JOURNAL OF VIROLOGY, Mar. 1988, p. 1076–1079
0022-538X/88/031076-04$02.00/0

Vol. 62, No. 3

Specific Interactions between Transcription Factors and the Promoter-Regulatory Region of the Human Cytomegalovirus Major Immediate-Early Gene

PETER GHAZAL, HENRYK LUBON, AND LOTHAR HENNIGHAUSEN*

Laboratory of Biochemistry and Metabolism, National Institute of Diabetes and Digestive and Kidney Diseases, Building 10, Room 9N-113, Bethesda, Maryland 20892

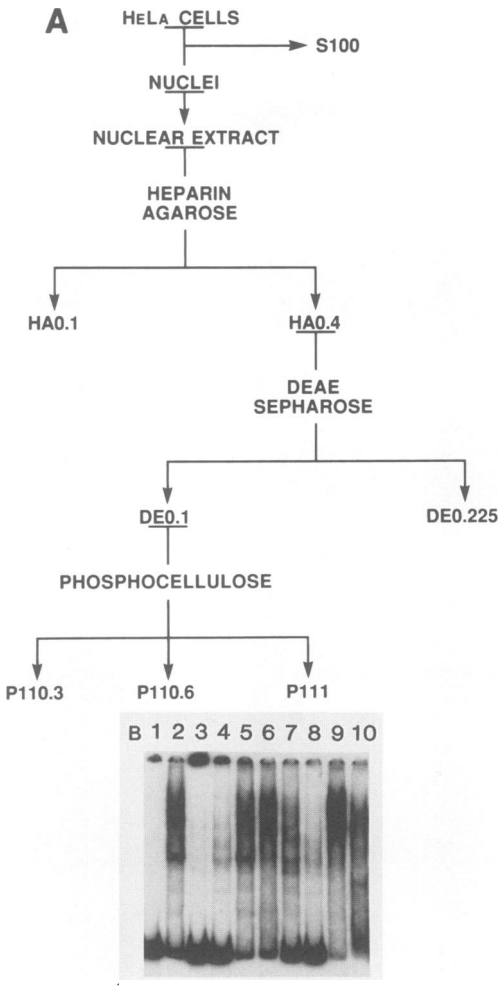
Received 12 August 1987/Accepted 11 November 1987

Repeat sequence motifs as well as unique sequences between nucleotides −150 and −22 of the human cytomegalovirus immediate-early 1 gene interact in vitro with nuclear proteins. We show that a transcriptional element between nucleotides −91 and −65 stimulated promoter activity in vivo and in vitro by binding specific cellular transcription factors. Finally, a common sequence motif, (T)TGG/AC, present in 15 of the determined binding sites suggests a particular class of nuclear factors associated with the immediate-early 1 gene.

The regulatory region of the human cytomegalovirus (HCMV) immediate-early 1 (IE1) gene contains several different functional domains. These include polymerase II promoter, enhancer, and repressor activities. The function of these regulatory domains has been shown in in vitro and in vivo expression systems (2, 9, 11, 12). A remarkable feature of the IE1 regulatory region is its complex repetitive sequence composition. Our previous experiments have shown that the enhancer domain, between nucleotides −524 and −185, interacts with specific transcription factors which recognize several of the repeated sequence motifs as well as a unique sequence (4). However, it was not clear and cannot be assumed from this study that repeat sequences, and in particular sequences outside the repeats, downstream of −185 bind specific transcription factors. Therefore, we have investigated the ability of DNA-binding proteins in transcriptionally active HeLa cell nuclear extract to interact with specific target sequences in the IE1 promoter-regulatory region.

Nuclear and S100 extracts from HeLa cells were prepared (3), and the nuclear extract was chromatographically fractionated as shown in Fig. 1A. We determined specific DNA-protein interactions by performing a mobility shift assay as described (8). The fragment (E3) between nucleotides −185 and −14 incubated with crude nuclear extract resulted in several distinct nucleoprotein complexes, suggesting the binding of multiple factors (Fig. 1B, lane 2). The cytoplasmic (S100) fraction did not contain significant binding activity (Fig. 1B, lane 3), confirming that the binding proteins were of nuclear origin. The various protein fractions obtained after fractionation of the crude nuclear extract also resulted in a complex band shift pattern (Fig. 1B). To delineate precisely at the nucleotide level sites of protein-DNA interaction, DNase I protection analyses were performed (Fig. 2) as described (4, 10). In two respects, these data are in agreement with the mobility shift experimental results. First, protected regions were observed in only those column fractions which developed specific band shift patterns. Second, multiple regions of protection were observed, which can explain the complex band shift pattern.

Four principal regions of protection were observed on the HCMV IE1 promoter-regulatory sequence between nucleotides −50 and −146 (Fig. 2; see Fig. 4). Interestingly, clear protection for the 18-base-pair (bp) repeat between nucleotides −172 and −155 was not observed, while that between



* Corresponding author.

nucleotides −109 and −92 was protected (Fig. 2). Immediately flanking and partly overlapping the 5′ end of this latter repeat was a strong binding site from nucleotides −120 to −107. The binding activity of this site could not be competed for by the simian virus 40 or human immunodeficiency virus enhancer elements (unpublished), which have sequence homologies with the 18-bp repeats (2, 4). This observation supports the suggestion that this sequence represents a unique binding site. In proximity to this binding site (between nucleotides −146 and −129) is a 19-bp repeat element that was clearly protected, in agreement with our previous observations for the other 19-bp repeats (4). All of these described binding sites were situated within the functionally characterized enhancer-regulatory domain (2, 11). A redundant form of the 19-bp palindromic repeat between nucleotides −70 and −58 (83% homology), in which the symmetry has been disrupted, was located within a weakly protected region. Additional confirmation of a factor capable of binding to this site was obtained after enrichment of the factor with a single-stranded DNA-cellulose column (data not shown). The binding property (degree of protection) of this site was clearly different from that of the other 19-bp repeats (Fig. 2) (4). We cannot at present distinguish between the possibilities that this factor is identical to or different from the factor binding to other 19-bp repeat elements. It is interesting that DNase I protection assays with mammary gland extract (see below) did not show protection of the 19-bp repeat elements within the characterized enhancer domain (2) but did show protection for the −70 to −50 region (Fig. 2) (unpublished). Binding sites further downstream of −50 were not detected with HeLa cell extract in the DNase I protection experiments. However, in vitro transcription competition experiments with the E3 fragment as specific competitor (see below) (4) indicated that additional *trans*-acting factors possibly interacted with sequences downstream of −50.

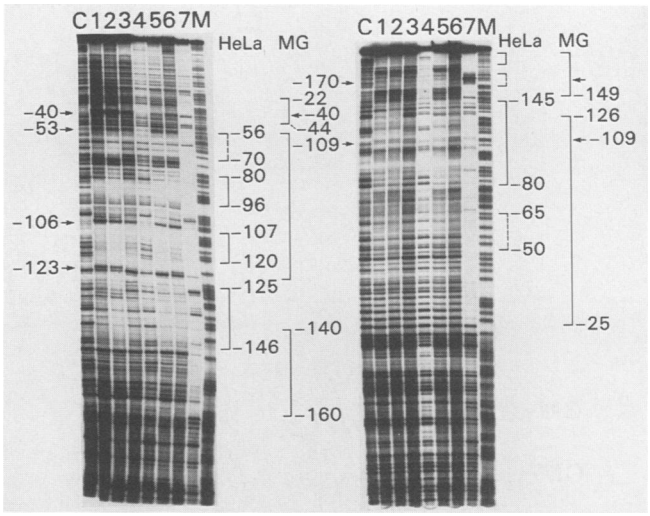


FIG. 2. Binding of nuclear proteins to the IE1 promoter as determined by DNase I protection. The *Nco*I-*Xba*I (−220 to +55) fragment was derived from pCMV(−524)CAT (4) and labeled at the 5′ ends with T4 polynucleotide kinase and [γ-^{32}P]ATP. The *Xba*I site at +55 was created by the addition of a polylinker to the *Hpa*II site at +55 of the IE1 gene. Protection of the upper (left panel) and lower (right panel) strand of the *Nco*I (−220) to *Xba*I (+55) fragment is shown. Lanes M, A/G Maxam and Gilbert sequencing reaction of the respective fragment. Lanes C, DNase I digestion products of the DNA fragments in the presence of bovine serum albumin. Lanes 1, 15 μl of HA0.4; lanes 2, 25 μl of DE0.1; lanes 3, 25 μl of DE0.225; lanes 4, 25 μl of P110.3; lanes 5, 25 μl of P110.6; lanes 6, 25 μl of P111; lanes 7, 6 μl of crude nuclear extract from lactating rat mammary glands (MG). Nuclear extract from the rat mammary gland was prepared as described (6). Arrows mark the positions of protein-induced DNase I-hypersensitive sites. Brackets indicate protected regions. Numbers refer to the nucleotide position with respect to the start site of transcription.

FIG. 1. (A) Schematic representation of the fractionation of crude nuclear extract. The buffer used in all chromatographic steps was buffer A (25 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesultonic acid, pH 7.9], 0.1 mM EDTA, 10% glycerol, 1 mM DTT, and KCl as indicated). HeLa cell nuclear extract (80 mg) was loaded on a 6.5-ml heparin-agarose (Bio-Rad) column equilibrated with buffer A and 0.1 M KCl. The column was washed with 0.1 M KCl in buffer A, followed by a step wash with 0.4 M KCl in buffer A. The proteins eluted with 0.4 M KCl (HA0.4 fraction) were dialyzed against 0.1 M KCl in buffer A. The HA0.4 fraction was applied to a 3.5-ml DEAE-Sepharose CL-6B (Pharmacia) column equilibrated with 0.1 M KCl in buffer A. The flowthrough fraction (DE0.1 fraction) was collected, and thereafter the column was washed with 0.1 M KCl in buffer A. The column was then step washed with 0.225 M KCl in buffer A, and the collected fraction (DE0.225 fraction) was dialyzed against 0.1 M KCl in buffer A. The DE0.1 fraction was applied to a 3.5-ml phosphocellulose (P11; Whatman) column eqilibrated with 0.1 M KCl in buffer A and washed with 0.1 M, 0.3 M, 0.6 M, and 1 M KCl, successively. Each step-washed fraction (fractions P110.3, P110.6, and P111) was dialyzed against 0.1 M KCl in buffer A. (B) Mobility shift complexes formed with nuclear proteins and sequences between nucleotides −185 and −14. The labeled E3 fragment (0.2 ng) and 1 μg of poly(dI-dC) were incubated with 1 μg of different protein fractions. Lane 1, probe in the absence of protein; lane 2, crude nuclear extract; lane 3, S100 fraction; lane 4, HA0.1 fraction; lane 5, HA0.4 fraction; lane 6, DE0.1 fraction; lane 7, DE0.225 fraction; lane 8, P110.3 fraction; lane 9, P110.6 fraction; lane 10, P111 fraction.

A plausible explanation for these apparently conflicting observations is that such *trans*-acting factors are a minor component of the HeLa cell nuclear extract and their detection is made possible by the extreme sensitivity of the transcription assay. In this connection, we asked whether these sequences had the ability to interact with nuclear proteins prepared from different sources. We found that mammary gland nuclear extract, similar to the HeLa cell extract, protected sequences between nucleotides −146 and −50. The exact boundaries of the protected regions were not the same as in HeLa extract, and additional hypersensitive sites were observed (Fig. 2, lanes 7). Moreover, HCMV IE1 promoter sequences between −50 and −22 were clearly protected with mammary gland extract. This observation supports the suggestion that these sequences can interact with specific nuclear factors.

Since we detected different binding properties between those 19-bp repeat elements within the characterized enhancer and that from −73 to −54, we asked whether sequences in this region could activate transcription from the promoter. Therefore, we tested two plasmids that specifically truncate the IE1 5′-flanking region at positions −91 [pCMV(−91)CAT] and −65 [pCMV(−65)CAT] in in vivo and in vitro expression systems. The template pCMV(−65) CAT directed a low but detectable level of transcription upon transient expression in HeLa cells, whereas pCMV(−91)CAT resulted in a 10-fold higher level of tran-

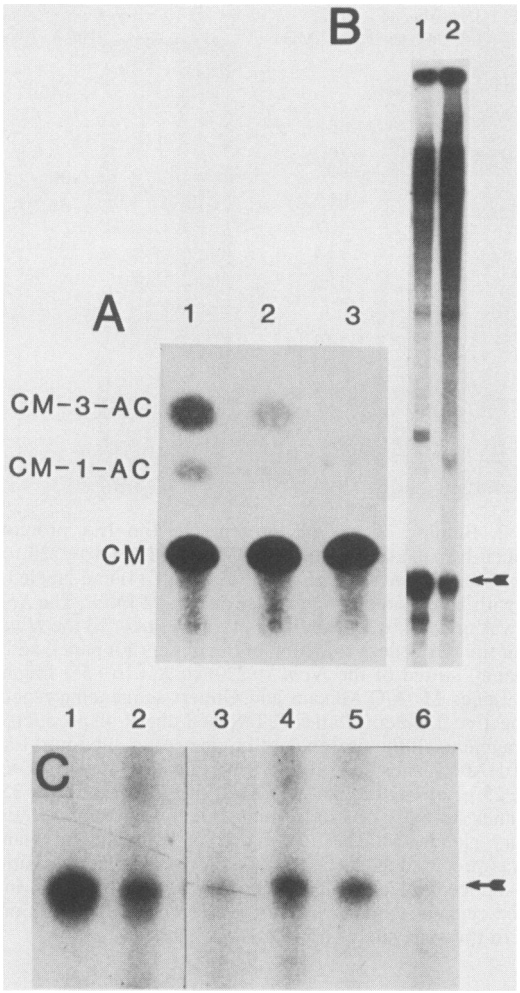


FIG. 3. Expression of the HCMV IE1 promoter in vivo and in vitro. Plasmids pCMV(−91)CAT and pCMV(−65)CAT were obtained from R. Ruger and B. Fleckenstein. These plasmids contain 5′ deletion fragments of the HCMV (strain AD169) IE1 gene linked to the chloramphenicol acetyltransferase (CAT) gene. Transfections were performed by the DEAE-dextran method (1). Cell extract was prepared, and CAT activity was determined by the method of Gorman et al. (5). (A) CAT activity directed by plasmids pCMV(−91)CAT (lane 1), pCMV(−65)CAT (lane 2), and pCATO (lane 3) after transient expression in HeLa cells. Chloramphenicol (CM) and its acetylated forms 1-acetate (CM-1-AC) and 3-acetate (CM-3-AC) were detected by autoradiography. Quantitation of CAT activity was determined by scintillation counting. (B) RNA synthesized in vitro from the templates pCMV(−91)CAT (lane 1) and pCMV(−65)CAT (lane 2), both of which had been truncated with *Pvu*II. In vitro transcription reactions were performed as described (4). The 220-nucleotide transcript marked by an arrow initiated at the IE1 mRNA start site used in vivo and terminated at the *Pvu*II site in the *cat* gene. Transcriptional activity was determined by densitometric scanning of autoradiographs. (C) RNA synthesized from templates pCMV(−91)CAT (lanes 1 to 3) and pCMV(−65)CAT (lanes 4 to 6) after preincubating the nuclear extract with competitor DNA. The conditions of the in vitro transcription competition reactions were the same as described previously (4). A 15-fold excess of the following competitors was used: pUC19 cut with *Hae*III (lanes 1 and 4), fragment E1 covering IE1 enhancer sequence between nucleotides −524 and −311 (lanes 2 and 5), and fragment E3 covering IE1 enhancer-promoter sequences between nucleotides −185 and −14 (lanes 3 and 6). The DNA fragments E1 and E3 were gel purified from their respective plasmids, pE1 and pE3 (4), after cleaving with the restriction endonucleases *Eco*RI and *Hin*dIII. The arrow marks the IE1 promoter-specific transcript.

scription (Fig. 3A). In agreement with this, runoff in vitro transcription analysis of these plasmids resulted in approximately 5- to 10-fold more specific transcript molecules from pCMV(−91)CAT than from the pCMV(−65)CAT template (Fig. 3B). Thus, in two indepedent assay systems, a promoter-proximal transcription regulatory sequence between −91 and −65 of the IE1 gene was functional in vivo and in vitro.

To determine whether the transcriptional stimulation mediated by these sequences involved the binding of *trans*-acting factors, we used an in vitro transcription competition assay (4). The same extract and DNA template preparations were assayed in the presence of nonspecific DNA and fragment E3 (−185 to −14) as specific competitor. The activity of the IE1 promoter in pCMV(−91)CAT and pCMV(−65)CAT was reduced below the basal level of transcription in the presence of specific competitor (Fig. 3C, lanes 1, 3, 4, and 6). This observation suggests that the transcriptional activities of both pCMV(−91)CAT and pCMV(−65)CAT involve the interaction of specific *trans*-acting factors. To examine whether the stimulatory activity of sequences present in pCMV(−91)CAT was associated with enhancer elements, competition assays were performed with an IE1 enhancer-specific fragment. Transcriptional stimulation of the IE1 promoter in pCMV(−91)CAT but not in pCMV(−65)CAT was selectively reduced by the enhancer sequences between −524 and −311 (Fig. 3C, lanes 2 and 5). The reduction of transcriptional activity of pCMV(−91)CAT to about the level observed with pCMV(−65)CAT suggests that a common or closely related *trans*-acting factor(s) interacted with the IE1 enhancer and upstream promoter sequences. Furthermore, this suggests that the enhancer domain extends to nucleotides between −91 and −65.

Three lines of evidence support the conclusion that the −70 to −50 region is important for stimulating transcription from the IE1 promoter. The E1 fragment contains two 19-bp palindromic repeats which showed homology with the sequence between −70 and −50 and competed effectively for transcriptional stimulation in the in vitro competition experiments; the transcriptionally active HeLa cell nuclear extract contained a factor(s) which specifically interacted with the −70 to −50 region; and the intact 5′ half of the protected palindromic sequence was deleted in pCMV(−65)CAT, which showed reduced transcriptional activity in vivo and in vitro (Fig. 4).

In summary, on the basis of this work and our previous experiments (4), we are able to conclude the following. Most, but not all, of the repetitive sequences which have been implicated in the activity of the enhancer (2, 11) are targets for specific transcription factors. The extent of protection for a repetitive sequence varies within members of a class of repeat elements (e.g., 19-bp repeats) as well as between different classes of repeat elements (e.g., 17- and 18-bp repeats). We show that sequences outside the repeat elements, between nucleotides −445 to −425 and −120 to −107, were also targets for specific nuclear proteins, further increasing the order of complexity of this region. The 5′-end half of 15 of the binding sites thus far determined (4, 6, 7; this work) have a common sequence motif, (T)TG(G/A)C, which may suggest a particular class of nuclear factors associated with the HCMV IE1 gene.

We thank Rudiger Ruger and Bernhard Fleckenstein for providing plasmids pCMV(−91)CAT and pCMV(−65)CAT and William Jakoby for generous support and comments on the manuscript.

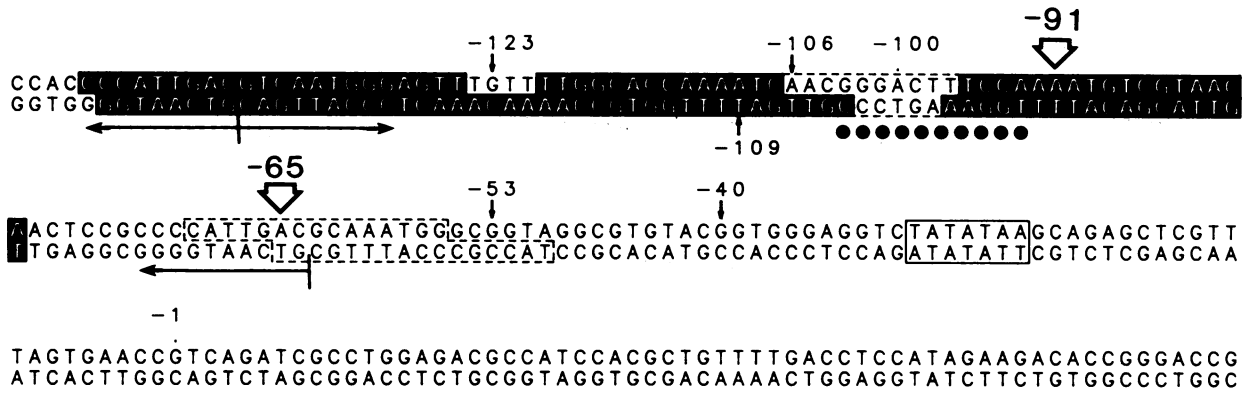


FIG. 4. Sites of protein-DNA interaction on the HCMV IE1 promoter. The DNA sequence of both strands is shown. Numbers represent the distance (in base pairs) upstream of the transcription start site (+1). Sequences protected from DNase I cleavage by HeLa cell nuclear proteins are marked by white letters on black background (strong protection) or by a hatched box (weak protection). Vertical arrows mark positions of enhanced protein-induced DNase I cleavage in vitro. Large open arrows indicate the 5′ endpoints of pCMV(−91)CAT and pCMV(−65)CAT. The horizontal arrow marks the 19-bp repeat element, and solid dots indicate the 18-bp repeat element (enhancer core).

LITERATURE CITED

1. **Banerji, J., L. Olson, and W. Schaffner.** 1983. A lymphocyte specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell **33**:729–740.
2. **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. Cell **41**:521–530.
3. **Dignam, J. D., R. M. Lebowitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
4. **Ghazal, P., H. Lubon, B. Fleckenstein, and L. Hennighausen.** 1987. Binding of transcription factors and creation of a large nucleoprotein complex on the human cytomegalovirus enhancer. Proc. Natl. Acad. Sci. USA **84**:3658–3662.
5. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
6. **Hennighausen, L., and B. Fleckenstein.** 1986. Nuclear factor 1 interacts with five DNA elements in the promoter region of the human cytomegalovirus major immediate-early gene. EMBO J. **5**:1367–1371.
7. **Jeang, K.-T., D. R. Rawlins, P. J. Rosenfeld, J. H. Shero, T. J. Kelly, and G. S. Hayward.** 1987. Multiple tandemly repeated binding sites for cellular nuclear factor 1 that surround the major immediate-early promoters of simian and human cytomegalovirus. J. Virol. **61**:1559–1570.
8. **Lubon, H., and L. Hennighausen.** 1987. Nuclear proteins from lactating mammary glands bind to the promoter of a milk protein gene. Nucleic Acids Res. **15**:2103–2121.
9. **Nelson, J. A., C. Reynolds-Kohler, and B. A. Smith.** 1987. Negative and positive regulation by a short segment in the 5′-flanking region of the human cytomegalovirus major immediate-early gene. Mol. Cell. Biol. **7**:4125–4129.
10. **Ohlsson, H., and T. Edlund,** 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. Cell **45**:35–44.
11. **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.
12. **Thomson, 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.