Regulation of Cytomegalovirus Late-Gene Expression: Differential Use of Three Start Sites in the Transcriptional Activation of ICP36 Gene Expression

FREDRICK S. LEACH AND EDWARD S. MOCARSKI*

Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305

Received 14 October 1988/Accepted 4 January 1989

We have investigated the transcriptional regulation of the human cytomegalovirus γ gene encoding the ICP36 family (p52, the major late DNA-binding protein). The ICP36 transcription unit initiates at three distinct sites which are separated by approximately 50 nucleotides and are differentially regulated during infection. At early times (8 h postinfection), only two of these start sites, the most proximal and distal site, were active whereas at late times (36 h postinfection), the middle start site was activated. Expression from this late start site was dependent upon DNA replication. Consensus TATA elements were located upstream of all three start sites, although the element upstream of the late start site was unusual in both sequence and position when compared with conventional TATA elements. Deletion analysis was used in conjunction with transient assays to define independent promoters in this region. The two early start sites and associated TATA elements functioned as separable independently regulated promoters. The region containing the late start site and TATA element but excluding either of the flanking TATA elements was inactive in transient assays. Our work establishes that the ICP36 gene is under complex early and late transcriptional regulation and that the sequences regulating transcriptional activation are temporally and spatially distinct.

Cytomegalovirus (CMV) has a large 230-kilobase-pair (kbp) DNA genome that carries nearly 200 open reading frames (ORFs) (57; M. Chee, A. T. Bankier, and B. G. Barrell, personal communication). In CMV, as in other herpesviruses, regulated expression of the α (immediate-early), β (delayed-early), and γ (late) genes occurs as an ordered cascade controlled at the transcriptional level by viral transactivators (13, 14, 22, 32, 40, 43, 46, 50, 52, 58, 64, 65; for recent reviews, see references 37 and 63). The most prominent locus of CMV α gene expression is transcriptionally controlled by a promoter with an adjacent strong enhancer (1, 6, 12, 53-55, 59, 61, 66) that is responsive to virion transactivation (50, 58). Transcriptional events also play a role in the regulation of β and γ gene expression (11, 34, 35, 50, 64, 65), and some γ gene transcription is strictly dependent on DNA replication (31). It appears that posttranscriptional events are important in the overall regulation of early and late gene expression (10, 15-17, 20, 64). The *cis*-acting regulatory elements controlling β and γ transcription have not been fully analyzed, although there is accumulating evidence that a family of related α gene protein products plays an important role as transactivators of viral gene expression (22, 43, 51, 52, 60).

Studies of regulation of herpes simplex virus (HSV) late-gene expression have concentrated on those few γ genes (so-called γ_2) whose expression is strictly dependent on viral DNA replication (9, 23, 26, 27), work that has centered on the detailed analysis of three viral genes: the US11 gene (28), the glycoprotein C gene (3, 4, 24, 25) and a gene encoding a 70-kilodalton regulatory protein (32, 46). In contrast to their behavior within the viral genome, γ_2 promoter-regulators act as β promoter-regulators when isolated and assayed outside of the viral genome (3, 4, 46). Studies directed toward dissecting authentic late-specific regulation have therefore depended on the construction of viral recombinants (24, 25,

32, 46). The *cis*-acting sequences that are important in the regulation of the HSV glycoprotein C gene are complex and map between -34 and -19 nucleotides (nt) relative to the transcription start site when the gene is within the viral genome (24, 25). In transient assays, additional sequences downstream of +22 in the glycoprotein C gene appear to be necessary for γ_2 -specific regulation (3). Therefore, multiple distinct elements are important in the regulation of at least one HSV late gene.

In CMV, the ICP36 transcript is expressed by 4 h postinfection (p.i.), long before the protein products are maximally synthesized (15, 19). Therefore, transcriptional activation of this gene appears to precede its full expression. ICP36 transcription initiates at three distinct sites spaced approximately 50 bp apart (15). Here we demonstrate that two of the start sites are activated early (by 8 h p.i.) during infection, while the third is activated more than 24 h later and is dependent upon DNA replication. Further, we show that the sequences important in the transcriptional regulation of ICP36 expression are spatially distinct and that two regions can independently support the regulated transcription of a *lacZ* reporter gene.

MATERIALS AND METHODS

Virus and cells. Human CMV (Towne) was grown, and the titers of the virus on human foreskin fibroblast (HF) cells were determined as previously described (49). COS-7 cells were grown and transfected with 10 to 30 μ g of plasmid using DEAE-dextran (Pharmacia, Inc., Piscataway, N.J.) as previously described (21, 50). Phosphonoformic acid (PFA) (a gift from D. Coen) was dissolved in medium at a concentration of 0.2 mg/ml and added to cells at the time of infection.

Plasmid constructs. Restriction enzymes (New England BioLabs, Inc., Beverly, Mass.), *Escherichia coli* DNA polymerase I Klenow fragment (Bethesda Research Laboratories, Inc., Gaithersburg, Md.), and T4 DNA ligase (New England BioLabs) were used as directed by the manufactur-

* Corresponding author.

ers or by established protocols (30). To construct pON2052, the 9-kbp CMV (Towne) *Xba*I L fragment (15, 38, 62) was cloned into the retrovirus vector, cistoneo (48), at a unique *Cl*aI site (which we converted to an *Xba*I site by linker addition). This intermediate clone was designated cistor-*Xba*L and is constructed such that the retrovirus 5' and 3' long terminal repeats are in the same transcriptional orientation as the ICP36 transcription unit. cistor-*Xba*L was then partially digested with *Xba*I, and after the addition of an *Xho*I linker, the sequences between this *Xho*I site and a natural *Xho*I site approximately 400 bp 5' of the first ICP36 initiation site (Fig. 1) were deleted to generate cistor-ICP36. A 800-bp *Dra*I fragment containing the origin and enhancer from simian virus 40 (21, 50) was inserted at the unique *Xho*I site in cistor-ICP36 to generate pON2052. Plasmid clones pON2053, pON2054, pON2055 and pON2056 were generated by digestion of pON2052 at convenient restriction sites shown in Fig. 2.

A series of ICP36 promoter-regulator fragments derived from pON237 (15) (Fig. 1) were directionally cloned into *Sal*I-*Xba*I-digested pON1 (21, 50), a β -galactosidase (β Gal) expression plasmid; pON268 was prepared by cloning an *Xho*I-*Pvu*II fragment from pON254 (Fig. 1) after converting the *Pvu*II site to an *Xba*I site with a linker to form pON262; pON289, which carried promoter-proximal sequences, was prepared by cloning the 271-bp *Kpn*I-*Xba*I fragment from pON262 (Fig. 1) after converting the *Kpn*I site to an *Xho*I site with linkers; and pON2011, which carries approximately 400 bp from a region upstream of the ICP36 gene, was constructed by cloning the *Xho*I-*Kpn*I fragment from pON254 after converting the *Kpn*I site to an *Xba*I site. One series of plasmids with BAL 31-generated deletions (pON2001, pON2003, pON2004, and pON2005) were generated by processive enzyme digestion from a unique *Sma*I site in pON289 (Fig. 1), followed by the addition of a *Sal*I linker to mark the deletion endpoint (see Fig. 6). pON296 was likewise generated by BAL 31 digestion from the unique *Kpn*I site in pON262 with the subsequent cloning of a *Xho*I-*Xba*I fragment into *Sal*I-*Xba*I-digested pON1. Additional plasmids with BAL 31-generated deletions (pON269 and pON270) were generated from the unique *Pvu*II site in pON254, followed by the addition of an *Xba*I linker to mark the deletion endpoint and cloning of *Xho*I-*Xba*I fragments into *Sal*I-*Xba*I-digested pON1. Other plasmids with deletions (pON2012, pON2013, and pON2014) were likewise generated, by starting with pON270 linearized at the unique *Hind*III site. After the insertion of a *Sal*I linker into the *Sma*I site of pON296 to generate clone pON299, a *Sal*I-*Xba*I fragment was removed and cloned into *Sal*I-*Xba*I-digested pON2012 or pON2013 to generate pON2030 or pON2031, respectively. To construct pON2032, pON2033, pON2034, and pON2035, BAL 31 deletions of *Hind*III-digested pON2001 or pON2003 were prepared (with an *Xba*I linker added to mark the deletion endpoint). The resulting *Sal*I-*Xba*I fragments were cloned into *Sal*I-*Xba*I-digested pON1. pON2048 was generated by inserting a *Sal*I linker into the *Kpn*I site of pON262 and then cloning the 400-bp *Xho*I-*Sal*I fragment in the proper orientation into the *Sal*I site of pON2034. All plasmid clones were sequenced to determine deletion endpoints (33, 45).

Transient assays and RNA analysis. Transfection-infection and β Gal assays were done as previously described (17, 50), except that for late time points (36 to 48 h p.i.), medium containing 4-methylumbelliferyl- β -D-galactoside (Sigma Chemical Co., St. Louis, Mo.) was not added until 24 h p.i. Total cellular RNA isolation and oligo(dT) selection were

done as described previously (7, 34). Cytoplasmic RNA was prepared by incubating cells on ice in a solution consisting of 50 mM Tris hydrochloride (pH 7.5), 12 mM MgCl₂, 150 mM KCl, 50 to 100 U of RNasin (Promega Biotec, Madison, Wis.) and 0.5% Nonidet P40 (BDH Chemicals Ltd., Poole, England) for 5 min, then scraping cells with glass beads, and centrifuging at $2,500 \times g$ for 10 min to remove nuclei. The nuclear pellet was discarded, and an equal volume of a solution of 50 mM Tris hydrochloride (pH 7.5), 400 mM NaCl, 20 mM EDTA, and 1.0% sodium dodecyl sulfate was added to the supernatant. Proteinase K (predigested for 2 h at 37°C) was added to a final concentration of 0.1 mg/ml, and after incubation at 37°C for 20 min, CsCl was added to 1.4 M and sodium acetate was added to 10 mM. The RNA was then pelleted through a 5.7 M CsCl cushion as previously described for whole-cell RNA (15). Primer extension (44) was performed with an end-labeled 79-nt primer prepared from *Xba*I-*Dpn*I-digested pON262. The hybridization conditions were as described for S1 nuclease analysis (5, 15), and the primer was strand separated prior to use (30).

Immunoblot analysis. Transfected or mock-transfected COS-7 cells were solubilized directly for sodium dodecyl sulfate-polyacrylamide gel electrophoresis and separated on 9.25% gels as previously described (15, 38). Western blot (immunoblot) analysis was also done as previously described (15). Monoclonal antibodies to ICP36 (CH16-1 and CH13-2) were kindly provided as ascites fluids by L. Pereira (42).

RESULTS

Position of the ICP36-coding sequences. The structure of the ICP36 transcription unit (Fig. 1) was previously determined by S1 nuclease mapping and RNA blot analysis (15, 38). In order to position the ICP36 protein-coding region within the 4.5-kb transcript, processive 3' deletions of the transcription unit were constructed and tested for expression in COS-7 cells without concurrent CMV infection (Fig. 2). At 72 h posttransfection, cell lysates were subjected to immunoblot analysis with ICP36-specific monoclonal antibodies CH16-1 and CH13-2 (15, 38). Constructs containing the first 1.5 kbp of the transcription unit (pON2052), as well as deletion mutants (pON2053 and pON2054) produced an immunoreactive protein that comigrated with the largest and most abundant species of the ICP36 family (ICP36b) (Fig. 2). pON2055 produced one major high-molecular-weight immunoreactive protein with an apparent molecular weight of 66,000 that, based on the nucleotide sequence in this region, represents an in-frame fusion between the ICP36-coding region and a downstream ORF encoding an α protein (29, 66). pON2055 also produced reactive bands that comigrated with smaller ICP36 species (ICP36e). As expected from the removal of the monoclonal antibody epitopes which map downstream of the *Bgl*II site (38), pON2056 did not express protein recognized by ICP36-specific antibodies. On the basis of the nucleotide sequence in this region, the deletion of 2.5 kbp in pON2056 would result in a 115-amino-acid polypeptide containing the ICP36 amino terminus and a short stretch of amino acids derived from a region downstream of the fusion point (Fig. 2; 15, 38; C. M. Brown, A. T. Bankier, and B. G. Barrell, personal communication). ICP36-specific RNA was expressed from pON2056 to similar levels as those of the other constructs, indicating that transcription was not affected by the deletion (data not shown). This analysis mapped the ICP36-coding sequences to the 5' 1.5 kb of the 4.5-kb ICP36 RNA (15). Sequence analysis in this region (Brown et al., personal communica-

FIG. 1. Structure of ICP36 transcription unit and promoter deletions. At the top of the figure is a restriction map of the ICP36 gene. The transcription unit is shown above with the ICP36-coding region (▨). Arrows upstream of the coding region represent the three transcription start sites. Complete *Bam*HI (B), *Bgl*II (Bg), *Eco*RI (E), *Hind*III (H), *Nco*I (N), *Sal*I (S), *Xba*I (X), and *Xho*I (Xh) maps and partial *Kpn*I (K), *Pst*I (P), and *Pvu*II (Pv) restriction maps are shown. MoAB, Monoclonal antibody. In the middle of the figure, pON237, pON254, and pON262, subclones used for the construction of ICP36 5'-flanking region deletions as described in Materials and Methods, are shown. Deletions within the 5'-flanking region for analysis of ICP36 promoter function are shown at the bottom of the figure. Diagrammatic representations of deletions are described in the text.

FIG. 2. (A) Map position of the ICP36-coding sequences. (A) Diagrammatic representation of deletions in *Xba*I-L (Materials and Methods) that were cloned into a simian virus 40-based vector and tested for expression in COS-7 cells. Symbols: □, SV40 origin/enhancer region; —, ICP36 gene with coding sequences (▨) as in Fig. 1; ▨, downstream ORFs that are fused to the ICP36 in pON2055 and pON2056 as described in the text; ■, retrovirus long terminal repeat from the plasmid cistron used here for transcription termination (48; Materials and Methods). Restriction enzyme abbreviations are as in the legend to Fig. 1. (B) Immunoblot analysis of transfected COS-7 cells for ICP36 expression. Each of the plasmid constructs was introduced into COS-7 cells, and at 72 h posttransfection, cell lysates were subjected to electrophoretic separation and immunoblot analysis with ICP36-specific monoclonal antibodies (Materials and Methods). Cells transfected with pON2052, pON2053, or pON2054 produced ICP36 that migrated in a position identical to authentic ICP36 produced during viral infection of HF cells (I-HF). The fusion protein produced by pON2055 is described in the text. The apparent immunoreactive product migrating with the mobility of ICP36f in the pON2056-transfected cells was not reproducible. Molecular weight standards (STD) are indicated to the left of the gel, and the multiple forms of ICP36 are indicated to the right of the gel. The higher-molecular-weight bands present in all COS-7 cells and the band at the position of ICP36f in the pON2056-transfected cells represent background reactivity due to the large amounts of protein loaded onto these lanes.

tion) of CMV strain AD169 is in agreement with our mapping of the ICP36-coding region and positions an ORF predicted to encode a 433-amino-acid protein to this region.

Precise mapping of ICP36 transcription start sites. Previous mapping of the ICP36 transcription unit by S1 nuclease and primer extension revealed the presence of three start sites spaced approximately 50 bp apart (15). Primer extension analysis was performed to accurately position these transcription start sites. Poly(A)⁺ RNA from uninfected and CMV-infected (48 h p.i.) HF cells was hybridized to a 5'-end-labeled 79-nt primer (see Materials and Methods). Consistent with the previous report (15), three major extended products (labeled start sites 1, 2, and 3) were observed in infected-cell RNA (Fig. 3A). Start site 2 consisted of one major and one minor species separated by 4 nt, which could represent either heterogeneous starts or a premature stop by reverse transcriptase. Nucleotide sequence analysis showed A-T-rich elements upstream of each start site (Fig. 3B). The A-T-rich elements upstream of start sites 1 and 3 clearly fit the consensus TATA box sequence (39) and are spaced the conventional distance of 38 or 39 bp from the first base of their respective TATA elements. The A-T-rich element upstream of start site 2 is less similar to a consensus TATA sequence, and the start sites are spaced further downstream (43 to 47 nt) than is usual. The only region of significant homology between start sites 1 and 3 is the TATA element. The translational start site for ICP36 is downstream of all three major transcription start sites, indicating that all of these transcripts could encode the same protein product.

Differential use of ICP36 transcription start sites. We used S1 analysis to investigate the differential use of ICP36 transcription start sites. S1 analyses were performed on total poly(A)⁺ RNA and cytoplasmic RNA was isolated from CMV-infected cells at 8, 24, and 48 h p.i. (Fig. 4A). Although all three initiation sites were apparently used at 48 h p.i., only transcripts corresponding to start sites 1 and 3 were readily detectable at 8 or 24 h p.i. [24 h p.i. poly(A)⁺ RNA data not shown].

The late activation of start site 2 accounted for approximately one-third of late ICP36 transcripts and was reminiscent of γ_2 gene expression in HSV (9, 23, 46, 63). To test the hypothesis that start site 2 usage was dependent upon viral DNA replication, we investigated transcription of this promoter in the presence of PFA. PFA is an inhibitor of DNA replication related to phosphonoacetic acid, a compound that blocks expression of ICP36 (15). Because it is not toxic, PFA can be used to treat cells for periods of time longer than 24 h and has been useful in studying the temporal characteristics of the simian CMV major late DNA-binding protein, the ICP36 equivalent (2, 19). To determine whether DNA replication was necessary for the use of start site 2, poly(A)⁺ RNA isolated from PFA-treated HF cells at 48 h p.i. was compared with RNA from untreated cells at 24, 36, and 48 h p.i. by S1 nuclease analysis [Fig. 4B; 48 h p.i. PFA-treated poly(A)⁺ RNA data not shown]. Transcripts initiating at start site 2 were not detectable in poly(A)⁺ RNA from infected cells maintained in the presence of PFA for 48 h, at least 12 h beyond the time ICP36 transcripts were first detectable in the absence of drug. Inhibition of DNA synthesis reduced to a much lesser extent steady-state levels of RNA initiating from start sites 1 and 3 compared with those of untreated samples. Consistent with the S1 nuclease results, RNA blot analysis of PFA-treated total or poly(A)⁺ RNA showed a reduction (by as much as 70%) of the 4.5-kb ICP36 transcript relative to the transcript level in untreated RNA (data not shown). The PFA sensitivity of start site 2 suggested that its use was dependent on viral DNA replication in a manner analogous to the γ_2 class of genes in HSV. Thus, the multiple start sites of ICP36 transcripts are differentially expressed at early and late times.

The ICP36 gene has two separable promoter-regulators. We characterized the *cis*-acting regulatory signals within the 5'-flanking region of the ICP36 gene using transient assay with various promoter fragments fused to the *E. coli lacZ* indicator gene. An *Xho*I-*Pvu*II fragment from pON237 (Fig. 1) was fused upstream of the *lacZ* gene (21, 50) to produce

FIG. 3. Fine mapping of the ICP36 start sites and sequence analysis of the 5'-flanking region. (A) Primer extension analysis to fine map the ICP36 initiation sites. A 79-nt primer (P) was annealed to poly(A)⁺ RNA prepared from 48-h infected (48) or uninfected (U) HF cells. Primed transcripts were extended by using avian reverse transcriptase and analyzed on a 6% polyacrylamide, 8 M urea sequencing gel. Start sites 1, 2, and 3 are seen only in the 48-h lane. Start site 2 is a doublet. Size standards (STD) are *Msp*I digests of plasmid pAT153 (44). (B) Nucleotide sequence of the ICP36 promoter-proximal region. Large arrows above the sequence indicate the positions of strong stops, while the small arrow indicates the position of the minor stop at start site 2. Heavy lines denote the position of the TATA elements upstream of the major start sites. A dyad symmetry element adjacent to TATA element 2 is indicated by the horizontal arrows below the sequence. The position of the ICP36 translation initiation codon (—) and the sequences corresponding to the 79-nt primer (---) are also shown.

pON268 as described in Materials and Methods. This construct was compared with two previously described constructs, pON249, an α promoter-regulator, and pON241, a β promoter-regulator (16, 50), by transient transfection-infection assay in HF cells. This promoter-regulatory region was transactivated by CMV and expressed β Gal in a temporal manner (Fig. 5), indicating it functioned as a β promoter in transient transfection-infection assays. The time of activation in transient assays corresponded to the time of activation of start sites 1 and 3 during natural CMV infection (Fig. 4). As would be expected, pON268 activation was not affected by PFA in this assay (data not shown). This DNA fragment contained the three start sites and more than 400 bp upstream of start site 1, as well as the ICP36 ATG codon (Fig. 3B). As a result of stop codons within the polylinker sequences downstream of the fusion to *lacZ* sequences in pON268, translational initiation at this ATG codon would be terminated upstream of the *lacZ* translation initiation codon (21, 50). Any translation initiation at this site did not substantially reduce expression of the *lacZ* insert in pON268 relative to pON269 in which the sequences including the ATG were deleted (Table 1). Thus, we conclude that

pON268 and pON269 carry the ICP36 promoter-regulator as indicated by the transient assay.

Deletion analysis was used to dissect the promoter-regulator region and identify *cis*-acting signals responsible for viral transactivation. The 600-bp ICP36 promoter-regulator fragment carried by pON268 was dissected into fragments containing none, one, two, or all three TATA elements with various amounts of flanking sequence (Fig. 1 and Table 1). Each of these constructs was analyzed by transient transfection-infection assay, followed by assaying for β Gal activity at 48 h p.i. The results of these experiments and the precise deletion endpoints are shown in Table 1. The β Gal activity for each clone is expressed as a percentage of the activity of clone pON268. The results are the average of at least two replicate transfections, and each clone has been assayed on at least three separate occasions. This analysis indicated that the TATA elements adjacent to either start site 1 or 3 functioned as independently regulated promoter elements. Deletion of sequences upstream of the *Kpn*I site 59 bp from start site 1 (pON289) did not substantially reduce the level of activation compared with that of pON268 (Table 1). Thus, this 266-bp promoter-regulator was sufficient for

FIG. 4. S1 analysis of ICP36 transcription initiation at early and late times. A 520-bp *Bgl/II-KpnI* fragment covering the region of transcription initiation (see Fig. 1) 5' end labeled at the *Bgl/II* site (P) was annealed to uninfected-cell RNA (U) and RNA from cells at 8, 24, or 48 h p.i. maintained in the presence or absence of PFA. (A) Differential expression from start site 2. The 8 and 24 h p.i. cytoplasmic samples contained 25 μ g of RNA and the 48 h p.i. cytoplasmic sample (CYTO) contained 20 μ g of RNA. The 8 and 48 h p.i. samples in the poly(A)⁺ (A⁺) lanes contained 1 μ g of oligo(dT)-selected RNA. Other bands represented protection due to reannealing of the 520-bp probe as well as protection of the probe by a read through transcript (16). The bands in the uninfected-cell RNA lane represent S1 nuclease digestion at the A-T-rich regions within the probe (corresponding to the TATA elements). (B) DNA replication is required for expression from start site 2. poly(A)⁺ RNA samples from 24 and 36 h p.i. as described above were compared with samples from cells infected and maintained in the presence of PFA for 48 h (48 PFA). All lanes contain equal amounts of RNA. Unmarked bands are as in panel A above.

FIG. 5. Kinetics of β Gal induction by the ICP36 promoter region resembles the kinetics of an early gene. An approximately 670-bp *XhoI-PvuII* fragment (see Fig. 1) was cloned into the β Gal expression vector pON1 (21, 50) and designated pON268. An α promoter construct (pON249) and a β promoter construct (pON241), both fused to *lacZ* (17, 50), are included for comparison. Plasmid (30 μ g) were transfected into permissive HF cells and subsequently infected with 10 to 20 PFU of CMV 48 h later. Medium containing the β Gal substrate 4-methylumbelliferyl- β -D-galactoside (MUG) (Materials and Methods) was overlaid onto the fibroblast monolayers and sampled at the indicated times during infection for fluorescence.

regulated expression. Consistent with these observations, a construct (pON2011) which carried the region upstream of this *KpnI* site and a construct devoid of TATA elements altogether (pON2004) were inactive in this assay. pON2012 and pON2013, both of which carry only the TATA element upstream of start site 1, were activated to levels comparable

TABLE 1. Activation of ICP36 promoter deletion mutants

Plasmid construct	Region carried ^a	Activity relative to pON268
pON268	<i>XhoI</i> -268	1.0
pON269	<i>XhoI</i> -183	1.2
pON270	<i>XhoI</i> -109	1.2
pON2014	<i>XhoI</i> -92	0.6
pON2013	<i>XhoI</i> -68	1.1
pON2012	<i>XhoI</i> -35	1.3
pON2011	<i>XhoI</i> -1	<0.1
pON289	5-268	0.8
pON2003	38-268	0.9
pON2001	55-268	0.7
pON2005	91-268	0.9
pON296	113-268	0.9
pON2004	144-268	<0.1
pON2030	<i>XhoI</i> -35 + 113-268	1.2
pON2031	<i>XhoI</i> -68 + 113-268	1.2
pON2032	51-105	<0.1
pON2033	51-96	<0.1
pON2034	38-113	<0.1
pON2035	38-99	<0.1
pON2048	<i>XhoI</i> -1 + 38-113	<0.1

^a Numbers refer to the nucleotide sequence shown in Fig. 3B.

with that of pON268. Likewise, pON296 and pON2005, both of which carry only start site 3 and the adjacent TATA element, were activated to high levels. Thus, either of these TATA elements and surrounding sequences independently directed expression to levels as high as that of the complete ICP36 promoter-regulator. Although fragments carrying TATA elements upstream of start sites 1 or 3 were independently transactivated, none of the constructs carrying only the TATA element upstream of start site 2 (pON2032, pON2033, pON2034, and pON2035) were activated in this assay, even when the region upstream of the *KpnI* site was present (pON2048). Constructs having two or three TATA elements (pON269, pON270, pON2014, pON2003, pON2001, pON2030, and pON2031) were consistently activated to levels comparable with those of pON268, pON269, and pON289. This plasmid deletion series established the independent regulation of two of the three TATA elements directing expression of ICP36 transcripts.

DISCUSSION

We have investigated the transcriptional regulation of ICP36, a CMV γ gene. Transcription of ICP36 was detectable as early as 8 h p.i., at which time we have demonstrated that two of three start sites were active. Starting between 24 and 36 h p.i., and dependent upon DNA replication, start site 2 was activated to direct expression of ICP36 transcripts. To add to this complex regulation of ICP36 transcription, post-transcriptional regulation of ICP36 expression results in a delay in the accumulation of protein until 24 to 36 h p.i. (15). Indeed, ICP36 is known as the major late DNA-binding protein and is one of only three CMV proteins that accumulate to high levels in infected cells (the others being the major capsid protein and an abundant tegument protein) (19, 56). The reason why the ICP36 transcript is synthesized and accumulates as long as 30 h prior to its active translation is unclear at present. The activation of start site 2 is coordinate with the expression of ICP36 in infected cells; however, we do not believe the two are linked. All sequences carried on transcripts initiating at start site 2 are already present on transcripts initiating at start site 1. This arrangement suggests that specific signals unique to the late transcript would be unlikely.

Transcriptional regulation of ICP36 expression occurs on at least two levels. The two transcription start sites that are activated early in infection presumably respond to α regulatory proteins. We have shown that the sequences responsible for this early transactivation may be separated into two independent promoter-regulators. Each contains a conventional TATA element which is the only region of significant homology between start sites 1 and 3 and is essential for expression. Late in infection, sequences within the ICP36 promoter that are dependent on DNA replication result in transcriptional activation at start site 2. This activation may also be dependent upon α regulatory proteins but also requires some function associated with DNA replication or indeed the process of replication itself. A temporal change in start site usage associated with the process of DNA replication has been reported for simian virus 40 early-gene expression (18) and adenovirus E1A expression (41); however, the mechanism of this switch is unclear.

The ability of the ICP36 promoter to maintain specific late transcriptional regulation at start site 2 while transcription is occurring 60 nt 5' and 45 nt 3' suggests that the γ -specific regulatory elements in this region are interspersed with and dominant over the early regulatory elements in this region.

There are several possible models for late transcriptional activation of start site 2, including the following. (i) A repressor could physically prevent initiation until late in infection. (ii) A transactivator could act late in infection to specifically induce transcription. (iii) A conformational change in the DNA sequence resulting from DNA replication could allow increased transcription. (iv) RNA polymerase II or transcription factors could become modified late in infection and recognize specific promoter sequences with higher affinity. There are several differences in the sequence surrounding the TATA element upstream of start site 2 which make these models attractive in the differential regulation of start site 2. First, its sequence is different from the TATA elements upstream of start sites 1 and 3. Second, a region of perfect dyad symmetry is located immediately 3' to the TATA element (Fig. 3B). Third, the positioning (43 to 47 nt) of the start site is unique when compared with those of the other TATA elements. Several groups working on other systems have implicated the role of the TATA element as critical in the regulation of gene expression (8, 24, 36, 47, 67), including the differential response of genes to transactivation (47). Our work suggests TATA-specific transcription factors may be differentially activated during CMV replication.

A possible mechanism consistent with our results involves the dependence of late activation on secondary structure of the ICP36 promoter such that the TATA element and the region of dyad symmetry upstream of the start site 2 are important for activation. However, in our transient transfection-infection results, a construct containing these sequences (pON2034) did not express β Gal. Therefore, we conclude that, at least for transient assays, sequences further 5' or 3' may be important in the activation of start site 2. More-specific delineation of these sequences will be needed to understand late transcriptional regulation of the ICP36 gene.

ACKNOWLEDGMENTS

We are grateful to L. Pereira for supplying monoclonal antibodies and to M. Kirichenko for excellent tissue culture assistance.

These studies were supported by grant MV230 from the American Cancer Society and Public Health Service grants AI20211 and HL33811 from the National Institutes of Health. F.S.L. was supported by a Medical Scientist Training Grant (GM07365) from the Public Health Service.

LITERATURE CITED

1. Akrigg, A., G. W. Wilkinson, and J. D. Oram. 1985. The structure of the major immediate early gene of human cytomegalovirus strain AD169. *Virus Res* 2:107-121.
2. Anders, D. G., A. Irmieri, and W. Gibson. 1986. Identification and characterization of a major early cytomegalovirus DNA-binding protein. *J. Virol.* 58:253-262.
3. Arsenakis, M., G. Campadelli-Fiume, M. T. Lombardo, and B. Roizman. 1988. The glycoprotein C gene of herpes simplex virus 1 resident in clonal L cell lines manifests two regulatory domains conferring a dominant β and a subordinate γ_2 regulation. *Virology* 162:300-310.
4. Arsenakis, M., L. F. Tomasi, V. Speziali, B. Roizman, and G. Campadelli-Fiume. 1986. Expression and regulation of glycoprotein C gene of herpes simplex virus 1 resident in a clonal L-cell line. *J. Virol.* 58:367-376.
5. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
6. Boshart, M., F. Weber, G. Jahn, K. Dorsch-Hasler, B. Fleckenstein, and W. Schaffner. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus.

- rus. Cell 41:521-530.
7. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18:5294-5299.
8. Coen, D. M., S. P. Weinheimer, and S. L. McKnight. 1986. A genetic approach to promoter recognition during trans-induction of viral gene expression. *Science* 234:53-59.
9. Conley, A. J., D. M. Knipe, P. C. Jones, and B. Roizman. 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of γ polypeptides. *J. Virol.* 37:191-206.
10. DeMarchi, J. M. 1983. Post-transcriptional control of human cytomegalovirus gene expression. *Virology* 124:390-402.
11. DeMarchi, J. M., C. A. Schmidt, and A. S. Kaplan. 1980. Patterns of transcription of human cytomegalovirus in permissively infected cells. *J. Virol.* 35:277-286.
12. Dorsch-Hasler, K., G. M. Keil, F. Weber, M. Jasin, W. Schaffner, and U. H. Koszinowski. 1985. A long and complex enhancer activates transcription of the gene coding for the highly abundant immediate early mRNA in murine cytomegalovirus. *Proc. Natl. Acad. Sci. USA* 82:8325-8329.
13. Elkareh, A., A. J. Murphy, T. Fichter, A. Efstratiadis, and S. Silverstein. 1985. "Transactivation" control signals in the promoter of the herpesvirus thymidine kinase gene. *Proc. Natl. Acad. Sci. USA* 82:1002-1006.
14. Everett, R. D., and M. Dunlop. 1984. Trans-activation of plasmid-borne promoters by adenovirus and several herpes group viruses. *Nucleic Acids Res.* 12:5969-5978.
15. Geballe, A. P., F. S. Leach, and E. S. Mocarski. 1986. Regulation of cytomegalovirus late-gene expression: γ genes are controlled by posttranscriptional events. *J. Virol.* 57:864-874.
16. Geballe, A. P., and E. S. Mocarski. 1988. Translational control of cytomegalovirus gene expression is mediated by upstream AUG codons. *J. Virol.* 62:3334-3340.
17. Geballe, A. P., R. R. Spaete, and E. S. Mocarski. 1986. A cis-acting element within the 5' leader of a cytomegalovirus β transcript determines kinetic class. *Cell* 46:865-872.
18. Ghosh, P. K., and P. Lebowitz. 1981. Simian virus 40 early mRNAs contain multiple 5' termini upstream and downstream from a Hogness-Goldberg sequence: a shift in 5' termini during the lytic cycle is mediated by large T-antigen. *J. Virol.* 40:224-240.
19. Gibson, W. 1983. Protein counterparts of human and simian cytomegaloviruses. *Virology* 128:391-406.
20. Goins, W. F., and M. F. Stinski. 1986. Expression of a human cytomegalovirus late gene is posttranscriptionally regulated by a 3'-end-processing event occurring exclusively late after infection. *Mol. Cell. Biol.* 6:4202-4213.
21. Hall, C. V., P. E. Jacob, G. M. Ringold, and F. Lee. 1983. Expression and regulation of *Escherichia coli* lacZ gene fusions in mammalian cells. *J. Mol. Appl. Genet.* 2:101-109.
22. Hermiston, T. W., C. L. Malone, P. R. Witte, and M. F. Stinski. 1987. Identification and characterization of the human cytomegalovirus immediate-early region 2 gene that stimulates gene expression from an inducible promoter. *J. Virol.* 61:3214-3221.
23. Holland, L. E., K. Anderson, C. Shipman, Jr., and E. Wagner. 1980. Viral DNA synthesis is required for the efficient expression of specific HSV-1 mRNA species. *Virology* 101:10-24.
24. Homa, F. L., J. C. Glorioso, and M. Levine. 1988. A specific 15-bp TATA box promoter element is required for expression of a herpes simplex virus type 1 late gene. *Genes Dev.* 2:40-53.
25. Homa, F. L., T. M. Otal, J. C. Glorioso, and M. Levine. 1986. Transcriptional control signals of a herpes simplex virus type 1 late (γ 2) gene lie within bases -34 to +124 relative to the 5' terminus of the mRNA. *Mol. Cell. Biol.* 6:3652-3666.
26. Honess, R. W., and B. Roizman. 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. *J. Virol.* 14:8-19.
27. Honess, R. W., and B. Roizman. 1975. Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. *Proc. Natl. Acad. Sci. USA* 72:1276-1280.
28. Johnson, P. A., and R. D. Everett. 1986. The control of herpes simplex virus type-1 late gene transcription: a TATA-box/capsid region is sufficient for fully regulated activity. *Nucleic Acids Res.* 14:8247-8264.
29. Kouzarides, T., A. T. Bankier, S. C. Satchwell, E. Preddy, and B. G. Barrell. 1988. An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology* 165:151-164.
30. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. Martinez, J., and S. C. St. Jeor. 1986. Molecular cloning and analysis of three cDNA clones homologous to human cytomegalovirus RNAs present during late infection. *J. Virol.* 60:531-538.
32. Mavromara-Nazos, P., and B. Roizman. 1987. Activation of herpes simplex virus 1 γ 2 genes by viral DNA replication. *Virology* 161:593-598.
33. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
34. McDonough, S. H., and D. H. Spector. 1983. Transcription in human fibroblasts permissively infected by human cytomegalovirus strain AD169. *Virology* 125:31-46.
35. McDonough, S. H., S. I. Staprans, and D. H. Spector. 1985. Analysis of the major transcripts encoded by the long repeat of human cytomegalovirus strain AD169. *J. Virol.* 53:711-718.
36. McKnight, S. E., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein coding gene. *Science* 217:316-324.
37. Mocarski, E. S. 1988. Biology and replication of cytomegalovirus. *Transfusion Med. Rev.* 2:229-234.
38. Mocarski, E. S., L. Pereira, and N. Michael. 1985. Precise localization of genes on large animal virus genomes: use of λ gt11 and monoclonal antibodies to map the gene for a cytomegalovirus protein family. *Proc. Natl. Acad. Sci. USA* 82:1266-1270.
39. Nevins, J. R. 1983. The pathway of eukaryotic mRNA formation. *Annu. Rev. Biochem.* 52:441-466.
40. O'Hare, P., and G. S. Hayward. 1985. Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. *J. Virol.* 53:751-760.
41. Osborne, T. F., and A. J. Berk. 1983. Far upstream initiation sites for adenovirus early region 1A transcription are utilized after the onset of DNA replication. *J. Virol.* 45:594-599.
42. Pereira, L., M. Hoffman, D. Gallo, and N. Cremer. 1982. Monoclonal antibodies to human cytomegalovirus: three surface membrane proteins with unique immunological and electrophoretic properties specify cross-reactive determinants. *Infect. Immun.* 36:924-932.
43. Pizzorno, M. C., P. O'Hare, L. Sha, R. L. LaFemina, and G. S. Hayward. 1988. *trans*-Activation and autoregulation of gene expression by the immediate-early region 2 gene products of human cytomegalovirus. *J. Virol.* 62:1167-1179.
44. Ravetch, J. V., R. Feder, A. Pavlovic, and G. Blobel. 1984. Primary structure and genomic organization of the histidine-rich protein of the malaria parasite *Plasmodium lophurae*. *Nature (London)* 312:616-620.
45. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
46. Silver, S., and B. Roizman. 1985. γ 2-Thymidine kinase chimeras are identically transcribed but regulated as γ 2 genes in herpes simplex virus genomes and as β genes in cell genomes. *Mol. Cell. Biol.* 5:518-528.
47. Simon, M. C., T. M. Fisch, B. J. Benecke, J. R. Nevins, and N. Heintz. 1988. Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation. *Cell* 52:723-729.
48. Sorge, J., D. Wright, V. D. Erdman, and A. E. Cutting. 1984. Amphotropic retrovirus vector system for human cell gene transfer. *Mol. Cell. Biol.* 4:1730-1737.

49. Spaete, R. R., and E. S. Mocarski. 1985. The α sequence of the cytomegalovirus genome functions as a cleavage/packaging signal for herpes simplex virus defective genomes. *J. Virol.* **54**: 817–824.
50. Spaete, R. R., and E. S. Mocarski. 1985. Regulation of cytomegalovirus gene expression: α and β promoters are *trans* activated by viral functions in permissive human fibroblasts. *J. Virol.* **56**:135–143.
51. Spector, D. J., and M. J. Tevethia. 1986. Identification of a human cytomegalovirus virus DNA segment that complements an adenovirus 5 immediate early mutant. *Virology* **151**:329–338.
52. Staprans, S. I., D. K. Rabert, and D. H. Spector. 1988. Identification of sequence requirements and *trans*-acting functions necessary for regulated expression of a human cytomegalovirus early gene. *J. Virol.* **62**:3463–3473.
53. Stenberg, R. M., and M. F. Stinski. 1985. Autoregulation of the human cytomegalovirus major immediate-early gene. *J. Virol.* **56**:676–682.
54. 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.
55. Stenberg, R. M., P. R. Witte, and M. F. Stinski. 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.
56. Stinski, M. F. 1977. Synthesis of proteins and glycoproteins in cells infected with human cytomegalovirus. *J. Virol.* **23**:751–767.
57. Stinski, M. F. 1983. The molecular biology of cytomegaloviruses, p. 67–113. *In* B. Roizman (ed.), *The herpesviruses*, vol. 2. Plenum Publishing Corp., New York.
58. Stinski, M. F., and T. J. Roehr. 1985. Activation of the major immediate early gene of human cytomegalovirus by *cis*-acting elements in the promoter-regulatory sequence and by virus-specific *trans*-acting components. *J. Virol.* **55**:431–441.
59. Stinski, M. F., D. R. Thomsen, R. M. Stenberg, and L. C. Goldstein. 1983. Organization and expression of the immediate early genes of human cytomegalovirus. *J. Virol.* **46**:1–14.
60. Tevethia, M. J., and D. J. Spector. 1984. Complementation of an adenovirus 5 immediate early mutant by human cytomegalovirus. *Virology* **137**:428–431.
61. Thomsen, D. R., R. M. Stenberg, W. F. Goins, and M. F. Stinski. 1984. Promoter-regulatory region of the major immediate early gene of human cytomegalovirus. *Proc. Natl. Acad. Sci. USA* **81**:659–663.
62. Thomsen, D. R., and M. F. Stinski. 1981. Cloning of the human cytomegalovirus genome as endonuclease XbaI fragments. *Gene* **16**:207–216.
63. Wagner, E. K. 1985. Individual HSV transcripts: characterization of specific genes, p. 45–104. *In* B. Roizman (ed.) *The herpesviruses*, vol. 3. Plenum Publishing Corp., New York.
64. Wathen, M. W., and M. F. Stinski. 1982. Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *J. Virol.* **41**:462–477.
65. Wathen, M. W., D. R. Thomsen, and M. F. Stinski. 1981. Temporal regulation of human cytomegalovirus transcription at immediate early and early times after infection. *J. Virol.* **38**: 446–459.
66. Wilkinson, G. W., A. Akrigg, and P. J. Greenaway. 1984. Transcription of the immediate early genes of human cytomegalovirus strain AD169. *Virus Res.* **1**:101–106.
67. Wu, L., S. E. Rosser, M. C. Schmidt, and A. Berk. 1987. A TATA box implicated in E1A transcriptional activation of a simple adenovirus 2 promoter. *Nature (London)* **326**:512–515.