Translational Inhibition Mediated by a Short Upstream Open Reading Frame in the Human Cytomegalovirus gpUL4 (gp48) Transcript

CATHERINE R. DEGNIN,[†] MARK R. SCHLEISS,[‡] JIANHONG CAO, AND ADAM P. GEBALLE*

Department of Molecular Medicine and Division of Clinical Research, Fred Hutchinson Cancer Research Center C2-023, 1124 Columbia Street, Seattle, Washington 98104-2092

Received 26 April 1993/Accepted 14 June 1993

The human cytomegalovirus (CMV) virion glycoprotein gpUL4 (gp48) gene expresses a transcript that contains three AUG codons upstream from the one used to initiate synthesis of the gp48 protein. Previously we reported that the second of these AUG codons, AUG2, was necessary but insufficient for inhibition of downstream translation (M. Schleiss, C. R. Degnin, and A. P. Geballe, *J. Virol.* 65:6782–6789, 1991). We now demonstrate that the coding information of the upstream open reading frame initiated by AUG2 (uORF2) is critical for the inhibitory signal. Several missense mutations, particularly those involving the carboxy-terminal codons of uORF2, inactivate the inhibitory signal, while mutations that preserve the coding content of uORF2 uniformly retain the inhibitory signal. The uORF2 termination codon is essential for inhibition, but leader sequences further downstream are not critical. Conservation of uORF2 among clinical strains of CMV suggests that uORF2 provides an important function in the CMV infectious cycle. Although these results indicate that the peptide product of uORF2 mediates the inhibitory effect, we demonstrate that the uORF2 signal acts only in *cis*, and we propose a model of inhibition by the gp48 uORF2 signal.

The 230-kb human cytomegalovirus (CMV) genome encodes approximately 200 open reading frames (ORFs) (5). During productive infection of human fibroblasts (HF) in cell culture, the expression of CMV genes is temporally regulated (30). In addition to transcriptional controls, posttranscriptional and translational events modulate the expression of viral proteins of α (or immediate-early), β (or early), and γ (or late) temporal classes (9–12, 19, 22, 27–29, 33). The observation that several CMV genes express transcripts long before their respective protein products are synthesized (9, 27) suggests that some CMV gene transcripts contain *cis*-acting signals that mediate a translational delay in expression.

We identified one translational inhibitory signal in the 5' leader of the predominant transcript of the gene encoding the virion glycoprotein gpUL4 (gp48) (22). Like several other CMV gene transcripts, the gp48 transcript contains upstream AUG codons and associated short upstream ORFs (uORFs) (4). The second of the three upstream AUG codons in the gp48 transcript leader (AUG2) is an essential component of a signal responsible for inhibition of downstream translation (22). The scanning model of eukaryotic translation suggests that ribosomes usually initiate only at the 5'-proximal AUG codon of a transcript and thus predicts that upstream AUG codons should inhibit downstream translation (15). However, analyses of several CMV transcript leaders containing upstream AUG codons demonstrated that not all such leaders repress translation of the downstream ORF (3). No established properties of eukaryotic translation account for

the different effects of upstream AUG codons in CMV gene transcripts. For example, the inhibitory effect of the gp48 AUG2 codon but neither of the other two upstream AUG codons (22) cannot be easily explained by differences in the nucleotide context surrounding the upstream AUG codons (14) or by differences in location of the upstream AUG codon within the transcript leader (15, 16, 23).

This study was designed to delineate which gp48 leader sequences are required in conjunction with AUG2 for inhibition of downstream translation by the gp48 leader. We demonstrate that the amino acid coding information of the uORF initiated by AUG2 (uORF2), but neither the nucleotide content of the gp48 transcript leader per se nor the 3' flanking sequences, is required. uORF2 is conserved among several clinical strains of CMV. Although these data strongly suggest that the peptide product of uORF2 mediates translational repression, the uORF2 signal operates only in *cis*.

MATERIALS AND METHODS

Cells, virus, and viral DNA. Human CMV(Towne) was grown in HF in Dulbecco's modified Eagle medium supplemented with 10% NuSerum (Collaborative Research, Inc., Bedford, Mass.) as described previously (26). DNAs from the CMV strains cultured in the virology laboratory at the Fred Hutchinson Cancer Research Center were kindly provided by B. Fries and B. Torok-Storb.

Plasmids and sequence analyses. Plasmids pEQ3, pEQ176, pEQ239, and pEQ325 (22) and pEQ134 (3) have been described previously. We constructed a series of gp48 leader- β -galactosidase (β -Gal) plasmids which contained deletion of the gp48 leader sequences from near the 3' end of uORF2 through either +121 or +197. After digestion of pEQ239 with *Afl*II, blunting with DNA polymerase (Klenow), and then cutting with *Hind*III, the resulting 94-bp fragment and the

* Corresponding author. Electronic mail address: ageballe@fred.hfrc.org.

[†] Present address: Department of Molecular and Medical Genetics, Oregon Health Sciences University, Portland, OR 97201.

[‡] Present address: Department of Pediatrics, Children's Hospital Research Foundation, Cincinnati, OH 45229.

FIG. 1. Effects of the uORF2 termination codon and downstream sequences on the uORF2 inhibitory signal. The indicated pEQ plasmids express transcripts with 5' ends (left) containing gp48 leader sequences upstream of the β -Gal ORF (open rectangle). The wild-type 228-nucleotide gp48 leader in pEQ239 contains three uORFs (rectangles). The uORF2 sequences (black) are entirely wild type (pEQ239, pEQ366, and pEQ365) or are fused to other gp48 leader sequences (stippled; pEQ353, pEQ357, and pEQ352) because of deletion of the uORF2 termination codon (+87 through +89). The numbers above the transcript leaders indicate the final nucleotide prior to the 5' end and the first nucleotide after the 3' end of each deletion. (A) The β -Gal activity (mean $\pm$ standard deviation) from triplicate 60-mm-diameter dishes was determined in cells transfected with the control plasmids (white bars) with no promoter (pEQ3) or with no CMV leader (pEQ176) or with plasmids containing (striped bars) or lacking (cross-hatched bars) the uORF2 wild-type termination codon and then infected with CMV for 24 h. Accumulated β -Gal transcripts expressed by the test plasmids (B) and from an enzymatically inactive control plasmids pEQ430 (data not shown) were detected by Northern blot analysis of whole cell RNA as described in Materials and Methods. Here and in Fig. 2 and 3, MUG stands for methylumbelliferyl- β -D-galactoside.

113-bp *RsaI*-*Asp718* 3' gp48 leader fragment from pEQ239 were ligated into *HindIII*-*Asp718*-digested pEQ176, resulting in pEQ366 (Fig. 1). Use of the same cloning steps but with the 37-bp *SspI*-*Asp718* fragment from the gp48 leader in place of the *RsaI*-*Asp718* 3' leader fragment resulted in pEQ365. To construct plasmids with deletions of the uORF2 termination codon, pEQ239 was digested sequentially with *AflII*, mung bean nuclease, and *HindIII*. Ligation of the resulting 5' gp48 leader fragment and the same 3' leader sequences used to construct pEQ366 and pEQ365 into *HindIII*-*Asp718*-digested pEQ176 resulted in plasmids pEQ357 and pEQ352, respectively (Fig. 1). During construction of pEQ357, the unexpected removal of an additional base pair (+85), presumably by mung bean nuclease, resulted in pEQ353. Plasmids with missense mutations in gp48 uORF2 were constructed by using the primers 18 (GAT CAAGCTTAATCAGATGCCGGCCTTGTGATGCAG) and 26 (TCACTTAAGGCGGGATGTATTTGCAAGTCAGCAAGACGACAGTTTTTTCGCCG) to amplify the gp48 leader from CMV(Towne) DNA (22). Primer 26 was synthesized by using a pool of deoxynucleoside triphosphates (dNTPs) containing 85% of the underlined nucleotides. For positions with underlined G or A, the remaining 15% was composed of equal quantities of each of the other three dNTPs, while for position with underlined T, the remaining 15% contained equal quantities of dCTP and dGTP. This oligonucleotide was designed to generate amplified products containing one to two mutations in the coding information of uORF2 per molecule, at codons 9 through 21. The amplified products were digested with *HindIII* and *AflII* and inserted into

HindIII-*AflII*-digested pEQ239, replacing the wild-type sequences to generate missense mutant plasmids depicted in Fig. 2 and 3 (pEQ393, pEQ394, pEQ397, pEQ398, pEQ400, pEQ402, pEQ403, pEQ404, pEQ406, pEQ408, pEQ412, pEQ414, pEQ415, and pEQ417).

Plasmids with mutations of nucleic acid content of uORF2 but not affecting the coding information were constructed by polymerase chain reaction (PCR) amplification of the gp48 leader in CMV(Towne) DNA by using primers 18 and 30 (TCACTTAAGG[AGT]GG[AT]ATATACTTGCAAG; the bracketed nucleotides represent equimolar mixtures of the indicated dNTPs). The amplified products were digested with *HindIII* and *AflII* and then inserted into pEQ239 digested with *HindIII* and *AflII*, resulting in pEQ418, pEQ419, pEQ420, and pEQ421 (Fig. 1). Digestion of pEQ239 with *AflII*, followed by blunting with mung bean nuclease and religation, resulted in pEQ341.

A truncated β -Gal construct containing the gp48 leader with mutation of AUG2 was constructed by digesting pEQ325 (22) partially with *EcoRI* and completely with *SacI*. After blunting of the ends with T4 DNA polymerase and religation, plasmid pEQ430, with a deletion of 1,064 nucleotides from the carboxy terminus of *lacZ*, was identified.

The content of the gp48 leader in each plasmid was verified by sequence analysis using primer 27 (GGAGGTC TATATAAGCAG) and the Sequenase II kit (U.S. Biochemical).

DNAs from HF monolayers infected with clinical CMV strains were amplified by PCR with primers gp48.3 (22) and 19 (GATCAAGCTTTGACTATAAGGATCGCGACCG),

flanking -236 through +234 of the CMV(Towne) gp48 leader. After gel purification, the nucleotide sequence of each strand from +1 to +192 was determined for the clinical isolates, using the *fml* DNA Sequencing System (Promega) with primer 41 (CCCCGTAAGATGATCCTCG), corresponding to the CMV(Towne) sequences -70 to -52, and primer gp48.3 (22).

Transfections and RNA analysis. β -Gal activity was assayed by adding the fluorogenic β -Gal substrate methylumbelliferyl- β -D-galactoside in medium after CMV infection of HF transfected with the indicated test plasmids and a control plasmid (either pEQ430 or pEQ134 [3]), using DEAE dextran as described previously (3, 22). Whole cell RNA was purified from transfected cells by guanidinium isothiocyanate solubilizing transfected cells, pooling the lysates from triplicate 60-mm-diameter dishes, and pelleting the RNA through CsCl (9). RNA was detected by Northern (RNA) blot hybridization as described previously (9).

Polysome distribution analyses. Polysomes from transfected and CMV-infected HF were fractionated through sucrose gradients as described previously (22) except that one experiment (Fig. 5B) was done with 5 to 47% sucrose gradients over a 2 M sucrose pad. RNA was harvested from each fraction was analyzed by Northern blot hybridization using a β -Gal-specific probe.

RESULTS

Inhibition by the gp48 leader requires the uORF2 termination codon but not further-downstream sequences. Previously we reported that the AUG2 codon in the leader of the predominant gp48 transcript was necessary but insufficient to inhibit translation of a downstream ORF (22). To clarify which additional gp48 leader sequences were required elements of the inhibitory signal, we constructed plasmids that express transcripts containing wild-type or mutant gp48 leaders upstream of the β -Gal ORF. As in previous studies (3, 10, 11, 22), the translational impact of these gp48 transcript leaders was assayed by measuring β -Gal activity after transfection of the expression plasmids into HF and subsequent infection with CMV.

We first investigated the possibility, suggested by studies of yeast *GCN4* translational control (18), that sequences immediately downstream from the stop codon of uORF2 are important for translational regulation by the gp48 leader. We constructed plasmids with deletions of the gp48 leader sequences immediately 3' of the uORF2 termination codon (Fig. 1). Plasmids pEQ366 and pEQ365 contain the gp48 leader with deletions of nucleotide +90 through +121 or +90 through +197, respectively. Despite the different nucleotides downstream from the uORF2 termination codon, both of these plasmids expressed low levels of β -Gal, similar to that expressed by the plasmid containing the full-length gp48 leader (pEQ239). Coupled with the demonstration of inhibition by using plasmids with alternate sequences 3' from uORF2 (Table 1), these data suggest that sequences downstream from the uORF2 termination codon are not critical for translational inhibition by the gp48 leader signal.

Previously we reported that deletion of nucleotides +86 through +228 (the 3' end of the gp48 leader) inactivated the inhibitory signal while deletion of nucleotides +90 through +228 retained the signal (22). These data suggested that nucleotides +86 through +89 (TTAA), containing the uORF2 termination codon (underlined), were required for inhibition. We constructed pEQ357 and pEQ352, which are identical to pEQ366 and pEQ365, respectively, except for

TABLE 1. Nucleotide sequence of the uORF2 termination codon and downstream 10 nucleotides in inhibitory gp48 leaders

Plasmid	gp48 leader sequence from the end of uORF2 (+87)	Figure or reference
pEQ239	TAA GTGATGAGTC	Fig. 1-3, reference 22
pEQ330	TAA GATCTCGAGC	Reference 22
pEQ365	TAA ATTTTGATCG	Fig. 1
pEQ366	TAA ACGGTAAAG	Fig. 1
pEQ341	TGA TGAGTCTATA	Fig. 2

the further deletion of nucleotides +86 through +89 (Fig. 1). The unexpected deletion of an extra nucleotide (+85) during cloning resulted in pEQ353. In transfection analyses, pEQ352, pEQ357, and pEQ353 each expressed high levels of β -Gal activity, similar to that expressed by the control with no CMV leader. These data verify that the sequences from +86 through +89 contain information essential for inhibition by the gp48 leader.

After determining β -Gal activity in transfected cells, we harvested whole cell RNA and analyzed the level of accumulated β -Gal RNA by Northern blot hybridization as described in Materials and Methods. All of the test plasmids expressed similar levels of β -Gal RNA (Fig. 1B). Simultaneous analysis of transcripts expressed from a cotransfected, truncated, enzymatically inactive β -Gal control plasmid (pEQ430) verified that transfection efficiency and RNA recovery were similar among the samples (data not shown). Thus, differences in transcript levels did not account for the differences of β -Gal activities among these constructs. Rather, these data indicate that the uORF2 termination codon but not gp48 leader sequences downstream from the termination codon is required for translational inhibition.

Amino acid coding information at the carboxy terminus of uORF2 is essential for inhibition of downstream translation. Coupled with results of previous studies (22), the results presented in Fig. 1 demonstrated that both the initiation and termination codons of uORF2 were required components of the gp48 inhibitory signal. The potential role of uORF2 was further highlighted by the loss of the inhibitory signal in a uORF2 frameshift mutant, with a drastically altered coding content but only slightly changed nucleotide content of uORF2 (22). To more precisely define which elements of uORF2 are critical for translational inhibition, we constructed a set of plasmids containing nucleotide substitutions resulting in missense mutations in uORF2. As described in Materials and Methods, we used a degenerate PCR primer with a mean of approximately two mutations per molecule in the first base of codons 9 through 22 of uORF2. We inserted the PCR products into the β -Gal expression vector pEQ176, isolated and sequenced individual plasmid clones, and analyzed the effects of the mutations in transfection assays.

The results of transfection assays of plasmids containing mutations near the carboxy terminus of uORF2 are shown in Fig. 2. As in previous studies (22), the wild-type gp48 leader (pEQ239) inhibited β -Gal expression approximately 10-fold compared with the control with no CMV leader (pEQ176). Plasmids with single-nucleotide substitutions at either the penultimate proline codon, P-21 (pEQ394 and pEQ397), or the carboxy-terminal proline codon, P-22 (pEQ393 and pEQ400), expressed high levels of β -Gal activity. A mutant with three nucleotide substitutions altering codons K-18, Y-19, and I-20 (pEQ398) also expressed nearly as much β -Gal as did the control with no CMV leader (pEQ176). Thus, nucleotide substitutions that altered the coding con-

FIG. 2. Translational effects of mutations near the carboxy terminus of uORF2. Plasmids with mutations that change (boxed) or preserve the coding content of uORF2 were transfected into HF in triplicate 60-mm-diameter dishes. β -Gal activity (A) and accumulated RNA (B) were assayed 24 h after CMV infection as described in the legend to Fig. 1 except that pEQ134 was cotransfected in place of pEQ340 to control for transfection efficiency and RNA recovery (data not shown).

tent near the carboxy terminus of uORF2 inactivated the inhibitory signal.

We next constructed a set of plasmids with nucleotide substitutions also involving the carboxy terminus of the uORF2 but not altering the coding content of uORF2. Plasmids pEQ418, pEQ419, pEQ420, and pEQ421 each contain substitutions in the third base of uORF2 codons 18 through 21 (Fig. 1). The coding information of uORF2 in each of these plasmids is identical to that of the wild-type gp48 leader construct (pEQ239). In transfection assays, each of these plasmids expressed low levels of β -Gal activity (Fig. 2).

The uORF2 coding content of pEQ341 (Fig. 2) is also identical to that of the wild-type leader. Deletion of the four nucleotides +86 through +89 (ttaa) at the 3' end of the wild-type uORF2 removes the final nucleotide (t) of the terminal proline codon (P-22) and the termination codon (taa). Fortuitously, the next four nucleotides, gtga, restore the carboxy-terminal proline codon (now ccg) and provide a new termination codon (tga) for uORF2. Thus, the wild-type uORF2 coding content is preserved in pEQ341. β -Gal expression from pEQ341 was repressed (Fig. 2), similar to the case for each of the other plasmids that contain the wild-type coding information of uORF2.

Analyses of transcript accumulation in transfected cells revealed that the plasmids expressed similar levels of β -Gal RNA (Fig. 2B). Thus, the differences in β -Gal activity

expressed from these plasmids cannot be explained by transcriptional or RNA stability differences. Rather, the amino acid coding content at the carboxy terminus of uORF2 is required for translational inhibition.

Other missense mutations of uORF2. We constructed and analyzed additional plasmids with missense mutations in uORF2 affecting codons K-9 through Y-19 (Fig. 3). Single-nucleotide substitutions which altered the coding information at positions K-9, K-10, L-11, T-16, and C-17 (pEQ415, pEQ408, pEQ412, pEQ403, and pEQ402, respectively) preserved the inhibitory signal of the gp48 leader. Contrarily, mutations at positions K-10 and S-12 (pEQ406) or L-14 and Y-19 (pEQ404) relieved much of the inhibitory impact of the gp48 leader. Mutation at L-14 (pEQ414) or L-15 (pEQ417) resulted in partial inhibition of downstream translation.

Northern analysis of RNA isolated from transfected cells in this experiment revealed modest variation in the level of β -Gal transcript accumulation. However, differences in β -Gal activity expressed from the plasmids could not be accounted for by alterations in β -Gal transcript accumulation. For example, the lower abundance of pEQ404 transcripts did not explain the increased β -Gal expression by pEQ404 compared with pEQ239. These data indicated that some but not all of the coding information of this central region of uORF2 is critical for translational inhibition.

The gp48 leader in clinical CMV isolates. Although most analyses of CMV use the laboratory-adapted strains CMV-

FIG. 3. Effects of missense mutations of the middle region of uORF2. Plasmids with the indicated missense mutations (boxed) were transfected into HF in triplicate 60-mm-diameter dishes. β -Gal activity (A) and accumulated RNA (B) were assayed 24 h after CMV infection as described in the legend to Fig. 2.

(Towne) and CMV(AD169), some studies have revealed significant differences in phenotype between laboratory-adapted strains and strains isolated from clinical samples (20, 21, 25). To begin investigating regulation of gp48 expression in clinical strains, we PCR amplified the gp48 leader from DNA of cells infected with five clinical isolates of CMV, using primers corresponding to -236 to -215 and complementary to +207 through +234 relative to the gp48 major transcript start site. The nucleotide sequences from +1 through +192 of these PCR products are shown in Fig. 4A. This sequence was at least 84% identical between each pair of isolates. AUG2 was conserved in all five strains, while AUG1 and AUG3 were present in only three of the strains.

The deduced amino acid sequence of uORF2 in these clinical isolates (Fig. 4B) was at least 82% identical in uORF2 coding content. Notably, the coding information of the carboxy-terminal six codons uORF2 was identical in all strains. However, variation among strains occurred at codons Q-2, L-6, S-7, K-9, K-10, S-12, and T-16. Although the significance of this variation in sequence is not yet known, the preservation of uORF2 in all five strains supports the likelihood that uORF2 plays a significant role during the CMV infectious cycle.

uORF2 acts only in *cis*. The requirement for a particular coding content of uORF2 suggests that the peptide product of uORF2 is synthesized and mediates the inhibitory effect on downstream translation. To investigate whether the

uORF2 inhibitory signal acts in *cis* or *trans*, we analyzed the polysomal distribution of transcripts containing and lacking the uORF2 inhibitory signal in transfected cells. pEQ430 expresses a gp48 leader-*lacZ* chimeric transcript with a mutation changing AUG2 to AAG and with a deletion in the 3' end of the β -Gal ORF. After cotransfection of pEQ430 and pEQ239, we infected cells with CMV by using cycloheximide-actinomycin D reversal conditions, under which no early or late viral gene products are synthesized (26). Thus, a peptide product of uORF2 could be made in these cells from the transcript expressed from pEQ239 but not from the viral genomic gp48 transcript. We analyzed the distribution of β -Gal transcripts on sucrose gradients as described in Material and Methods. pEQ239 transcripts sedimented mainly with the ribosomal subunits, monosomes and disomes (Fig. 5A). In contrast, pEQ430 sedimented with larger polysomes (≥ 7 -mers), similar to the position of efficiently translated β -Gal transcripts in previous studies (22). If uORF2 expressed on the pEQ239 transcripts acted in *trans*, we would have expected pEQ430 transcripts to migrate with smaller ribosomes.

In a similar experiment, we analyzed the polysomal distribution of pEQ239 and pEQ430 transcripts CMV infection for 18 h without drug blockades (Fig. 5B). Under these conditions, uORF2 is expressed in cells on both pEQ239 transcripts and viral gp48 transcripts. The migration of pEQ430 on larger polysomes substantiates that uORF2 acts only in *cis*.

FIG. 4. The gp48 leader sequence in CMV strains. (A) The gp48 leader sequence corresponding to the first 192 nucleotides of the predominant CMV(Towne) gp48 transcript leader, including the three upstream AUG codons (underlined) and uORF2 (overlined, uppercase), was determined by using DNA from five clinical strains (C12, C2, C4, C1, and C3). For comparison, this same region of laboratory strains CMV(Towne)(T [4]) and CMV(AD-169) (AD [GenBank accession number X17403]) is shown. Nucleotides identical to those in CMV(Towne) (.) or deleted compared with CMV(Towne) (-) are indicated. (B) The deduced amino acid sequences of uORFs from these same strains. Codons identical to those of CMV(Towne) (.) and mutations that preserve the amino acid of a codon (*) are indicated. The presence (+) or absence (-) of AUG1 and AUG3 is shown.

DISCUSSION

Approximately 5 to 10% of eukaryotic genes express transcripts that, like CMV gp48 transcripts, contain upstream AUG codons and associated short uORFs (17). For most of these genes, the translational effects of the upstream AUG codons and uORFs have not been investigated. Suggested models of eukaryotic translation (15) could not predict the impact of upstream AUG codons in several CMV gene transcripts. Neither the context of nucleotides flanking the upstream AUG codons nor the locations of the uORFs within the transcript leader differentiate which upstream AUG codons inhibit downstream translation (3, 22). The unpredictable effect of upstream AUG codons is apparent in analyses of the gp48 transcript leader, in which AUG2, but neither AUG1 nor AUG3, is required for inhibition of downstream translation in vivo (22).

Transcript leader sequences other than the upstream AUG codons must distinguish inhibitory transcript leaders from those which do not alter downstream translation. In the translationally regulated yeast *GCN4* gene, the different effects of the various uORFs map, in part, to the 10 nucleotides downstream from the uORF termination codons (18). In contrast, the nucleotides immediately downstream from the gp48 uORF2 termination codon do not contribute to the inhibitory effect of uORF2 (Fig. 1). Translation of transcripts in which any of five different sequences follow the uORF2 termination codon is inhibited (Table 1).

Our results also suggest that secondary structure of the gp48 transcript leader is not critical for the inhibitory signal. All plasmids containing the intact uORF2, regardless of other gp48 leader sequences, contain the inhibitory signal. The analyses shown in Fig. 1 suggest that nucleotides +86

FIG. 5. Polysome analysis of cotransfected plasmids containing and lacking the uORF2 signal. A plasmid with the wild-type gp48 leader and full-length β -Gal (pEQ239) was cotransfected with one containing an AUG2 $\rightarrow$ AAG mutation in the gp48 leader and a truncated β -Gal ORF (pEQ430). Polysomes were separated on sucrose gradients as described in Materials and Methods either after infection in the presence of cycloheximide (50 μ g/ml) for 8 h, cycloheximide (50 μ g/ml) and actinomycin (10 μ g/ml) for 1 h, and finally actinomycin D (10 μ g/ml) for 3 h (A) or after CMV infection without drug blockade for 17 h (B). The positions corresponding to monosomes and disomes and to polysomes with at least seven ribosomes were determined from the absorbance profiles (not shown).

through +89 are required for translational inhibition (e.g., compare pEQ366 with pEQ357 or pEQ365 with pEQ352). The possibility that these four nucleotides are required because they are part of an important RNA secondary-structure element is unlikely because plasmid pEQ341 contains a deletion of the same four nucleotides yet retains the inhibitory signal (Fig. 2). Fortuitously, this plasmid preserves the wild-type coding information of uORF2, suggesting that the essential contribution of nucleotides +86 through +89 to the inhibitory signal is the uORF2 termination codon.

Most intriguingly, analyses of mutations affecting the body of uORF2 (Fig. 2 and 3) indicate that the inhibitory signal depends on the amino acid coding information of uORF2. The carboxy-terminal codons, particularly P-21 and P-22, are essential components of the signal (Fig. 2). In contrast, mutations that involve the carboxy-terminal nucleotides of uORF2 but preserve the uORF2 coding content all retain the inhibitory signal. The critical nature of the carboxy terminus of uORF2 to inhibitory function is further revealed in the analysis of uORF2 stop codon deletions (Fig. 1). Addition of only a single codon to uORF2 inactivates the inhibitory signal (Fig. 1, pEQ357 and pEQ352). Although we have tested only a small subset of all possible missense mutations of uORF2, these data demonstrate that some of the coding information of uORF2 is required for inhibition.

Analysis of the gp48 leader in clinical isolates suggests that uORF2 plays an important role in CMV infection, in contrast to uORF1 and uORF3, which are dispensable for CMV replication. Notably, the coding information of the carboxy-terminal region of the gp48 leaders is identical in all clinical isolates analyzed. However, additional studies are necessary to determine whether the variation in gp48 leader from these clinical isolates alters the translational effects of the gp48 leader.

Although there is no direct evidence for the existence of the peptide product of uORF2, the conclusion that uORF2 coding information is essential for translational inhibition implicates this potential peptide as a mediator of the inhibitory effect. However, in cotransfection assays, uORF2 acts only in *cis* (Fig. 5). These results suggest that the peptide might inhibit translation only when still attached to the ribosome-mRNA complex. In other eukaryotic genes, nascent peptides encoded at the 5' end of proteins are known to influence gene expression. For example, the signal peptide at the amino terminus of a protein destined for translocation across the endoplasmic reticulum membrane mediates translational pausing (2). Another explanation for the apparent *cis*-only effect of uORF2 is that the free peptide may be very unstable and thus act only locally. A less likely possibility is that the uORF2 does act in *trans* but the *cis*-acting signal on the target transcript that is responsive to the peptide requires AUG2. Since pEQ430 lacks AUG2, uORF2 peptide would be unable to inhibit translation of pEQ430 transcripts and we would not detect inhibition in our cotransfection assay (Fig. 5). However, this explanation could not account for absence of *trans*-acting inhibition of expression from the missense mutants in the experiments shown in Fig. 2 and 3, since the uORF2 peptide would be expressed from gp48 transcripts encoded by the viral genome introduced into these cells by infection.

Translational regulation is known to depend on the coding information of a uORF for only two other eukaryotic genes. The 25-codon uORF in the yeast *cpa-1* gene is required for inhibition of Cpa-1 expression in the presence of arginine (31). The *S*-adenosylmethionine decarboxylase gene in lym-

phocytes is similarly regulated by a six-codon uORF (13). Sequence comparisons fail to reveal common features of the predicted product of these two uORFs and the gp48 uORF2. In each case, the amino acid coding information of the uORF is required for inhibition of downstream translation, yet the signal appears to act only in *cis*. Deletions of the uORF termination codons in both gp48 (Fig. 1) and *S*-adenosylmethionine decarboxylase transcripts, which extend the uORFs by only one or a few codons, inactivate the inhibitory signal. Coupled with the preservation of the signal in wobble mutants, these results demonstrate that the inhibitory effect is not caused by a peculiar codon usage in the uORF. These striking similarities among these inhibitory signals suggest that a common mechanism may regulate expression of these three genes and some of the other eukaryotic genes expressing transcripts containing uORFs.

Because gp48 uORF2 sequences near and including the termination codon are critical for inhibition by the uORF2 signal, we propose that the nascent peptide product of uORF2 acts by preventing efficient termination of translation or release of the ribosome from the mRNA. The ribosome-peptide complex may create a barrier to translation of the downstream ORF of the gp48 transcripts. An analogous inhibitory effect on downstream translation by slow translation of an upstream ORF is thought to occur during translation of simian virus 40 16S and reovirus sigma-1 hemagglutinin-p14 dicistronic transcripts (1, 7, 24). Although it is unknown whether regulation at termination of translation occurs in eukaryotes, termination is a slow step (6, 32) and is regulated in some prokaryotic genes (8). We expect that further investigations into the mechanism of gp48 uORF2-mediated inhibition will disclose fundamental mechanisms of eukaryotic translation.

ACKNOWLEDGMENTS

We thank Alan Hinnebusch, David Morris, and Michael Katze for helpful discussions, Timmothy Dellitt for technical assistance, Stephanie Child for critical review of the manuscript, and Bettina Fries and Beverly Torok-Storb for DNA from CMV clinical strains.

This work was supported by Public Health Service grant AI26672 from the National Institutes of Health. M.R.S. was supported by NIH Academic Pediatric training grant T32-HD07233.

REFERENCES

1. Barkan, A., and J. E. Mertz. 1984. The number of ribosomes on simian virus 40 late 16S mRNA is determined in part by nucleotide sequence of its leader. *Mol. Cell. Biol.* 4:813-816.
2. Bernstein, H. D., M. A. Poritz, K. Strub, P. J. Hoben, S. Brenner, and P. Walter. 1989. Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature (London)* 340:482-486.
3. Biegalka, B. J., and A. P. Geballe. 1990. Translational inhibition by cytomegalovirus transcript leaders. *Virology* 177:657-667.
4. Chang, C.-P., D. H. Vesole, J. Nelson, M. B. A. Oldstone, and M. F. Stinski. 1989. Identification and expression of a human cytomegalovirus early glycoprotein. *J. Virol.* 63:3330-3337.
5. Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169, p. 125-169. *In* J. K. McDougall (ed.), *Cytomegaloviruses*. Springer-Verlag, New York.
6. Doohan, J. P., and C. E. Samuel. 1992. Biosynthesis of reovirus-specified polypeptides: ribosome pausing during the translation of reovirus S1 mRNA. *Virology* 186:409-425.
7. Fajardo, J. E., and A. Shatkin. 1990. Translation of bicistronic

- viral mRNA in transfected cells: regulation at the level of elongation. *Proc. Natl. Acad. Sci. USA* **87**:328–332.
8. Gavini, N., and L. Pulakat. 1991. Role of ribosome release in the basal level of expression of the *Escherichia coli* gene *pheA*. *J. Gen. Microbiol.* **137**:679–684.
 9. 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.
 10. 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.
 11. 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.
 12. 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.
 13. Hill, J. R., and D. R. Morris. 1993. Cell-specific translational regulation of S-adenosylmethionine decarboxylase mRNA: dependence on translation and coding capacity of the *cis*-acting upstream open reading frame. *J. Biol. Chem.* **268**:726–731.
 14. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283–292.
 15. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* **108**:229–241.
 16. Kozak, M. 1990. A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Exp.* **1**:111–115.
 17. Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* **115**:887–903.
 18. Miller, P. F., and A. G. Hinnebusch. 1989. Sequences that surround the stop codons of upstream open reading frames in GCN4 mRNA determine their distinct functions in translational control. *Genes Dev.* **3**:1217–1225.
 19. Mocarski, E. S. 1991. Evidence for posttranscriptional modulation of cytomegalovirus gene expression, p. 287–299. *In* E. Wagner (ed.), *Herpesvirus transcription and its regulation*. CRC Press, Boca Raton, Fla.
 20. Mocarski, E. S., M. Bonyhadi, S. Salimi, J. M. McCune, and H. Kaneshima. 1993. Human cytomegalovirus in the SCID-hu mouse: thymic epithelial cells are prominent targets of viral replication. *Proc. Natl. Acad. Sci. USA* **90**:104–108.
 21. Plotkin, S. A., S. E. Starr, H. M. Friedman, K. Brayman, S. Harris, S. Jackson, N. B. Tustin, R. Grossman, D. Dafoe, and C. Barker. 1991. Effect of Towne live virus vaccine on cytomegalovirus disease after renal transplant. *Ann. Intern. Med.* **114**:525–531.
 22. Schleiss, M., C. R. Degnin, and A. P. Geballe. 1991. Translational inhibition mediated an upstream reading frame within the cytomegalovirus gp48 gene transcript leader. *J. Virol.* **65**:6782–6789.
 23. Sedman, S. A., G. W. Gelembiuk, and J. E. Mertz. 1990. Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located very close to the 5' cap. *J. Virol.* **64**:453–457.
 24. Sedman, S. A., P. J. Good, and J. E. Mertz. 1989. Leader-encoded open reading frames modulate both the absolute and relative rates of synthesis of the virion proteins of simian virus 40. *J. Virol.* **63**:3884–3893.
 25. Simmons, P., K. Kaushansky, and B. Torok-Storb. 1990. Mechanisms of cytomegalovirus-mediated myelosuppression: perturbation of stromal cell function versus direct infection of myeloid cells. *Proc. Natl. Acad. Sci. USA* **87**:1386–1390.
 26. 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.
 27. Spaete, R. R., R. M. Thayer, W. S. Probert, F. R. Masiarz, S. H. Chamberlain, L. Rasmussen, T. C. Merigan, and C. Pachl. 1988. Human cytomegalovirus strain Towne glycoprotein B is processed by proteolytic cleavage. *Virology* **167**:207–225.
 28. Stamminger, T., E. Puchtler, and B. Fleckenstein. 1991. Discordant expression of the immediate-early 1 and 2 gene regions of human cytomegalovirus at early times after infection involves posttranscriptional processing events. *J. Virol.* **65**:2273–2282.
 29. Stenberg, R. M., A. S. Depto, J. Fortney, and J. A. Nelson. 1989. Regulated expression of early and late RNAs and proteins from the human cytomegalovirus immediate-early gene region. *J. Virol.* **63**:2699–2708.
 30. Stinski, M. F. 1990. Cytomegalovirus and its replication, p. 1959–2010. *In* B. N. Fields and D. M. Knipe (ed.), *Virology*, vol. 2. Raven Press, New York.
 31. Werner, M., A. Feller, F. Messenguy, and A. Pierard. 1987. The leader peptide of yeast gene *cpa1* is essential for the translational repression of its expression. *Cell* **49**:805–813.
 32. Wolin, S. L., and P. Walter. 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J.* **7**:3559–3569.
 33. Wright, D. A., and D. H. Spector. 1989. Posttranscriptional regulation of a class of human cytomegalovirus phosphoproteins encoded by an early transcription unit. *J. Virol.* **63**:3117–3127.