JOURNAL OF CLINICAL MICROBIOLOGY, Sept. 1992, p. 2307–2310
0095-1137/92/092307-04$02.00/0

Vol. 30, No. 9

Effect of Interstrain Variation on Diagnostic DNA Amplification of the Cytomegalovirus Major Immediate-Early Gene Region

SUNWEN CHOU

Medical and Research Services, VA Medical Center, and Division of Infectious Diseases, Oregon Health Sciences University, Portland, Oregon 97201

Received 6 April 1992/Accepted 22 June 1992

The immediate-early region exon 4 sequences of six clinical cytomegalovirus strains were determined and compared with those of laboratory strains AD169 and Towne. Of 407 codons in exon 4, 33 (8.1%) showed interstrain variation at the peptide level and 74 (18%) showed interstrain variation at the nucleotide level. Variation occurred sporadically throughout the exon, and no grouping of strains was apparent. Published oligonucleotide primers proposed for diagnostic detection of cytomegalovirus by polymerase chain reaction have often been based on exon 4 sequences. Some of these primers show sequence mismatches with strains sequenced here. Amplification sensitivity for mismatched strains was reduced up to 100-fold. More-uniform detection sensitivity was achieved with primers of conserved sequence.

Amplification of cytomegalovirus (CMV) sequences from clinical specimens is widely performed for rapid diagnosis as well as for exploration of pathogenesis. Despite numerous reports of successful use of the polymerase chain reaction (PCR) for this purpose, there has been no consensus as to the optimal primers and targets for diagnostic DNA amplification. The large size of the CMV genome (230 kb) offers an effectively unlimited choice of primers. There is, however, sequence variation among CMV strains which affects every region of the genome. Ideally, primer selection should be based on knowledge of interstrain variability in candidate primer sequences and the consequences of mismatching at primer sites. In published work to date, diagnostic PCR primers have frequently been derived from locations within the CMV major immediate-early (MIE) region exon 4 (2–4, 8–11, 14, 16). This exon is 1,221 bp (407 codons) in length and has been assumed to be well conserved on the basis of its crucial function in CMV gene regulation (13) and the almost identical published exon 4 sequences from laboratory strains AD169 and Towne (1, 15).

In this report, MIE exon 4 sequences for six additional clinical strains of CMV are presented. Conserved and variable codons of exon 4 were mapped, and the performance of published and unpublished exon 4 primers was tested against variant strains in an effort to test the effect of sequence mismatch within the primers used for PCR.

MATERIALS AND METHODS

CMV strains. Six genetically distinct clinical CMV strains isolated from urine or blood specimens of kidney, heart, or liver recipients were selected for sequencing on the basis diversity of restriction enzyme digest patterns of PCR products from amplification of the MIE gene region (5). Four of the strains have also been sequenced in the glycoprotein B and H regions (6, 7). Except for strains C325 and C327, which were plaque purified directly from the original urine specimen, strains were not plaque purified.

DNA sequencing. The sequence of the MIE exon 4 was determined by dideoxy chain termination reactions on unpurified PCR-amplified biotinylated templates attached to magnetic streptavidin-coated particles as previously described (7). Templates varied in length from 0.4 to 2.7 kb. Overlapping partial sequences of exon 4 were determined by using synthetic oligonucleotide primers spaced 300 or fewer bases apart and representing both DNA strands of the region sequenced (Table 1). New sequence information was aligned with existing data (including reference strains AD169 and Towne), and any ambiguities in alignment or sequence variation were rechecked with additional sequencing reactions.

Standard PCR templates. For evaluation of primer performance, templates of known sequence and concentration were prepared by PCR amplification of a 0.8-kb target (primers MIE2444 and MIE4Z) from exon 4 (Fig. 1). Four CMV strains which had been shown to have sequence diversity in this region were selected. In each case, the amplified product showed a single ethidium-stained band on polyacrylamide gel electrophoresis and was purified by removal of primers and small molecules by ultrafiltration (Ultrafree-MC; 10,000-molecular-weight cutoff; Millipore). The DNA concentration was measured by determining the A_{260}, and standard dilutions were made of each template, containing 10 fg, 1 fg, 100 ag, and 10 ag of template per μl. The diluent was 10 mM Tris (pH 8.0)–1 mM EDTA containing 1 μg of salmon sperm DNA per ml as a carrier. These dilutions were used to test the detection sensitivity of matched and mismatched diagnostic PCR primers.

Comparative sensitivity of diagnostic PCR primers. Detection of exon 4 sequences was attempted with several primers, including ones previously described as well as those newly selected for this study. The forward (coding-strand) primers were MIE2783 and IE-1, and the reverse (noncoding)-strand primers were MIE3114 and MIE-5 (Table 1 and Fig. 1). MIE2783 and MIE3114 completely match the sequenced strains, whereas IE-1 (14) and MIE-5 (8) have various degrees of sequence mismatch. Each of the four combinations of forward and reverse primers was tested. Each primer pair was used to amplify 4 dilutions of the standardized templates defined above. PCR amplifications were done in 20-μl volumes in buffer consisting of 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.05% Tween 20,

TABLE 1. Primers for diagnostic PCR and sequencing

Primer	Sequence[a]	Strand[b]
MIE544[c]	TTTTGACCTCCATAGAAGAC	F
MIE2024	CAGATTAAGGTTCGAGTG	F
MIE4A	AGCATGTATGAGAACTAC	F
MIE2444	TGCAGTCAGGCCATGGCGGC	F
MIE2783	CGCCGCATTGAGGAGATCTGC	F
IE-1	CCACCCGTGGTGCCAGCTCC	F
MIE2977	GTCCCCTGTACCCGCGACTA	F
MIE2200	AATACACTTCATCTCCTC	R
MIE48	GCTGAGTTCTTGGTA	R
MIE42	GCGTGAGCACCTTGTCTC	R
MIE3114[c]	GACTTGACAGACACAGTG	R
MIE-5	CAGCACCATCCTCCTCTTCCTCTGG	R
MIE4Z	ACTGGTCAGCCTTGCTTC	R
MIE4R[c]	ATAGTGACGTGGGATCCAT	R

[a] All sequences are shown in the 5′-to-3′ direction.
[b] F, coding strand; R, noncoding strand.
[c] Biotinylated for PCR template production.

and 0.2 mM (each) dATP, dCTP, dGTP, and TTP. Each reaction mixture contained 1 ng of each primer per μl, 0.3 μl of the selected template dilution, and 0.5 U of *Taq* polymerase. Thermal cycling was done in an automated cycler (model 50; Coy Labs, Ann Arbor, Mich.) at 94.5°C for 1.5 min, 55°C for 2 min, and 72°C for 1 min for 35 cycles. The PCR product (2 μl) was visualized after electrophoresis through a 4% polyacrylamide gel and staining with ethidium bromide. The endpoint of detection sensitivity was determined from the lowest concentration of standard template which was amplified to an ethidium-stained product band of the appropriate size and clearly discernible above the background. Negative controls without added template were included to check for contamination of primers and reagents.

Nucleotide sequence accession numbers. The GenBank accession numbers for the new sequences are M95634 through M95639.

RESULTS

Interstrain variation in MIE exon 4. Sequence variation was assessed in the six sequenced clinical strains and laboratory strains AD169 (1) and Towne (15). Among the 407 codons in exon 4, 33 (8.2%) varied at the peptide level and 74 (18.3%) varied at the nucleotide level. This is comparable to CMV envelope glycoproteins gp55 (part of gB) and gH, in which variation is approximately 5% at the peptide level and 21% at the nucleotide level (6, 7). In the MIE region, variant codons were scattered throughout exon 4 (Fig. 1) and not strongly clustered or grouped in the manner previously described for gB and gH (6, 7). There were stretches of 10 to 20 codons which showed base conservation in all strains sequenced, located mostly near the beginning of the exon and at its 3′ one-third portion. Several distinctive structural features of the encoded peptide, which are also noted in other transcription factors, are preserved among all strains, including a zinc finger motif (codons 180 to 200) and serine-rich and polyglutamic acid tracts (one strain deleted one of 5 contiguous glutamic acid residues at codon 348). Amino acid changes at variant codons were mostly conservative and commonly affected more than one strain of those sequenced.

Diagnostic sensitivity of matched and mismatched PCR primers. On the basis of the comparative sequence data obtained in this study, previously published diagnostic PCR primers derived from exon 4 (2–4, 8–11, 14, 16) were noted to have various degrees of mismatch with the sequenced CMV strains, ranging from 0 to 5 bases. Two of these were selected to examine the effect of variation on amplification sensitivity (Table 2). The first, MIE-5 (8), is from the noncoding strand and is mismatched at the 3′-terminal base with five strains and an additional base in one strain. The second, IE-1 (14), is from the coding strand and has at least a 2-base mismatch with every strain because of an error in the published Towne strain sequence (codon 307) on which the primer was based. This codon was reported (15) as CGT (numbered 392), but resequencing of a sample of Towne strain during this study revealed it to be GCT, in concordance with all other strains. Additionally, IE-1 has mismatches with various strains at 3, 6, and 7 bases from its 3′ end (1 base per strain; Table 2). Primers MIE2783 and MIE3114, which match all strains sequenced, were used for comparison.

Short targets were amplified cleanly by the four diagnostic primer pairs to yield one predominant fragment of the expected mobility on polyacrylamide gel electrophoresis (Fig. 2), regardless of whether the CMV strain had a completely matching sequence. The diagnostic sensitivities of various primer pairs are shown in Table 2 and compared with the extent of mismatch within the primer sequences. Matched primers detected 3 ag (approximately 4 copies) of target sequence in all four strains, as did the primer pair MIE2783–MIE-5, where MIE-5 is mismatched with the 3′-terminal base in two strains (C005 and C076). Primer IE-1, which has the most extensive and diverse mismatching, was consistently 1 to 2 orders of magnitude less sensitive, even though amplification should be favored by the shorter target molecule. However, with the primer pair IE-1–MIE-5, the least mismatched strain (Towne, 2 bases) was amplified with equal sensitivity (3 ag) as with matched primers.

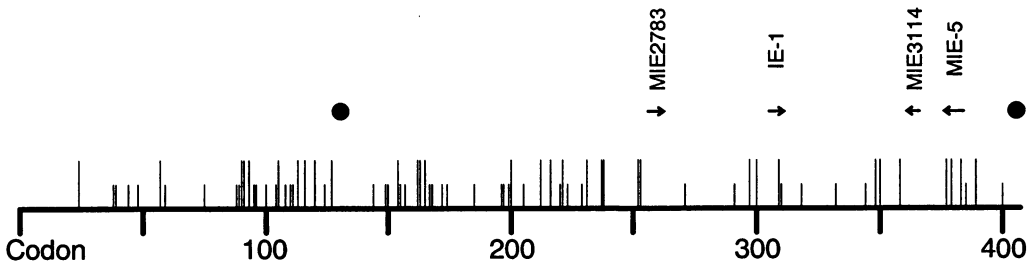

FIG. 1. Map of variant codons in exon 4 of the CMV MIE region (eight strains). Variation is at the peptide level (longer vertical lines) or at the nucleotide level only (shorter lines). Proposed diagnostic PCR primers are indicated by arrows. A target DNA sequence between the two closed circles was amplified from several strains for use as standard templates.

TABLE 2. Diagnostic amplification with matched and mismatched PCR primers

Strain	Sequence at primer site[a]		Minimum template (ag) detected with primer pairs			
	IE-1	MIE-5	MIE2783-MIE3114	IE-1-MIE3114	IE-1-MIE-5	MIE2783-MIE-5
C005	GCTGGTGCTAGCTCC	CTCCTCTTTCTCTGA	3	300	30	3
C076	GCTGGTGCCAGTTCC	CTCCTCTTCCTCTGA	3	300	30	3
C145	GCTGGTGTCAGCTCC	CTCCTCTTCCTCTGG	3	30	30	3
Towne	GCTGGTGCCAGCTCC	CTCCTCTTCCTCTGG	3	30	3	3

[a] Fifteen bases corresponding to the 3' end of the primer sequence. Variation from the primer sequence is shown in boldface type.

DISCUSSION

The main goals of this study were to provide a data base of MIE exon 4 sequences useful in designing diagnostic PCR primers for CMV and to illustrate the sensitivity differences that might result from interstrain variation within primer sequences. It was found that despite the fortuitous >99% sequence homology in published sequences of laboratory strains AD169 and Towne, the extent of nucleotide variation in this exon is comparable to that seen in the coding sequences of envelope glycoproteins gB and gH. Selection of primers from the more conserved portions of exon 4 should enable the amplification of a broad range of CMV isolates.

Compared with the highly clustered variation seen in the gB and gH, resulting in the clear delineation of group relationships among strains, variation in exon 4 appears more diffuse and sporadic. Most of the changes occur in isolated codons rather than in several contiguous ones. They are probably the result of point mutations that have accumulated over time, and these mutations appear to affect peptide encoding more often in exon 4 than in gB or gH. The sequencing of more clinical strains will likely reveal additional mutations.

To date, many published diagnostic PCR primers have been selected from MIE exon 4, apparently without the benefit of much comparative sequence data (2–4, 8–11, 14, 16). Some of the primers match the sequence of all the strains studied here, and in others, there are various degrees of mismatch with some strains. As the results show, the effect of mismatch at primer sequences is variable and can be subtle, even when there appears to be obvious mismatching. For example, primer MIE-5 performs as well as a nearby matched primer despite mismatching of the 3'-terminal base in most strains. It has been previously reported that a 3'-terminal G-T mismatch, as was the case here, is not detrimental to primer performance, whereas some other mismatches seriously affect sensitivity (12). Primer MIE-5 is also slightly longer than other primers studied; greater primer length may facilitate annealing in the presence of base mismatching. On the other hand, primer IE-1 lost 1 to 2 orders of magnitude of sensitivity (depending on the strain) but was still able to amplify three different strains each mismatched at three positions. The very high degree of sensitivity achieved with matched primers (4 copies detected) must be considered idealized, because little extraneous DNA is present in the standard templates. Sensitivity will likely be less when amplifying from cell extracts.

Although primer mismatching does not necessarily prevent PCR amplification, use of primers having a known high degree of sequence conservation, such as those proposed here and elsewhere (6, 7), should reduce assay variability in clinical practice. For routine diagnostic use, there is no special reason for favoring primers from MIE exon 4. In this and other regions, conserved sequences can be used as amplification primers, whereas the more variable sequences allow the differentiation of CMV strains. Amplification of target sequences, followed by restriction or sequence analysis, is valuable in authenticating signals from clinical specimens.

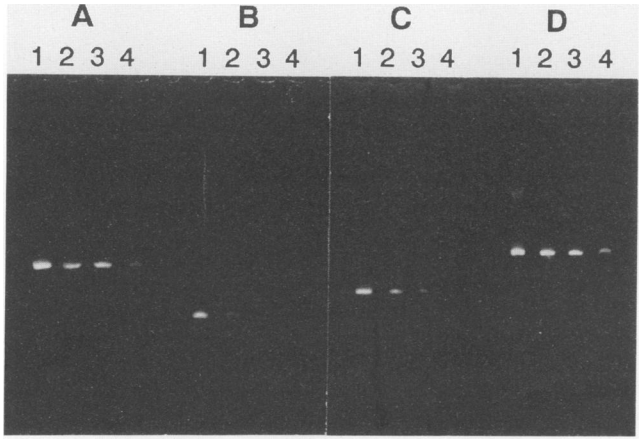

FIG. 2. PCR amplification of MIE targets from strain C076. Standard dilutions of template were 3 fg (lanes 1), 300 ag (lanes 2), 30 ag (lanes 3), and 3 ag (lanes 4). Primer pairs were MIE2783-MIE3114 (A), IE-1-MIE3114 (B), IE-1-MIE-5 (C), and MIE2783-IE-5 (D). The amplified product was analyzed by polyacrylamide gel electrophoresis and ethidium bromide staining.

ACKNOWLEDGMENTS

This work was supported by Department of Veterans Affairs research funds.

Marcia Kennedy, Taylor Dunn, and Gail Marousek provided technical assistance.

REFERENCES

1. **Akrigg, A., G. W. Wilkinson, and J. D. Oram.** 1985. The structure of the major immediate early gene of human cytomegalovirus strain AD-169. Virus Res. **2**:107–121.
2. **Buffone, G. J., G. J. Demmler, C. M. Schimbor, and J. Greer.** 1991. Improved amplification of cytomegalovirus DNA from urine after purification of DNA with glass beads. Clin. Chem. **37**:1945–1949.
3. **Burgart, L. J., M. J. Heller, M. J. Reznicek, T. C. Greiner, C. J. Teneyck, and R. A. Robinson.** 1991. Cytomegalovirus detection in bone marrow transplant patients with idiopathic pneumonitis. A clinicopathologic study of the clinical utility of the polymerase chain reaction on open lung biopsy specimen tissue. Am. J. Clin. Pathol. **96**:572–576.
4. **Cassol, S. A., M. C. Poon, R. Pal, M. J. Naylor, J. Culver-James, T. J. Bowen, J. A. Russell, S. A. Krawetz, R. T. Pon, and D. I. Hoar.** 1989. Primer-mediated enzymatic amplification of cyto-

megalovirus (CMV) DNA. Application to the early diagnosis of CMV infection in marrow transplant recipients. J. Clin. Invest. **83:**1109–1115.

5. **Chou, S.** 1990. Differentiation of cytomegalovirus strains by restriction analysis of DNA sequences amplified from clinical specimens. J. Infect. Dis. **162:**738–742.
6. **Chou, S.** 1992. Comparative analysis of sequence variation in gp116 and gp55 components of glycoprotein B of human cytomegalovirus. Virology **188:**388–390.
7. **Chou, S.** Molecular epidemiology of envelope glycoprotein H of human cytomegalovirus. J. Infect. Dis., in press.
8. **Demmler, G. J., G. J. Buffone, C. M. Schimbor, and R. A. May.** 1988. Detection of cytomegalovirus in urine from newborns by using polymerase chain reaction DNA amplification. J. Infect. Dis. **158:**1177–1184.
9. **Einsele, H., M. Steidle, A. Vallbracht, J. G. Saal, G. Ehninger, and C. A. Muller.** 1991. Early occurrence of human cytomegalovirus infection after bone marrow transplantation as demonstrated by the polymerase chain reaction technique. Blood **77:**1104–1110.
10. **Gerna, G., D. Zipeto, M. Parea, M. G. Revello, E. Silini, E. Percivalle, M. Zavattoni, P. Grossi, and G. Milanesi.** 1991. Monitoring of human cytomegalovirus infections and ganciclovir treatment in heart transplant recipients by determination of viremia, antigenemia and DNAemia. J. Infect. Dis. **164:**488–498.
11. **Jiwa, N. M., G. W. Van Gemert, A. K. Raap, F. M. Van de Rijke, A. Mulder, P. F. Lens, M. M. M. Salimans, F. E. Zwaan, W. Van Dorp, and M. Van der Ploeg.** 1989. Rapid detection of human cytomegalovirus DNA in peripheral blood leukocytes of viremic transplant recipients by the polymerase chain reaction. Transplantation **48:**72–76.
12. **Kwok, S., D. E. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson, and J. J. Sninsky.** 1990. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. **18:**999–1005.
13. **Nelson, J. A., J. W. Gnann, and P. Ghazal.** 1990. Regulation and tissue-specific expression of human cytomegalovirus. Curr. Top. Microbiol. Immunol. **154:**75–100.
14. **Shibata, D., W. J. Martin, M. D. Appleman, D. M. Causey, J. M. Leedom, and N. Arnheim.** 1988. Detection of cytomegalovirus DNA in peripheral blood of patients infected with human immunodeficiency virus. J. Infect. Dis. **158:**1185–1192.
15. **Stenberg, R. M., D. R. Thomsen, and M. F. Stinski.** 1984. Structural analysis of the major immediate early gene of human cytomegalovirus. J. Virol. **49:**190–199.
16. **Taylor-Wiedeman, J., J. G. P. Sissons, L. K. Borysiewicz, and J. H. Sinclair.** 1991. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. J. Gen. Virol. **72:**2059–2064.