The Major Open Reading Frame of the \( \beta 2.7 \) Transcript of Human Cytomegalovirus: In Vitro Expression of a Protein Posttranscriptionally Regulated by the 5' Region

Giovanna Bergamini, Marko Reschke, Maria Concetta Battista, Maria Cristina Boccuni, Fabio Campanini, Alessandro Ripalti and Maria Paola Landini

The Major Open Reading Frame of the β2.7 Transcript of Human Cytomegalovirus: In Vitro Expression of a Protein Posttranscriptionally Regulated by the 5′ Region

GIOVANNA BERGAMINI,* MARKO RESCHEK,† MARIA CONCETTA BATTISTA, MARIA CRISTINA BOCCUNI, FABIO CAMPANINI, ALESSANDRO RIPALTI, AND MARIA PAOLA LANDINI

Department of Clinical and Experimental Medicine, Division of Microbiology, University of Bologna, St. Orsola Hospital, Bologna, Italy

Received 11 May 1998/Accepted 10 July 1998

β2.7 is the major early transcript produced during human cytomegalovirus infection. This abundantly expressed RNA is polysome associated, but no protein product has ever been detected. In this study, a stable peptide of 24 kDa was produced in vitro from the major open reading frame (ORF), TRL4. Following transient transfection, the intracellular localization was nuclear and the expression was posttranscriptionally inhibited by the 5′ sequence of the transcript, which harbors two short upstream ORFs.

The major early transcript of HCMV. Among β class genes of human cytomegalovirus (HCMV), an unspliced polyadenylated RNA of 2.7 kb originates within the two inverted repeats flanking the long unique segment (8, 39) (Fig. 1a). The two copies of the β2.7 transcript in the viral DNA each have one open reading frame (ORF), named TRL4 or IRL4 (EMBL accession no. X17403). β2.7 is the most abundant transcript, representing more than 20% of the total viral mRNA made during infection (17, 28). Its promoter element, referred to as the β2.7 promoter, is contained within a region beginning 213 bp upstream from the start site of transcription and has homologies to known transcription factor-binding sites (20, 38). This promoter is transactivated by immediate-early 1 and 2 gene products of HCMV, but other viral factors are necessary for its full, high-level expression (19). Starting from 4 h postinfection this transcript accumulates progressively throughout the replication cycle; it shows maximal amplification at between 8 and 14 h (29).

Following infection of nonpermissive cells, the β2.7 transcript seems to be exclusively confined to the nucleus (40). Nevertheless, during productive infection it is predominantly localized in the cytoplasm and is associated with the polysomes (25, 39). Although this localization pattern is consistent with an active translation during productive infection, no specific translation product has been detected so far (14), supporting an alternative functional hypothesis in which the RNA itself might have some regulatory role during infection (30).

In addition to TRL4, which is 513 nucleotides (nt) long (14) and is here also referred to as ORF3, two short upstream ORFs (uORFs), ORF1 and ORF2, have been identified in the sequence of the β2.7 transcript. ORF1 is located 81 nt from the transcription start site, and its 24-nt sequence is conserved in both the Towne and AD169 strains (5, 13). In contrast, ORF2 differs considerably in the two laboratory strains. In Towne it starts 20 nt downstream from the end of ORF1 and is 18 nt long, while in AD169 it starts 34 nt downstream from ORF1 and is composed of 108 nt.

Previous analyses, using a transient-transfection assay with lacZ as an indicator gene, have identified regulatory domains within the 5′ leader of the β2.7 transcript (1, 13). These studies demonstrated the existence of an inhibitory cis-acting signal which operates at a posttranscriptional level by repressing translation from the downstream reporter gene. This repression also seemed to alter the kinetics of expression during the infection cycle. The sequence causing this effect required an intact ORF1 and 32 downstream nucleotides including the AUG codon of ORF2 (11).

mRNAs containing one or more short uORFs have been characterized for both viral and cellular systems (21, 22). In some cases the AUG codons of these uORFs appear to negatively regulate downstream translation when they are recognized as valid start codons by eukaryotic ribosomes (7, 15, 16, 33, 36). According to Kozak’s model, the inhibitory influence of these uORFs might therefore be due either to the provocation by the short intercistronic space of an inefficient reinitiation at subsequent internal start sites or to the complete dissociation of the ribosome from the mRNA after efficient translation of the uORF (24). Alternatively, the nascent peptide encoded by the uORF could interact with the ribosome and prevent its disassembly, thus blocking the scanning mechanism, as proposed by Geballe and Morris (10, 12).

In this study we investigated the ORFs of the β2.7 transcript. A specific product of approximately 24 kDa was synthesized following eukaryotic expression of TRL4 in a cell-free assay; this is the first evidence that a stable protein can be produced in vitro from this sequence. Following transient transfection of various cell types, the TRL4 product, pTRL4, tagged with an immunoreactive epitope, FLAG, was found to be localized mainly within intranuclear bodies (the nucleoli). Importantly, TRL4 is largely conserved in HCMV strains, consistent with its predicted role in viral infection. The expression of this protein seemed to be highly regulated at a posttranscriptional level by the 5′ leader sequence of its mRNA, which bears the two short uORFs. This study was a preliminary assessment of the putative protein coded for by the β2.7 transcript, conducted with a view to carrying out experiments to define TRL4 expression in...
TRL4 codes in vitro for a protein. TRL4, coding for a putative product of 19.6 kDa, was cloned into the vector pcDNA3 (Invitrogen), under the transcriptional control of the T7 promoter and the major immediate-early promoter/enhancer element of HCMV (Fig. 1b). The resulting plasmid, pTo/ORF3, was subjected to an in vitro transcription and translation assay in rabbit reticulocyte lysates (RRL) (TNT System; Promega), and a stable product of approximately 24 kDa was detected (Fig. 2, lane 2). Although the expression in RRL was not very efficient, these data indicate that the TRL4 start codon is recognized by the eukaryotic translational machinery. In prokaryotic systems, despite using different fusion partners, we could not obtain a stable product with a full-length peptide derived from TRL4. Since no polyclonal antibody was available, an immunoreactive octapeptide, termed FLAG (indicated in construct names by an asterisk), was fused to the carboxy terminus of pTRL4, yielding the construct pTo/ORF3*.

The tagged protein exhibited the expected molecular mass of about 25 kDa, indicating that addition of this epitope did not affect the stability of the product (Fig. 2, lane 3).

Human astrocytoma (U373-MG) cells, which are permissive for HCMV replication, were transiently transfected with the construct pTo/ORF3*. Indirect immunofluorescence with the anti-FLAG monoclonal antibody (MAb) M2 (Eastman Kodak Company) revealed the expression of a specific product showing a characteristic intracellular localization. The protein accumulated in subnuclear structures, which were demonstrated to correspond to nucleoli by phase-contrast microscopy (Fig. 3a and c). Fusion protein was not detected within the nucleoplasm, and approximately 50% of positive cells displayed a diffuse cytoplasmic staining in addition to nucleolar staining (Fig. 3b and d). This particular pattern was also observed in human embryonal lung fibroblasts and monkey kidney (COS7) cells, implying the existence of similar mechanisms for targeting pTRL4 in different cell types.

Proteins smaller than 40 to 60 kDa like pTRL4 can diffuse freely through the nuclear pore complexes (34). However, the observed nuclear localization was very distinctive, suggesting a specific transport rather than a passive diffusion. Analysis of the deduced primary structure of pTRL4 highlighted the presence of a short stretch of basic amino acids (KRVRKRKK; amino acids [aa] 88 to 94) closely related to the prototype nuclear localization signal of the simian virus 40 large T antigen (18). Moreover, this domain resembles a bipartite nuclear
localization signal with a second cluster of basic amino acids (RRIQSRR; aa 107 to 113) located at a distance of 12 residues (35). Furthermore, a positively charged region (RRIQSRRFP TRENRKT; aa 107 to 124) having homologies with reported nuclear localization signals (6, 27, 37) is present. This region is also an arginine-rich motif known to display RNA-binding activities (2). Furthermore, both the nuclear localization and the nucleolar localization of pTRL4 could be determined by specific motifs present in its amino acid sequence.

Two putative N glycosylation sites (aa 119 to 121 and aa 141 to 143) have been previously identified (14). Since we detected TRL4 product within the nucleus, it seems unlikely that it enters the endoplasmic reticulum, where these signals could be processed by membrane-associated glycosyl transferases.

mRNA sequences were derived from the strain AD169. In performing DNA sequencing (with Sequenase 2.0; Amersham) of TRL4 from Towne, we found three nucleotide transitions, three at positions 4429, 4415, and 4346 and an additional one at position 4450 (13). Furthermore, there were differences in the nucleotide sequences, confirming the results obtained with the cell-free assay.

The same effect was obtained with sequences from both Towne and AD169, which show identical uORF1s but considerably different uORF2s, while the entire secondary structure of the RNA is most likely conserved. This may suggest that inhibition depends predominantly on translation of uORF1. According to Geballe and Morris’s model (12), ribosome stalling (26) can explain the marked inhibitory effect even if the AUG codon of uORF1 is seldom recognized because of its suboptimal consensus context (4). Alternatively, a particular secondary structure of the mRNA could account for the enhancement of the initiation efficiency (23).

While the presence of the 5′ leader sequence in vivo seems unlikely to explain the expression of TRL4 in the in vitro assay, inhibition was complete in transfected cells. This difference could be attributed to the thresholds of sensitivity in these two systems. As suggested by Cao and Geballe, the impact of a stalled ribosome could be more evident in cells than in cell-free systems. The low expression of TRL4 detected in RRL could be explained by a leaky scanning mechanism, by which some ribosomes fail to initiate translation at uORF1 and continue scanning along the messenger (24).

Since an RNA synthesized in vitro or from a designed construct in transfection experiments cannot reflect what occurs in vivo, our findings to not answer the question of whether the β2.7 messenger expresses a protein during infection. In fact, examples of in vitro-expressed viral proteins have been reported in the literature (9), but confirmation in vivo experiments has not been obtained.
On the basis of our findings and previous reports showing that the expression of an ORF downstream from the 5' leader of the b2.7 transcript is temporarily regulated during viral replication (13), we suggest that the constitutive inhibitory effect we observed could be partially or completely released in lytic infection or latent and persistent infection. Such a post-transcriptional regulation of gene expression is not unusual, since some key eukaryotic genes, oncoproteins, receptors, and transcription factors that are constitutively repressed by the 5' leader region (20) can be modulated by physiological conditions and during cellular differentiation (31). Therefore, translation of TRL4, most likely inhibited by the uORF's, might occur under particular conditions related to cell type and/or cell differentiation.

We thank A. P. Geballe (Fred Hutchinson Cancer Research Center, Seattle, Wash.) and E. S. Mocarski (Stanford University School of Medicine, Stanford, Calif.) for reviewing the manuscript, R. Luhrmann and K. Radsak (Philipps University, Marburg, Germany) for valuable discussions, and M. La Placa (Bologna, Italy) for encouragement. We thank Luisa Bertacchi for excellent technical assistance.

FIG. 3. Intracellular localization of pTRL4 in U373-MG cells. Following transfection with pAD/ORF3+, cells were fixed and probed with the anti-FLAG MAb M2. The fluorescence signal is always localized to the nuclei (a and b), as verified by comparison with the phase-contrast images (c and d). In approximately 50% of the positive cells, some positivity was also detectable throughout the cytoplasm (b).

FIG. 4. Northern blot analysis of total RNA extracted from COS7 cells transfected with the control vector (lane 1) and with plasmids pTo/ORF3+ (lane 2), pTo/ORF1-2-3+ (lane 3), and pAD/ORF1-2-3+ (lane 4). Detection of glyceraldehyde-3-phosphate dehydrogenase (G3PDH) transcript confirmed the presence of equal quantities of cellular RNA. Autoradiography showed that transcripts produced from all of the three plasmids were present in similar amounts in transfected cells (lanes 1, 2, and 3).
This work was partially supported by the Italian Ministry of University and Scientific Research, 60% and 40% grant; by the ECC Project Bio-med 2; and by the AIDS Project of the Italian Ministry of Public Health.

REFERENCES


