman primate CMV to each other (69.9 versus 54.3% for exons 5 and 4, respectively) and comparisons of both to HCMV [RhCMV: 58 (exon 5), 43.3% (exon 4); AGM CMV: 60.6 (exon 5), 44.6% (exon 4)]. In fact, exon 4 was the

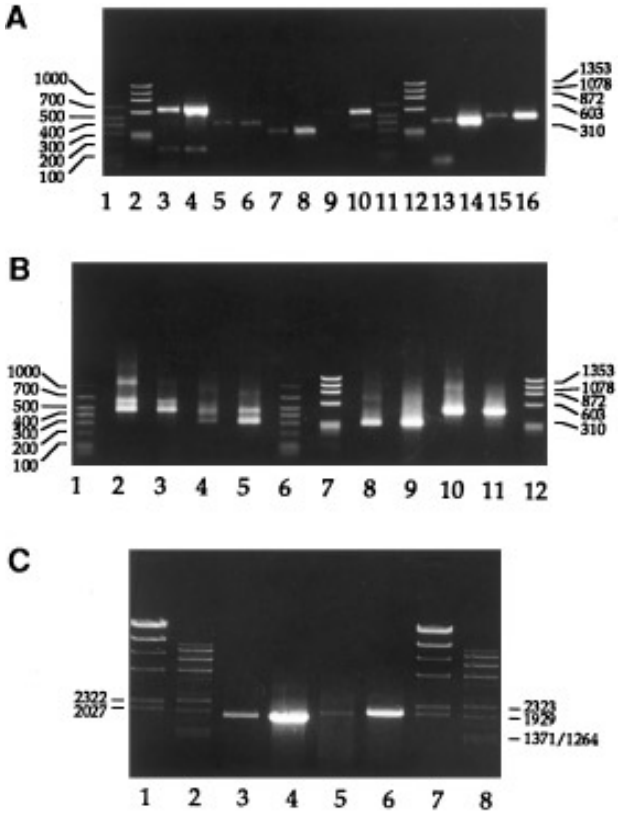

FIG. 2. Mapping of IE1 and IE2 transcripts. (A) The splice sites were mapped by RT/PCR (see Materials and Methods) using RNA from cells transfected with pRhIE 9.4-2 (lanes 3, 5, 7, 9, 13, 15) and from cells infected with RhCMV strain 68-1 (lanes 4, 6, 8, 10, 14, 16). Samples were amplified with primer pairs 57/53 (lanes 3, 4), 60/48 (lanes 5, 6), 60/170 (lanes 7, 8), 60/S3 (lanes 9, 10), 57/S2 (lanes 13, 14), and 175/S2 (lanes 15, 16). Primer sequences and locations are given in Table 1. Sizes for molecular weight markers XI (Promega) (lanes 1 and 11) and ϕX174 DNA cleaved with *Hae*III (lanes 2 and 12) are presented. (B) The polyadenylation sites were mapped by 3′ RACE (see Materials and Methods) using RNA from cells transfected with pRhIE 9.4-2 (lanes 2, 4, 8, 10) and cells infected with RhCMV strain 68-1 (lanes 3, 5, 9, 11). Samples were amplified with primer pairs 55/(dT)17 (lanes 2, 3), 143/(dT)17 (lanes 4, 5), and S7/(dT)17 (lanes 8, 9), 58/(dT)17 (lanes 10, 11). (C) IE1 and IE2 transcripts were amplified using RNA from cells transfected with pRhIE 9.4-2 (lanes 3 and 5) and cells infected with RhCMV strain 68-1 (lanes 4 and 6). Samples were amplified with primer pairs 175/176 (lanes 3, 4) and 175/S5 (lanes 5, 6). Sizes for λ DNA cleaved with *Hin*dIII (lanes 1 and 7) and *Bst*EII (lanes 2 and 8) are presented.

second-least conserved region between RhCMV and AGM CMV; only intron 2 sequences were less conserved (41.4%). In nonhuman primate CMV, introns 1 and 3 had higher nucleic acid homology (65.7 and 77.1%, respectively) than exon 4, as did exons 2 and 3 (68 and 59%, respectively). This was not the case when RhCMV and AGM CMV introns 1–3 and exon 2 were compared to HCMV; homologies ranged between 38 and 44%.

Some of the introns and exons were conserved in size among all three CMV, whereas others were not. For instance, exons 1 and 3 and intron 2 were close in size, but exons 2, 4, and 5 and introns 1 and 3 displayed some plasticity (Table 1). For instance, exon 2 ranged between 153 (AGM CMV) and 88 nt (HCMV) in length. Intron 3 was almost identical in size for AGM CMV and RhCMV (86 and 88 nt, respectively), while it was 170 nt for HCMV. There was no consistent pattern to which species had the largest or smallest exons or introns.

HCMV codes for an abundant transcript, expressed at late times after infection, which originates from a promoter within exon 5 (Puchtler and Stamminger, 1991; Stenberg *et al.,* 1989). Both RhCMV and AGM CMV (Chang *et al.,* 1995) contained identical promoter motifs (TATTAA) in exon 5. We determined by 5′ RACE that this promoter was active at late times of RhCMV infection (data not shown), although the precise start site of transcription has not been mapped. Recently, a novel spliced IE2 transcript was identified that resulted from previously unrecognized splice sites within exon 5 (Kerry *et al.,*

TABLE 2

Nucleic Acid Homology for IE Exons and Introns

CMV	Region	Length (nt)		Nucleic acid homology with RhCMV	Nucleic acid homology with HCMV
AGM	Exon 1	109		86	70.1
Rh		107			69.8
H		121			
AGM	Intron 1	895		65.7	41.6
Rh		967			42.9
H		827			
AGM	Exon 2	153		68	38.6
Rh		128			42.9
H		88			
AGM	Intron 2	113		41.4	37.8
Rh		111			39.8
H		114			
AGM	Exon 3	194		59	48.4
Rh		185			46.9
H		185			
AGM	Intron 3	86		77.1	44.2
Rh		88			39.8
H		170			
		Coding[a]	3′ Untrans.[b]		
AGM	Exon 4	1319	194/313	54.3	44.6
Rh		1373	249/388		43.3
H		1217	98		
AGM	Exon 5	1427	185	69.9	60.6
Rh		1562	215		58
H		1483	87		

[a] First base of exon to first base of stop codon.

[b] First base of stop codon to first base of polyadenylation sequence.

1995). We have been unable to detect by RT/PCR any evidence for such a splicing variant in our analysis using RNA from cells infected with 68-1 in the presence (lane 6, Fig. 2C) or absence of cycloheximide (data not shown). Although the nucleotide sequence for the splice donor region was strongly conserved with HCMV, there was no apparent splice acceptor at the corresponding site identified in HCMV (Kerry *et al.,* 1995). In addition, we have not detected any splice variants within exon 5 that corresponded to the 1.7-kb transcript of HCMV (Stenberg *et al.,* 1985). Using primers that would flank potential splicing signals within exon 5, there was no evidence of a spliced transcript removing a corresponding 434-nt intron (Stenberg *et al.,* 1985) from RhCMV (see lanes 5 and 6, Fig. 2C); only the predicted full-length transcripts were detected. There is no evidence for such a splicing event in AGM CMV (Chang *et al.,* 1995).

Identification of polyadenylation sites

The same sources of RNA were used to identify polyadenylation sites by 3′ RACE and sequence analysis. Using primer 55 and the $(dT)_{17}$-adapter primer (Frohman *et al.,* 1988), two PCR products of 612 and 473 nt were identified for both pRhIE 9.4-2 and 68-1 (lanes 2 and 3, Fig. 2B). A third band of greater than 1000 nt was also observed but was not analyzed further. Analysis with a different primer pair, $143/(dT)_{17}$ also resulted in two prominent bands of 459 and 320 nt (lanes 4 and 5, Fig. 2B). For each primer pair, the larger amplification product resulted from polyadenylation 15 nt downstream from a consensus polyadenylation signal (AATAAA, beginning at position +3327). Smaller amplification products were due to polyadenylation 18 nt downstream from a nonconsensus polyadenylation signal (AATATA, beginning at position +3185). A similar pattern of two polyadenylation signals has been described for AGM CMV IE1, including the use of the sequence AATATA as a polyadenylation motif (Chang *et al.,* 1995). Only a single polyadenylation site has been described for HCMV (Stenberg *et al.,* 1984). The predicted size for the RhCMV IE1 transcript (2042 nt for the shorter exon 4) was very close to the size for the IE1 transcript estimated from Northern blot analysis (2.0 kb) (Alcendor *et al.,* 1993). Only a single polyadenylation site was identified for the IE 2 transcript. Amplification of the 3′ end of the IE2 transcript with either primer pair $S7/(dT)_{17}$ or primer pair $58/(dT)_{17}$ yielded PCR products of 312 and 475 nt, respectively (lanes 8–11, Fig. 2B), resulting from polyadenylation 15 nt downstream from a consensus polyadenylation signal (AATAAA, beginning at +5162).

For both nonhuman primate CMV, the 3′ untranslated region for IE1 and IE2 were considerably longer than for HCMV (Table 1). The IE1 3′ untranslated regions for RhCMV and AGM CMV were 2–4 times longer than in

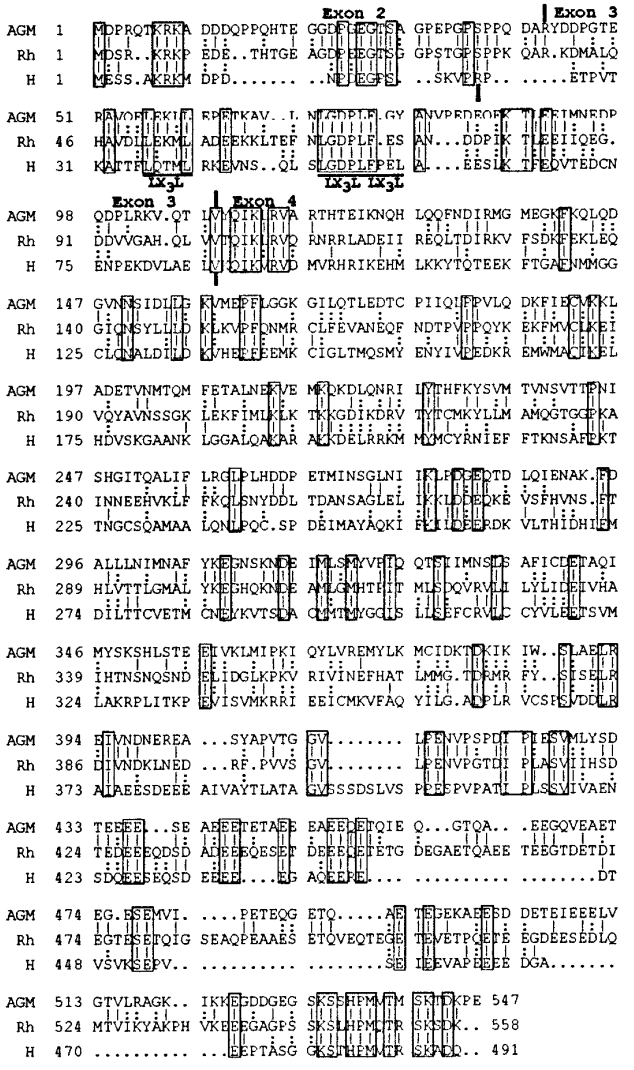

FIG. 3. Amino acid comparisons for IE1 of AGM CMV, RhCMV, and HCMV. The predicted amino acid sequences were aligned with the GCG Pileup and Gap programs. Amino acid identities are indicated by a vertical line, and regions of amino acid similarities are indicated by two vertical dots. Amino acids conserved between all three isolates are boxed. Potential regulatory features are underlined and exons are delineated by vertical lines.

HCMV (depending on polyadenylation site), and the 3′ untranslated region for IE2 was over 2 times longer in the nonhuman primate CMV.

Amino acid sequence comparisons

The amino acid alignments for the IE exons are presented in Figs. 3 (IE1) and 4 (IE2). RhCMV coded for the largest IE1 and IE2 proteins. For IE1, the RhCMV predicted protein was 558 amino acids (aa), compared to 547 and 491 aa for AGM CMV and HCMV, respectively. As with the nucleic acid comparisons (Table 3, Figs. 3 and 4), the two nonhuman primate IE1 and IE2 proteins shared a higher degree of amino acid identity and simi-

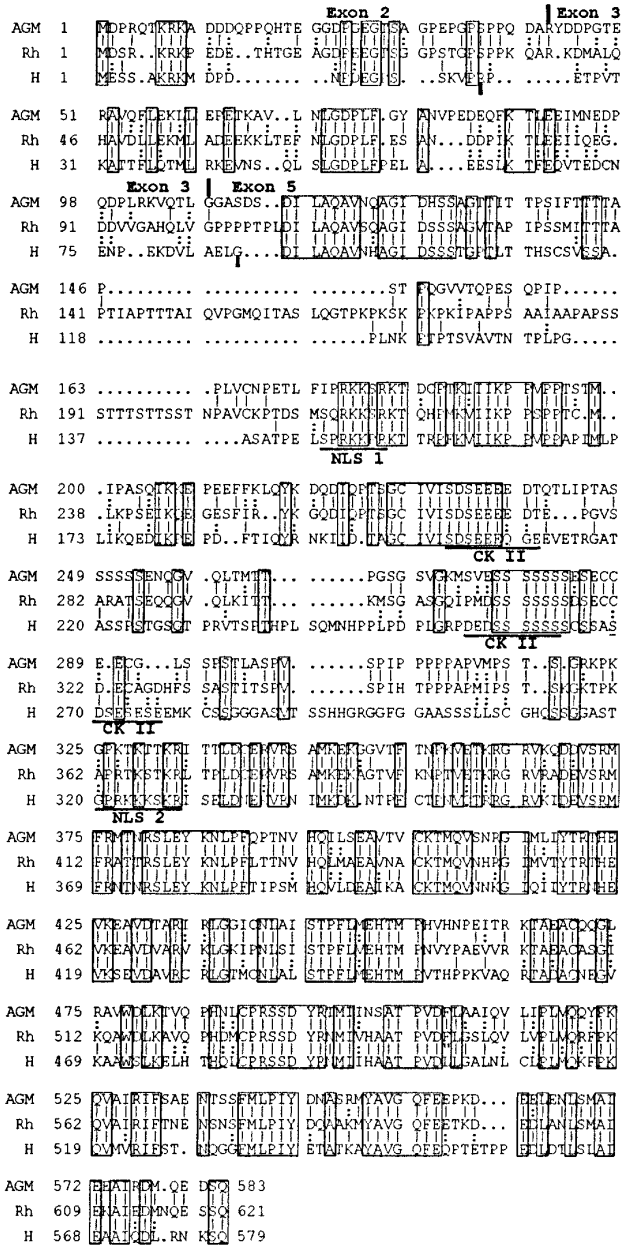

FIG. 4. Amino acid comparisons for IE2 of AGM CMV, RhCMV, and HCMV.

larity than either had with the HCMV protein. For instance, the IE1 proteins of RhCMV and AGM CMV were 40% identical and 61% similar (Table 3). Each nonhuman primate IE1 was almost equidistant from the HCMV IE1 protein; the RhCMV was 29% identical/52% similar with HCMV IE1, and the AGM CMV IE1 protein was 28% identical/49% similar with HCMV. The IE1 proteins were considerably less homologous than the IE2 proteins in every pairwise combination. While the two nonhuman IE1 proteins were 40% identical/61% similar, the IE2 proteins were 65% identical/77% similar (Table 3). Similarly, the RhCMV and HCMV IE2 proteins were 48% identical/64% similar, compared to 29% identical/52% similar for IE1. The same pattern of IE1 and IE2 homologies was true for AGM CMV compared to HCMV (Table 3).

The increased sizes of the two nonhuman primate CMV IE1 relative to HCMV were the result of additional amino acids at two sites. Most of the additional amino acids were very near to the carboxyl termini (exon 4) (Fig. 3), requiring the introduction of three large gaps in HCMV IE1 for optimal alignment with nonhuman primate IE1. The other region with additional bases was within exon 2, requiring the introduction of three gaps for alignments of HCMV to RhCMV or AGM CMV. Identical and conservative changes in the three IE1 proteins were uniformly distributed throughout the coding region (Fig. 3). Significantly, motifs thought to be important for IE1 function have been conserved. For instance, exon 3 of HCMV has been shown to contain a transcription activator domain of the IE2 protein (Malone *et al.,* 1990; Pizzorno *et al.,* 1991; Stenberg *et al.,* 1990); the activation capacity may be related to three domains with the structure of leucine-X_3-leucine (Stenberg *et al.,* 1990). Two of these domains have been conserved in RhCMV and AGM CMV (Fig. 3). Similarly, exon 4 of all three CMV were glutamate-rich with multiple polyglutamate stretches, particularly in the carboxyl one-third of the exon; these motifs may be important acidic activation domains of the IE1 protein (Stenberg *et al.,* 1990). Within this same region of the IE1 protein, many glutamate residues align with aspartate codons, substituting one negative charge for another. Perhaps significantly, within the three gaps of HCMV IE1 required for optimal alignments at the carboxyl termini, the RhCMV IE1 sequence contained 10, 6, and 5 additional negatively charged residues, compared to HCMV. The AGM CMV sequence contained 4, 3, and 4 additional negatively charged residues in these same gaps.

For IE2, the predicted RhCMV IE2 protein was 621 aa versus 583 aa for AGM CMV and 579 aa for HCMV (Table 3). The larger size of the RhCMV IE2 protein was due to two blocks of amino acids between positions 140–168 and 185–204 (RhCMV coordinates, Fig. 4). The second of these blocks contained a stretch of 12 contiguous serine or threonine residues. Serine and threonine residues constituted 19% of the RhCMV and HCMV IE2 proteins and 18.5% of AGM CMV IE2. Regions of identity/similarity between the three IE2 proteins exhibited a bipartite nature. The carboxyl half of IE2 was more conserved than the amino half (Fig. 4). For instance, the carboxyl half of RhCMV IE2 (aa 362–621) was 60% identical/75% similar with HCMV IE2 and 71% identical/83% similar with AGM CMV IE2. The amino half was 36% identical/58% similar with the amino half of HCMV IE2 and 59% identical/76% similar with AGM CMV. Multiple domains for the many functions of the HCMV IE2 protein have been mapped to the exon 5 portion. These functions include transactivation, autoregulation, dimerization, nu-

clear localization, protein interactions, and DNA binding (Chiou *et al.,* 1993; Macias and Stinski, 1993; Pizzorno *et al.,* 1991; Scully *et al.,* 1995). Other features in the amino half of the exon, postulated to be important for IE2 functions, have also been conserved. For instance, two nuclear localization signals identified in HCMV IE2 (Pizzorno *et al.,* 1991) have been retained in the nonhuman primate IE2 (Fig. 4). Similarly, two of three potential sites for phosphorylation by casein kinase II have been strongly conserved, while a third potential site in both RhCMV and AGM CMV has diverged considerably from HCMV (Fig. 4).

Sequence comparison of UL121-117

Translation of the RhCMV sequence (Fig. 1B) identified additional open reading frames downstream from the IE transcription region homologous to the UL121–117 ORF of HCMV. The UL121–117 region of HCMV encompasses five ORF of unknown function (Chee *et al.,* 1990). Nucleic acid sequences for these RhCMV ORF ranged in identity from 43.4% (UL119) to 54.9% (UL117) with the corresponding HCMV ORF (not shown). The cloned insert in pRhIE 9.4-2 terminated within the UL117 ORF. Predicted translation products of the ORF ranged between 29.1 (UL121) to 46.7% identical (UL120) and 52.1 (UL121) to 66.3% (UL120) similar with HCMV (Table 3).

The predicted translation products of RhCMV UL121 and 120 (RhUL121 and RhUL120) were similar in size to the cognate ORF of HCMV (180 amino acids for both UL121, and 202 versus 201 for UL120, respectively) (Table 3). The hydropathy plots of the predicted translation products of UL121 and 120 revealed conserved hydrophobic and hydrophilic domains. Although there has been considerable drift between the two UL121 proteins (29% identity overall, 36% carboxyl half only), the substitutions have preserved potential structural motifs. For the region between residues 91 and 156 of RhUL121, in particular, the pattern of hydrophobic and hydrophilic residues very closely matched that of HUL121 (Fig. 5). The amino acid identity was less in the amino half of the proteins, although there were smaller stretches of conserved hydrophobic/hydrophilic domains. The UL120 proteins were very conserved throughout the entire sequences, 47% identical/66% similar. The hydropathic plots were synchronous throughout almost the entire proteins; the most significant deviations were near the amino termini (Fig. 5). UL121 of both RhCMV and HCMV contained 1 potential site for N-linked glycosylation (NXS or NXT), although the sites were not conserved in location. RhUL120 contained 8 potential glycosylation sites and HUL120 contained 10 sites; 6 of these sites were conserved in location between the two viruses (Fig. 5).

Portions of the UL119, 118, and 117 ORF of HCMV form part of a larger spliced transcript (which includes UL116 and UL115) that is present at all times of infection (Leatham *et al.,* 1991; Rawlinson and Barrell, 1993). Although there has been no published analysis of any proteins encoded by this transcript, the predicted amino acid sequences are glycoprotein-like. The RhUL119 ORF was 238 aa long, compared to 142 aa for HCMV; regions of similarity were uniformly distributed, although two large gaps were required for optimal alignment. Overall, RhUL119 was 38% identical/52% similar with HUL119. When the hydropathic plots were compared with the two large alignment gaps included, the two proteins were very similar at both the amino and the carboxyl termini. For amino acid residues 141–224 (RhUL119), the transitions from hydrophobic to hydrophilic residues were coincident with those of HUL119, with two exceptions. HUL119 contains 4 potential N-glycosylation sites, compared to 11 for RhCMV; only 1 of these was conserved in location between the two viruses (Fig. 5). Five of the potential glycosylation sites within RhUL119 were within the 28-aa alignment gap between residues 106 and 133. Transcription of the HUL119-115 transcript initiates 22 nucleotides downstream of a TATA box (Leatham *et al.,* 1991). A TATA-like element (TTATAT) was located upstream of RhUL119, although it is not known if this is part of a functional promoter for UL119. RhUL119/118 contained potential splice donor and acceptor sequences such that translation of a RhUL119/118 translation would maintain the same reading frame as HUL119/118. An interesting feature of UL119 was the preponderance of serine (S) and threonine (T) residues; HUL119 consisted of 32% S+T and RhUL119 was 43% S+T. Most of these residues were within the amino half of the proteins.

UL118 of RhCMV and HCMV were of similar length, 216 and 209 aa, respectively (Fig. 5). The amino two-thirds of the ORF were about 46% similar, whereas the carboxyl third was 85% similar. Both the RhCMV and the HCMV UL118 contained eight potential N-linked glycosylation sites; four of these were conserved in location between both viruses (Fig. 5). The hydropathic plots for UL118 were discordant for most of the protein sequences, except for the region corresponding to amino acids 141–216 of RhUL118. What was especially significant was that a splice donor site identified in HUL118, placing the HUL118 ORF in frame with UL117 (Leatham *et al.,* 1991; Rawlinson and Barrell, 1993), corresponded to residue 179 of RhUL118. The high degree of similarity between RhUL118 and HUL118 downstream from the identified splice donor suggests that other potential splice donor sites downstream of the HUL118 ORF (Leatham *et al.,* 1991; Rawlinson and Barrell, 1993) may be utilized in other transcripts. Two additional potential splice donor sites have been identified downstream from HUL118 (Leatham *et al.,* 1991; Rawlinson and Barrell,

TABLE 3

Amino Acid Homology for UL123–117

			Amino Acid			
			with RhCMV		with HCMV	
CMV	Protein	Length (aa)	Identity	Similarity	Identity	Similarity
AGM	IE 1	547	40.4	61	27.9	49
Rh		558			29.3	51.8
H		491				
AGM	IE2	583	65.1	>6.8	49.2	64.5
Rh		621			48.2	63.5
H		579				
Rh	UL121	180			29.1	52.1
H		180				
Rh	UL120	202			46.7	66.3
H		201				
Rh	UL119	238			38.1	52.5
H		142				
Rh	UL118	216			37.2	60.7
H		209				
Rh	UL117[a]	[a]			43.7[a]	58.9[a]

[a] Sequence comparison determined only for that portion of RhUL117 within p9.4-2.

1993), although it was not determined whether these sites were actually used. The RhUL118 sequence has been strongly conserved within different RhCMV isolates. There has been only a single, silent nucleotide change from +7345 to +7635 (aa residues 70–178) between the isolate presented in this report [strain 22659 (Alcendor *et al.,* 1993)] and the 68-1 strain (Asher *et al.,* 1974) of RhCMV (data not shown).

The sequenced region of RhUL117 coded for 237 amino acids; however, the molecular clone for this region of the RhCMV genome ended at an *Xho*I restriction site within the middle of the UL117 ORF. HUL117 codes for 423 amino acids (Chee *et al.,* 1990). The UL119–115 transcript identified for HCMV (Leatham *et al.,* 1991) splices from the UL118 ORF into the UL117 ORF at a point downstream from the end of the RhCMV clone. However, alignment of the RhCMV and HCMV UL117 amino acid sequence revealed that the amino termini of the two UL117 ORF were very highly conserved (not shown). Of the first 113 amino acids of overlap, there was 64% identity and 71% similarity. The hydropathic plots were also very similar for much of this same region. The plots were most discordant carboxyl to aa residue 85.

DISCUSSION

In this report, we have analyzed a region of the RhCMV genome that spans the immediate-early gene, the UL121–118 ORF, as well as part of UL117. This work has demonstrated that several key aspects of sequence, transcription, structural motifs, and genome organization for the UL123–117 region have been strongly conserved between the RhCMV and the HCMV genomes. Since UL121–118 have also been identified in the AGM CMV genome (G. Hayward, personal communication), these ORF appear to have been conserved in sequence and colinearity during speciation of the primate family of CMV. We have recently determined that most, if not all, of the long unique region of HCMV is colinear with the RhCMV genome (R. Kravitz, P. Luciw, P. Barry, unpublished data). The lineages for the two natural hosts of RhCMV and AGM CMV (*Macaca mulatta* and *Cercopithecus aethiops,* respectively) diverged approximately 8 million years ago (Cronin *et al.,* 1980); descendants of the *Macaca* lineage migrated from Africa to Asia and Europe, while the African green monkeys (*Cercopithecus*) remained in Africa. Comparisons of these two nonhuman primate CMV with each other and with HCMV provide a basis for a better understanding of the evolution of the primate family of CMV and for identifying sites within ORF critical for function.

A significant finding of this report was that the IE1 protein was less conserved than the IE2 proteins and less conserved than UL121–118 ORF (Table 2). This was true when just the simian CMV IE1/2 sequences were compared to each other or when either was compared

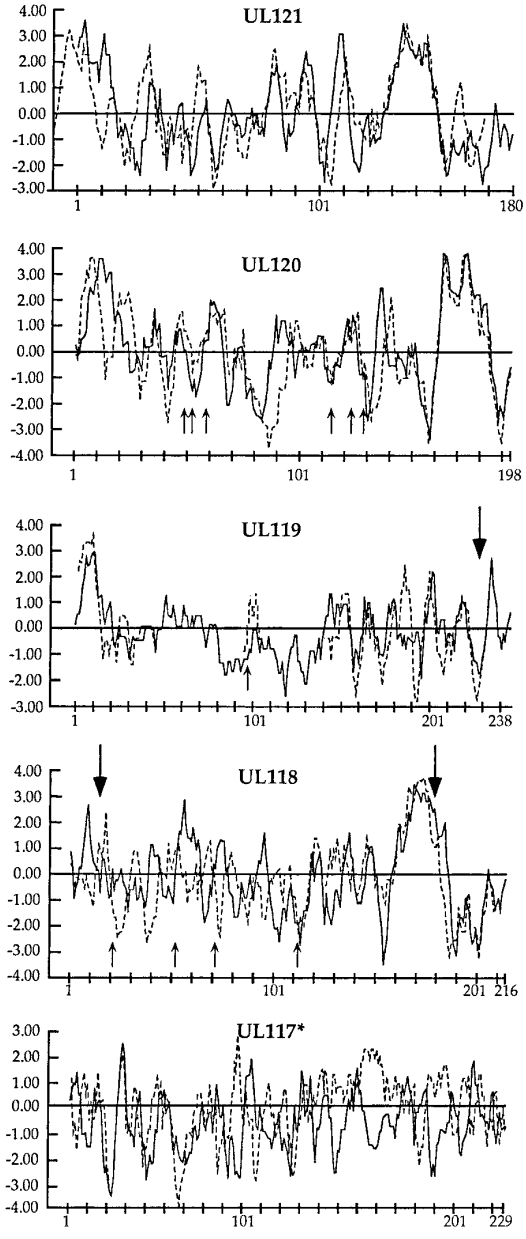

FIG. 5. Hydropathy plots for UL121–117 of RhCMV and HCMV. Kyte and Doolittle hydropathy plots were calculated using the MacDNAsis program (V. 3.5; Hitachi, San Bruno, CA). Values, averaged over a window of six residues (vertical axis), were plotted against RhCMV residue number (RhCMV, solid lines; HCMV, dashed lines). Positive values represent hydrophobic regions; negative values represent hydrophilic regions. Large arrows in UL119 and 118 indicate splice sites identified in HCMV (Leatham *et al.,* 1991; Rawlinson and Barrell, 1993). Small arrows represent conserved, potential N-linked glycosylation sites. The HUL119 plot was split to represent the two large gaps required for optimal alignment. The cloned fragment within p9.4-2 terminated within the RhUL117 sequence, upstream of the identified splice site of HUL117 (Leatham *et al.,* 1991; Rawlinson and Barrell, 1993).

to HCMV. The magnitude of IE1 evolution was highlighted by the fact that the IE1 proteins of RhCMV and HCMV have been less conserved than the predicted ORF of UL121–118 (Table 3). Similarly, the IE1 proteins of AGM CMV and RhCMV have been less conserved than UL121–118 (data not shown; G. Hayward, personal communication). The relatively high degree of IE1 variation would appear to be inconsistent with the fact that the IE1 protein plays an essential role in viral replication via transactivation of the major IE promoter (Cherrington and Mocarski, 1989; Sambucetti *et al.,* 1989) and β and γ promoters (Klucher *et al.,* 1993; Malone *et al.,* 1990; Stenberg *et al.,* 1990). Since HCMV IE1 transactivation of the major IE promoter/enhancer involves the NF-κB transcription factor (Sambucetti *et al.,* 1989), it might be expected that there would be tight constraints on IE1 evolution.

Sequence evolution of IE1 may be due, in part, to immunological pressures. The extremely abundant IE1 protein accumulates within the nucleus of infected cells throughout the cell cycle (Otto *et al.,* 1988). In both humans (Borysiewicz *et al.,* 1988) and mice (Koszinowski *et al.,* 1987), cytotoxic T-lymphocytes recognize IE1 epitopes. Part of the drive in IE1 variation during primate CMV speciation may have been the result of evasion from immune detection.

The utility of infection of rhesus macaques with RhCMV as a model for the study of HCMV biology and regulation requires close similarities between the biology of these viruses in their natural hosts and close conservation of genomic sequence, gene structure, and function. Previous publications have demonstrated that the biology of RhCMV in rhesus macaques is very similar to that of HCMV in humans (Asher *et al.,* 1974; Baskin, 1987; London *et al.,* 1986; Osborn *et al.,* 1984; Vogel *et al.,* 1994). This report and a previous one from our group (Alcendor *et al.,* 1993) establish that selective pressures during the speciation of RhCMV, AGM CMV, and HCMV have contributed to conservation of nucleic and amino acid sequences and genomic organization. Nonhuman primate CMV models should prove to be extremely useful for a better understanding of HCMV biology, gene regulation, and pathogenesis.

ACKNOWLEDGMENTS

These studies were supported in part by a grant to P.A.L. from the National Heart, Lung, and Blood Institute (HL43609); a National Research Service Award to P.A.B. (AI07893); and the base grant to the California Regional Primate Research Center from the National Institutes of Health (RR00169).

REFERENCES

Alcendor, D. J., Barry, P. A., Pratt-Lowe, E., and Luciw, P. A. (1993). Analysis of the rhesus cytomegalovirus immediate-early gene promoter. *Virology* 194, 815–821.

Asher, D. M., Gibbs, J., Lang, D. J., and Gadjusek, D. C. (1974). Persistent shedding of cytomegalovirus in the urine of healthy rhesus monkeys. *Proc. Soc. Exp. Biol. Med.* 145, 794–801.

Baskin, G. B. (1987). Disseminated cytomegalovirus infection in immunodeficient rhesus macaques. *Am. J. Pathol.* 129, 345–352.

Borysiewicz, L. K., Hickling, J. K., Graham, S., Sinclair, J., Cranage, M. P., Smith, G. L., and Sissons, J. G. P. (1988). Human cytomegalovirus-specific cytotoxic T cells. Relative frequency of stage-specific CTL recognizing the 72-kD immediate early protein and glycoprotein B expressed by recombinant vaccinia viruses. *J. Exp. Med.* 168, 919–931.

Chang, Y.-N., Jeang, K.-T., Lietman, T., and Hayward, G. S. (1995). Structural organization of the spliced immediate-early gene complex that encodes the major acidic nuclear (IE1) and transactivator (IE2) proteins of African green monkey cytomegalovirus. *J. Biomed. Sci.* 2, 105–130.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., Horsnell, T., Hutchison, C. A., Kouzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, K. M., and Barrell, B. G. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* 154, 125–169.

Cherrington, J. M., and Mocarski, E. S. (1989). Human cytomegalovirus ie1 transactivates the α-promoter-enhancer via an 18-base-pair repeat element. *J. Virol.* 63, 1435–1440.

Chiou, C.-J., Zong, J., Waheed, I., and Hayward, G. S. (1993). Identification and mapping of dimerization and DNA-binding domains in the C terminus of the IE2 regulatory protein of human cytomegalovirus. *J. Virol.* 67, 6201–6214.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* 162, 156–159.

Covell, W. (1932). Occurrence of intranuclear inclusions of monkeys unaccompanied by specific signs of disease. *Am. J. Pathol.* 8, 151–158.

Cowdry, W., and Scott, G. (1935). Nuclear inclusions in kidneys of *Macacus rhesus* monkeys. *Am. J. Pathol.* 11, 659–668.

Cronin, J. E., Cann, R., and Sarich, V. M. (1980). Molecular evolution in the genus *Macaca*. *In* "The Macaques. Studies in Ecology, Behavior and Evolution" (D. G. Lindberg, Ed.). Van Nostrand–Reinhold, New York.

Eizuru, Y., Tsuchiya, K., Mori, R., and Minamishima, Y. (1989). Immunological and molecular comparisons of simian cytomegaloviruses isolated from African green monkey (*Ceropithicus aethiops*) and Japanese macaque (*Macaca fuscata*). *Arch. Virol.* 107, 65–75.

Frohman, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85, 8998–9002.

Gibson, W. (1983). Protein counterparts of human and simian cytomegaloviruses. *Virology* 128, 391–406.

Hendrickx, A. G., and Binkerd, P. E. (1990). Nonhuman primates and teratological research. *J. Med. Primatol.* 19, 81–108.

Ho, M. (1991). "Cytomegalovirus: Biology and Infection." 2nd ed. Plenum, New York.

Jeang, K.-T., and Gibson, W. (1980). A cycloheximide-enhanced protein in cytomegalovirus-infected cells. *Virology* 107, 362–374.

Jeang, K. T., Chin, G., and Hayward, G. S. (1982). Characterization of cytomegalovirus immediate-early genes. I. Nonpermissive rodent cells overproduce the IE94K protein from CMV (Colburn). *Virology* 121, 393–403.

Jeang, K.-T., Cho, M.-S., and Hayward, G. S. (1984). Abundant constitutive expression of the immediate-early 94K protein from cytomegalovirus (Colburn) in a DNA-transfected mouse cell line. *Mol. Cell. Biol.* 4, 2214–2223.

Jeang, K.-T., Rawlins, D. R., Rosenfeld, P. J., Shero, J. D., Kelly, T. J., and Hayward, G. S. (1987). Multiple tandemly repeated binding sites for cellular nuclear factor 1 that surround the major immediate-early promoters of simian and human cytomegalovirus. *J. Virol.* 61, 1559–1570.

Kalter, S. S., and Heberling, R. L. (1990). Primate viral diseases in perspective. *J. Med. Primatol.* 19, 519–535.

Kerry, J. A., Sehgal, A., Barlow, S. W., Cavanaugh, V. J., Fish, K., Nelson, J. A., and Stenberg, R. M. (1995). Isolation and characterization of a low-abundance splice variant from the human cytomegalovirus major immediate-early gene region. *J. Virol.* 69, 3868–3872.

Kessler, M. J., London, W. T., Madden, D. L., Dambrosia, J. M., Hilliard, J. K., Soike, K. F., and Rawlins, R. G. (1989). Serological survey for viral diseases in the Cayo Santiago rhesus macaque population. *Puerto Rican Health Sci. J.* 8, 95–97.

Klucher, K. M., Sommer, M., Kadonaga, J. T., and Spector, D. H. (1993). In vivo and in vitro analysis of transcriptional activation mediated by the human cytomegalovirus major immediate-early proteins. *Mol. Cell. Biol.* 13, 1238–1250.

Koszinowski, U. H., Keil, G. M., Schwarz, H., Schickedanz, J., and Reddehase, M. J. (1987). A nonstructural polypeptide encoded by immediate-early transcription unit 1 of murine cytomegalovirus is recognized by cytolytic T lymphocytes. *J. Exp. Med.* 166, 289–294.

Leatham, M. P., Witte, P. R., and Stinski, M. F. (1991). Alternate promoter selection within a human cytomegalovirus immediate-early and early transcription unit (UL119-115) defines true late transcripts containing open reading frames for putative viral glycoproteins. *J. Virol.* 65, 6144–6153.

London, W. T., Martinez, A. J., Houff, S. A., Wallen, W. C., Curfman, B. L., Traub, R. G., and Sever, J. L. (1986). Experimental congenital disease with simian cytomegalovirus in rhesus monkeys. *Teratology* 33, 323–331.

Mach, M., Stamminger, T., and Jahn, G. (1989). Human cytomegalovirus: Recent aspects from molecular biology. *J. Gen. Virol.* 70, 3117–3146.

Macias, M. P., and Stinski, M. F. (1993). An in vitro system for human cytomegalovirus immediate early 2 protein (IE2)-mediated site-dependent repression of transcription and direct binding of IE2 to the major immediate early promoter. *Proc. Natl. Acad. Sci. USA* 90, 707–711.

Malone, C. L., Vesole, D. H., and Stinski, M. F. (1990). Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: Mutational analysis of the viral proteins. *J. Virol.* 64, 1498–1506.

Osborn, K. G., Prahalada, S., Lowenstine, L. J., Gardner, M. B., Maul, D. H., and Henrickson, R. V. (1984). The pathology of an epizootic of acquired immunodeficiency in rhesus macaques. *Am. J. Pathol.* 114, 94–103.

Otto, S. M., Sullivan-Tailyour, G., Malone, C. L., and Stinski, M. F. (1988). Subcellular localization of the major immediate early protein (IE1) of human cytomegalovirus at early times after infection. *Virology* 162, 478–482.

Pizzorno, M. C., Mullen, M. A., Chang, Y. N., and Hayward, G. S. (1991). The functionally active IE2 immediate-early regulatory protein of human cytomegalovirus is an 80-kilodalton polypeptide that contains two distinct activator domains and a duplicated nuclear localization signal. *J. Virol.* 65, 3839–3852.

Puchtler, E., and Stamminger, T. (1991). An inducible promoter mediates abundant expression from the immediate-early 2 gene region of human cytomegalovirus at late times after expression. *J. Virol.* 65, 6301–6306.

Rawlinson, W. D., and Barrell, B. G. (1993). Spliced transcripts of human cytomegalovirus. *J. Virol.* 67, 5502–5512.

Sambucetti, L. C., Cherrington, J. M., Wilkinson, G. W., and Mocarski, E. S. (1989). NF-kappa B activation of the cytomegalovirus enhancer is mediated by a viral transactivator and by T cell stimulation. *EMBO J.* 8, 4251–4258.

Sanger, F., Nicklen, S., and Coulsen, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.

Scully, A. L., Sommer, M. H., Schwartz, R., and Spector, D. H. (1995). The human cytomegalovirus IE2 86-kilodalton protein interacts with an early gene promoter via site-specific DNA binding and protein–protein associations. *J. Virol.* 69, 6533–6540.

Smith, K. O., Thiel, J. F., Newman, J. T., Harvey, E., Trousdale, M. D., Gehle, W. D., and Clark, G. (1969). Cytomegalovirus as common adventitious contaminants in primary African green monkey kidney cell cultures. *J. Natl. Cancer Inst.* 42, 489–496.

Stenberg, R. M., Depto, A. S., Fortney, J., and Nelson, J. A. (1989). Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. *J. Virol.* 63, 2699–2708.

Stenberg, R. M., Fortney, J., Barlow, S. W., Magrane, B. P., Nelson, J. A., and Ghazal, P. (1990). Promoter-specific transactivation and repression by human cytomegalovirus immediate-early proteins involves common and unique protein domains. *J. Virol.* 64, 1556–1565.

Stenberg, R. M., Thomsen, D. R., and Stinski, M. F. (1984). Structural analysis of the major immediate early gene of human cytomegalovirus. *J. Virol.* 49, 190–199.

Stenberg, R. M., Witte, P. R., and Stinski, M. F. (1985). Multiple spliced and unspliced transcripts from human cytomegalovirus immediate-early region 2 and evidence for a common initiation site within immediate early region 1. *J. Virol.* 56, 665–675.

Swack, N. S., Liu, O. C., and Hsiung, G. D. (1971). Cytomegalovirus infections of monkeys and baboons. *Am. J. Epidemiol.* 94, 397–402.

Vogel, P., Weigler, B. J., Kerr, H., Hendrickx, A., and Barry, P. A. (1994). Sero-epidemiology of rhesus cytomegalovirus in a population of captive macaques. *Lab. Anim. Sci.* 44, 25–30.

Weiner, D., and Gibson, W. (1981). Identification of a primate cytomegalovirus group-common protein antigen. *Virology* 115, 182–191.