Cellular or Viral Protein Binding to a Cytomegalovirus Promoter Transcription Initiation Site: Effects on Transcription

MIMI P. MACIAS, LIPING HUANG, PHILIP E. LASHMIT, AND MARK F. STINSKI*

Department of Microbiology, College of Medicine, University of Iowa, Iowa City, Iowa 52242

We have previously shown that the IE2 protein of human cytomegalovirus (CMV) represses its own synthesis by binding to the major immediate-early promoter (M. P. Macias and M. F. Stinski, Proc. Natl. Acad. Sci. USA 90:707–711, 1993). The binding of a viral protein (IE2) and a cellular protein in the region of the transcription start site was investigated by site-specific mutational analysis and electrophoretic mobility shift assay. The viral protein and the cellular protein require different but adjacent core DNA sequence elements for binding. In situ chemical footprinting analysis of DNA-protein interactions with purified CMV IE2 protein or HeLa cell nuclear extracts demonstrated binding sites that overlap the transcription start site. The IE2 protein footprint was between bp −15 and +2, relative to the transcription start site, and the cellular protein was between bp −16 and +7. The ability of the unknown human cellular protein of approximately 150 kDa to bind the CMV major immediate-early promoter correlates with an increase in the level of transcription efficiency. Mutations in the core DNA sequence element for cellular protein binding significantly reduced the level of in vitro transcription efficiency. Mutations upstream and downstream of the core sequence moderately reduced the transcription efficiency level. Negative autoregulation of the CMV promoter by the viral IE2 protein may involve both binding to the DNA template and interference with the function of a cellular protein that binds to the transcription start site and enhances transcription efficiency.

Human cytomegalovirus (CMV) is a member of the herpesvirus family and infects 40 to 100% of the population. The CMV genome remains in a latent state in monocytes or granulocyte-macrophage precursor cells (14, 38). Infection of a healthy adult is frequently asymptomatic. However, CMV causes serious congenital birth defects and severe or fatal infections in immunocompromised hosts (10).

The cloned CMV enhancer-containing promoter is used in many laboratories for research purposes because of its extremely high transcription levels in a variety of primary and stable cell lines and in cell-free transcription systems (6, 37, 39). Upstream of the promoter is an enhancer region which contains repetitions of consensus binding sites for a number of cellular transactivating proteins such as CREB or ATF, NF-κB or Rel, and SP1 (reviewed in reference 36). The CMV major immediate-early promoter (MIEP) is responsible for producing the most abundant viral RNAs in infected host cells at immediate-early (IE) times after infection (43). The IE gene products, designated IE1 (491 amino acids) and IE2 (579 amino acids), are translated from alternatively spliced mRNAs (25, 31, 32). These proteins play a critical role in activating the cascade of viral gene expression leading to virus production (reviewed in reference 35).

The viral IE2 protein, which has been shown to be a potent transactivator of homologous and heterologous viral promoters as well as of several cellular promoters (8, 9, 20, 24, 28, 33), serves to negatively autoregulate the MIEP (2, 18, 23). Repression of the MIEP requires IE2 protein to bind to a *cis* repression signal (crs) located between the TATA box and the transcription start site (16, 17, 19). Minor groove contacts are essential for IE2 protein binding (17). Repression of MIEP transcription by IE2 protein is crs position dependent but orientation independent (2, 18, 23).

Characteristic of a core promoter element is an initiator sequence (1, 30), which usually overlaps or flanks the transcription start site. A consensus initiator sequence has been more precisely defined to contain the sequence Py Py A N T/A Py Py (11). This sequence is present in the CMV MIEP adjacent to the transcription initiation site (34). In TATA containing promoters, an initiator functions to enhance initiation of transcription two- to threefold (1, 30). Little is known about the cellular proteins that bind to initiator sequences. Cellular transcription factors E2F, TFII-I, USF, and YY-1 have been shown to bind to initiator-like sequences (4, 22, 26, 29, 40). However, it also has been proposed that none of the above proteins represent the true initiator binding protein (11). A TATA box-binding protein-associated factor (TAFII150) for polymerase II in *Drosophila* cells was proposed to be an initiator DNA-binding protein that contributes significantly to the promoter recognition process (7, 41, 42). dTAFII150 and other TATA box-binding protein-associated factors may modulate the stability of initiation complexes and govern transcriptional regulation.

In this report, we compare the binding of the viral IE2 protein with a nuclear protein(s) from HeLa or human foreskin fibroblast (HFF) cells. Binding of the viral IE2 protein, as reported previously (19), has a negative effect on in vitro transcription. Mutational analysis indicates that binding of the unknown human cellular protein(s) has a positive effect. The viral and cellular protein binding sites are adjacent and overlap. By identifying the positive- and negative-acting factors involved in regulating the extremely strong CMV MIEP, we hope to better understand the mechanisms which determine the level of gene expression and influence viral latency and reactivation.

MATERIALS AND METHODS

Plasmids. The construction of pCAT760 and pCAT760dlRE was described previously (18, 37). To construct the mutant plasmid pCAT760m5, the parent plasmid pCAT760dlRE was first cleaved with *Sal*I restriction endonuclease and filled in with dCTP and TTP with the Klenow fragment of *Escherichia coli* polymerase I. After *Sph*I enzyme cleavage, the DNA was incubated with T4

* Corresponding author. Phone: (319) 335-7792. Fax: (319) 335-9006. Electronic mail address: CMDMFS@VAXA.WEEG.UIOWA.EDU.

DNA polymerase to flush the single-stranded ends for subsequent blunt-end ligation. The mutant plasmids pGEMREm1, pGEMREm2, pGEMREm3, and pGEMREm4 were constructed by inserting the mutant 27-bp *Sst*I-*Hin*dIII fragments (m1, m2, m3, and m4) into pGEM-7zf(+) (Promega, Madison, Wis.). The mutant DNA fragments (m1, m2, m3, and m4) were generated by PCR with the sense primers 5′-TTGGAGCTCaggTAGTGAAC-3′ (m1) and 5′-TTGGAGC TCGTTTAGTGAgaaGTCAGATC-3′ (m2) and the antisense primers 5′-GAT AAGCTTCTAGAGATCcccCGGTTCAC-3′ (m3) and 5′-GATAAGCTTCTA GAcccCTGACGG-3′ (m4), respectively. The mutations in the primers are shown as lowercase letters. The M13-20 primer (Stratagene, La Jolla, Calif.) was used as the antisense primer for generating the m1 and m2 DNA fragments. The M13 reverse primer (Stratagene) was used as the sense primer for generating the m3 and m4 DNA fragments. Plasmids pGEMRE(wt) and pGEMREm5 were constructed by inserting the 27-bp *Sst*I-*Hin*dIII fragments from plasmids pCAT760 and pCAT760m5, respectively, into pGEM-7zf(+) (Promega). All PCR products were digested with *Sst*I and *Hin*dIII restriction endonucleases (New England Biolabs, Beverly, Mass.) and gel purified. All mutations were confirmed by dideoxynucleotide sequencing.

r-IE2 protein. The construction, expression, and purification of recombinant maltose-IE2 fusion protein (r-IE2) containing amino acids 290 to 579 of the viral IE2 protein has been described previously (19). An r-IE2 protein mutated in the putative C_2H_2 zinc finger motif (r-IE2HL) has also been described (19). The r-IE2552 protein is a modification of r-IE2 because of a stop codon at amino acid position 552. r-IE2552 lacks the carboxyl-terminal 28 amino acids.

Synthesis and labeling of DNA probes. The DNA probes used in electrophoretic mobility shift assays (EMSA) were synthesized by PCR. The plasmids pGEMRE(wt), pGEMREm1, pGEMREm2, pGEMREm3, pGEMREm4, and pGEMREm5 were the templates for PCR amplification of DNA fragments wt, m1, m2, m3, m4, and m5, respectively, with primer pairs. After digestion with *Mlu*I and *Cla*I restriction endonucleases (New England Biolabs), the 49-bp DNA fragments were gel isolated and end labeled with either [α-^{32}P]dCTP (6,000 Ci/mmol) (Amersham, Arlington Heights, Ill.) and the Klenow fragment of *E. coli* polymerase I or [γ-^{32}P]ATP (3,000 Ci/mmol) (Amersham) and T4 polynucleotide kinase (New England Biolabs). The DNA fragments and probes were quantitated by the ethidium dot method (27). To produce RET7H used in footprinting, the 27-bp *Sst*I-*Hin*dIII fragment of pCAT760 was subcloned into pBluescript KSII+ (Stratagene) and confirmed by dideoxynucleotide sequencing. PCR amplification with T7 and M13 reverse primers was followed by cleavage with *Hin*dIII restriction endonuclease and gel purification. The 59-bp fragment was filled in on the coding strand at the *Hin*dIII site with dATP and dGTP, [α-^{32}P]dCTP (3,000 Ci/mmol), and the Klenow fragment of *E. coli* polymerase I. The bromodeoxyuridine (BrdU)-incorporated RET7H probe, used for cross-linking experiments, was produced with the same primers and plasmid template used to produce RET7H DNA, but the reaction mixtures included BrdU in place of TTP during PCR amplification. REBH DNA used in footprinting experiments was synthesized from the same pBluescript KSII+ subclone described above. Here, the M13-20 and M13 reverse primers were used during PCR amplifications. The amplified product was cleaved with both *Bss*HII and *Hin*dIII restriction endonucleases, and the resulting 61-bp, REBH DNA fragment was gel purified. The noncoding strand was labeled with the Klenow fragment of *E. coli* polymerase I by filling in [α-^{32}P]dCTP (3,000 Ci/mmol).

Nuclear extract preparation. Crude HeLa and HFF cell nuclear extracts were prepared by a modification described previously (19). Crude HeLa cell nuclear extracts were also purchased from Promega. Briefly, cell pellets were suspended in buffer A (10 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 7.9] containing 1.5 mM MgCl₂, 10 mM KCl, and 0.5 mM dithiothreitol) with protease inhibitors (0.5 mM phenylmethylsulfonyl fluoride, 5 μg of leupeptin per ml, 2 μg of pepstatin per ml, and 1.0 mM TLCK [*N*α-*p*-tosyl-L-lysine chloromethyl ketone]) and allowed to swell at 4°C for 20 min. Cells were lysed in a Kontes all-glass Dounce homogenizer with a type B pestle. The crude nuclei were centrifuged to form pellets and suspended in buffer C (20 mM HEPES, pH 7.9, containing 25% [vol/vol] glycerol, 0.42 M NaCl, 1.5 mM MgCl₂, and 0.2 mM EDTA) with the same protease inhibitors noted above. After 30 min at 4°C, a clear supernatant was obtained by centrifugation and dialyzed against buffer D (20 mM HEPES, pH 7.9, containing 20% [vol/vol] glycerol, 0.1 M KCl, and 0.2 mM EDTA) with protease inhibitors. Precipitates were centrifuged to form pellets, and the supernatant was stored as aliquots at −134°C.

EMSA. For detection of a cellular protein factor(s), 15-μl binding reaction mixtures included 3.5 or 7.0 μg of crude HeLa nuclear extract and 1.0 μg of sheared salmon sperm DNA preincubated for 10 min at room temperature in 25 mM Tris-HCl, pH 8.0, containing 0.5 mM EDTA, 6.25 mM MgCl₂, 0.5 mM dithiothreitol, 9% (vol/vol) glycerol, and 0.01% Nonidet P-40. For r-IE2 protein binding assay, 3.7 or 5.8 pmol of purified r-IE2 was incubated with 2 μg of sheared salmon sperm DNA at room temperature for 10 min in 15 μl of 25 mM HEPES, pH 7.9, containing 0.2 mM EDTA, 100 mM NaCl, 10 mM KCl, 8 mM MgCl₂, 1 mM dithiothreitol, and 5% glycerol. Probe DNA was added (1.5 × 10⁴ to 5.0 × 10⁴ cpm), and incubation continued an additional 15 min at room temperature. For cold competition assay, a 10- or 50-fold molar excess of a nonradioactive DNA fragment was added to the preincubation reaction mixture. DNA-protein complexes were separated from the free probe by electrophoresis through 5% polyacrylamide (36:1 acrylamid-bisacrylamide ratio) gels with 0.5×

TBE (45 mM Tris-borate [pH 8.3]–1.0 mM EDTA) at room temperature. Gels were dried and exposed to Hyperfilm MP (Amersham).

UV cross-linking of protein to ³²P-labeled DNA. The standard binding reaction was performed with 6 pmol of r-IE2 and 2 × 10⁴ to 5 × 10⁴ cpm of the BrdU-incorporated RET7H probe. Immediately after an EMSA, the wet gel was exposed to medium-wave (312 nm) UV irradiation for 15 min at a distance of 4.5 cm. After exposure to Kodak X-Omat AR film at 4°C, a gel containing a band of interest was excised. The DNA-protein complexes were eluted from the gel, denatured by boiling for 5 min in sample buffer containing 125 mM Tris (pH 6.8), 1% sodium dodecyl sulfate (SDS), and 3 mM dithiothreitol, and then subjected to electrophoresis in a 7% polyacrylamide–SDS gel. The gel was dried and exposed to film with an intensifying screen at −70°C.

Chemical footprinting. The procedure for chemical footprint analysis of specific DNA-protein complexes using 1,10-phenanthroline-copper ion has been described previously (15). Following separation by electrophoresis, the wet mobility shift gel was soaked in a 1,10-phenanthroline-copper ion solution which penetrates the gel matrix and cleaves the DNA in positions not protected by the bound protein. Autoradiography of the treated gel was used to identify and excise the specifically shifted DNA-protein complexes and the band corresponding to the free probe. The chemically cleaved DNA fragments were passively eluted from the polyacrylamide gel slices by soaking them overnight at 37°C in 400 μl of 0.5 M ammonium acetate–1 mM EDTA. The DNA fragments were precipitated with 2.5 volumes of 95% ethyl alcohol and 1 μl of glycogen (20 mg/ml). Samples were resuspended in 80% formamide–10 mM NaOH–1.0 mM EDTA and loaded on a 15 or 20% polyacrylamide–7 M urea sequencing gel. Maxam-Gilbert G/C sequencing ladders of the end-labeled DNA were run in parallel on the gels to orient the positions of the chemical footprints.

In vitro runoff transcription and analysis of RNA. Plasmid pCAT760 contains wild-type DNA sequences from −760 to +7 relative to the transcription start site upstream of the chloramphenicol acetyltransferase gene as previously described (19, 37). Mutations m1, m2, m3, m4, and m5 were substituted for wild-type sequences in plasmid pCAT760. Template DNA concentrations were estimated by UV absorbance spectroscopy (Beckman; model DU-8) followed by ethidium dot quantitation (27). The mean of three independent estimates of ethidium dot quantitations was used to determine the template concentration. Visual confirmation of the DNA quality was also confirmed by agarose gel electrophoresis and ethidium bromide staining. Transcription mixtures, which contained 150 ng (0.04 pmol) of truncated template DNA and 10 μCi of [α-³²P]UTP, were as described previously (19). Reaction mixtures were incubated at 30°C for 60 min. DNA templates were degraded with RNase-free DNase (Promega). The runoff products were extracted from the protein by the addition of 100 μl of 4.0 mM Tris, pH 7.5, containing 7.6 M urea, 0.5% SDS, 10 mM EDTA, 0.14 M NaCl, 5 μg of calf thymus tRNA, and 60 μl of phenol-chloroform-isoamyl alcohol. The radiolabeled RNAs were precipitated with 3 volumes of 95% ethyl alcohol. RNA pellets were dissolved in diethyl pyrocarbonate-treated distilled H₂O and fractionated in 6% polyacrylamide–7 M urea gels. The gels were dried and exposed to Kodak X-Omat AR film or Hyperfilm MP (Amersham). Signals were either visualized by autoradiography or quantitated by the AMBIS (San Diego, Calif.) Image Acquisition system. Transcription assays were repeated multiple times to control for variability among extracts.

RESULTS

r-IE2 protein binding to DNA bearing the MIEP. We (19) and others (12, 16, 17) have determined that the CMV IE2 protein binds to a region encompassing the transcription initiation site of the CMV MIEP. This region is required for IE2 protein-mediated repression of transcription from the MIEP in vivo (2, 18, 23) and in vitro (19) and was termed the crs. The crs is located between −13 and −1 bp, relative to the transcription start site. To determine essential base pairs for IE2 protein-DNA binding in vitro, we made 3-bp site-specific mutations upstream and downstream of the transcription start site or replaced the entire region with heterologous DNA (Fig. 1A). A ³²P-labeled 49-bp DNA probe bearing 26 bp of wild-type CMV MIEP sequences from bp −18 to +8, relative to the transcription start site, or mutant DNA probes were combined with 5.8 pmol of either purified wild-type r-IE2 or mutant r-IE2HL. Only 26 bp of wild-type DNA were used, because CMV DNA has binding sites for other cellular proteins both upstream and downstream of the 26-bp region. The mutant r-IE2HL protein failed to bind the crs-containing DNA as described previously (19). The r-IE2 protein efficiently bound wild-type DNA and mutant DNAs m3 and m4. The level of binding of r-IE2 protein to mutant m2 DNA was reduced by approximately 60%. The 3-bp mutation from −12 to −10 (m1)

A


```
                              HCMV DNA
                    ┌──────────────────────────┐
                              CRS
                    ┌─────────────────┐  ┌→
        MluI                                              ClaI
  wt  5'CGCGTTGGGAGCTCGTTTAGTGAACCGTCAGATCTCTAGAAGCTTATCG3'
  m1  5'CGCGTTGGGAGCTCaggTAGTGAACCGTCAGATCTCTAGAAGCTTATCG3'
  m2  5'CGCGTTGGGAGCTCGTTTAGTGAgaaGTCAGATCTCTAGAAGCTTATCG3'
  m3  5'CGCGTTGGGAGCTCGTTTAGTGAACCGgggGATCTCTAGAAGCTTATCG3'
  m4  5'CGCGTTGGGAGCTCGTTTAGTGAACCGTCAGgggTCTAGAAGCTTATCG3'
  m5  5'CGCGTTGGGAGCTCGgtacccggggatcctctagagtccAAGCTTATCG3'
```

B

C

FIG. 1. EMSA with r-IE2 protein. (A) DNA sequences of the wild-type and mutant *Mlu*I-*Cla*I DNA fragments from the MIEP. The wild-type DNA fragment wt bears the CMV promoter sequence from −18 to +8, which includes both the repressor element and the transcription initiation site. The crs is marked above with a line, and the transcription start site is indicated by an arrow showing the direction of transcription. The mutant DNA fragment m1, m2, m3, or m4 contains the mutations in the MIEP from −12 to −10, −3 to −1, +2 to +4, or +6 to +8, relative to the transcription start site, respectively. The mutant DNA fragment m5 has a mutation which replaces the CMV sequence from −11 to +8 and five of the flanking base pairs with heterologous DNA. All mutations in the DNA sequence are indicated by lowercase boldface letters. The *Mlu*I and *Cla*I restriction endonuclease sites are indicated. The human CMV (HCMV) wild-type DNA is indicated by a bracket. (B) Comparison of r-IE2 binding to wild-type or mutant probes by EMSA. Lanes 1, 3, 5, 7, 9, and 11 contain a wt, m1, m2, m3, m4, or m5 DNA probe, respectively, plus 5.8 pmol of r-IE2HL; lanes 2, 4, 6, 8, 10, and 12 contain a wt, m1, m2, m3, m4, or m5 DNA probe, respectively, plus 5.8 pmol of r-IE2. r-IE2-DNA complexes are indicated. (C) Cold competition assay using r-IE2. Lane 1 contains the wt DNA probe plus 3.7 pmol of r-IE2; lanes 2, 3, 4, 5, 6, and 7 contain a probe plus 3.7 pmol of r-IE2 in the presence of a 50-fold molar excess of a cold wt, m1, m2, m3, m4, or m5 DNA fragment, respectively. r-IE2-DNA complexes are indicated. The results in this panel represent a single experiment on the same gel.

and the replacement with heterologous DNA (m5) eliminated any detectable r-IE2 protein binding (Fig. 1B). The purified r-IE2 protein formed higher-order complexes on the DNA probe.

Cold competition-EMSA indicated that 50-fold molar excesses of wt, m2, m3, and m4 DNAs significantly reduced r-IE2

A

B

FIG. 2. EMSA with cellular protein(s). (A) Comparison of a cellular protein factor(s) in HeLa nuclear extract binding to wild-type or mutant probes by EMSA. Lane 1 contains the free wild-type DNA probe wt; lanes 3, 5, 7, 9, and 11 contain free mutant DNA probe m1, m2, m3, m4, or m5, respectively; lanes 2, 4, 6, 8, 10, and 12 contain a wt, m1, m2, m3, m4, or m5 DNA probe, respectively, plus 7.0 μg of HeLa cell nuclear extract. The position of the human CMV-specific complex is indicated. Also indicated is the position of a complex common to all probes. The results in this panel represent a single experiment on the same gel. (B) Cold competition assay using HeLa cell nuclear extract. Lane 1 contains the free wild-type DNA probe; lane 2 contains probe wt plus 3.5 μg of HeLa cell nuclear extract; lanes 3, 5, 7, and 9 contain probe wt plus 3.5 μg of HeLa cell nuclear extract in the presence of a 10-fold molar excess of the cold wt, m1, m2, or m3 DNA fragment, respectively; lanes 4, 6, 8, and 10 contain probe wt plus 3.5 μg of HeLa cell nuclear extract in the presence of a 50-fold molar excess of the cold wt, m1, m2, or m3 DNA fragment, respectively. Specific and common complexes are indicated by arrows. Significant separation of the specific and common complexes required a long period of electrophoresis, which placed the free probe a long distance from the bands of interest.

binding to the wild-type DNA probe, but mutant DNAs m1 and m5 had no effect (Fig. 1C). These data indicate that the 5′ end of the crs from −12 to −10 is essential for r-IE2 protein binding in vitro and that mutations to the 3′ end reduce the level of binding efficiency. Mutations to the 5′ end, center, or 3′ end of the crs were reported to effect IE2-mediated repression of downstream expression from the CMV MIEP in transiently transfected cells (2). Our in vitro IE2 protein-DNA binding results are consistent with these observations from in vivo experiments and indicate that the bases from −12 to −10 are required for IE2 protein binding.

Cellular protein binding to DNA bearing the CMV promoter. Using crude HFF (data not shown) or HeLa cell nuclear extracts for EMSA, we were able to detect specific binding of a cellular protein to the 49-bp probe bearing 26 bp of wild-type CMV promoter sequences (Fig. 2A). The probe bearing only 26 bp of wild-type DNA contains the crs and the initiator-like DNA sequence but lacks the TATA box to the 5′ side and potential wild-type DNA-protein binding sequences to the 3′ side. A common DNA-protein complex, which ap-

peared to have approximately the same intensity in each of the lanes, was detected by wild-type and mutant probes that were ^{32}P labeled to equivalent specific activities. Mutant probes m1, m2, and m4 also detected the specific DNA-protein complex but at a lower intensity. In contrast, the specific DNA-protein complex was not detectable with mutant probes m3 and m5 (Fig. 2A).

Cold competition-EMSA indicated that 10- and 50-fold molar excesses of wild-type DNA significantly reduced cellular protein binding and that 50-fold molar excesses of mutants m1, m2, and m4 (data not shown) had a similar effect. In contrast, mutant DNAs m3 and m5 (data not shown) had no effect (Fig. 2B). The 3-bp mutation from +2 to +4 (m3) and the replacement with heterologous DNA (m5) eliminated any detectable cellular protein binding.

Determining the molecular masses of the viral and cellular DNA-binding proteins. To determine the sizes of the viral and cellular binding proteins, we covalently cross-linked nonradioactive proteins to radioactive DNA. PCR was used to incorporate BrdU into a 59-bp DNA probe bearing 26 bp of wild-type DNA (RET7H). Probe DNA was labeled with ^{32}P and used in gel mobility shift assays, and specific complexes were subsequently cross-linked by ultraviolet (UV) irradiation as described in Materials and Methods. r-IE2 and a truncation mutant of r-IE2 protein designated r-IE2552 were produced in *E. coli* and purified. These two forms of IE2 protein are able to bind the CMV sequence of the wild-type probe in EMSA, and these viral proteins are capable of repressing the CMV MIEP in transient transfection assays (19a). Coomassie blue staining of an SDS-polyacrylamide gel showed that the recombinant proteins were greater than 90% pure. Measured against prestained molecular mass standards (Fig. 3A, lane 1), r-IE2 had a molecular mass of 70 kDa (Fig. 3A, lane 2) and r-IE2552 had a molecular mass of 67 kDa (Fig. 3A, lane 3). Binding reactions were performed with wild-type ^{32}P-labeled BrdU-RET7H DNA, and each of the purified recombinant viral proteins. Following electrophoresis, the wet gel was exposed to medium-wave (312 nm) UV irradiation to covalently cross-link the protein to DNA. The approximate molecular masses of the cross-linked proteins were determined by the method of Williams et al. (44) as described in Materials and Methods. The sizes of the complexes formed by the r-IE2 and r-IE2552 proteins corresponded to those predicted for the dimer forms of IE2 protein, approximately 140 kDa (Fig. 3B, lane 1) and 134 kDa (Fig. 3B, lane 2), respectively. Although Fig. 1B indicates that oligomers of IE2 protein can form on the DNA template, a dimer of the IE2 protein is in close contact with the viral DNA. This result was consistent with our expectations, since both recombinant IE2 proteins retained the domains required for dimerization and DNA binding (3, 5).

To study the cellular binding protein(s), binding reactions were performed with HeLa cell nuclear extract and wild-type BrdU-RET7H DNA. We excised both the specific complex and the common complex from the wet mobility shift gel for size analysis by SDS-polyacrylamide gel electrophoresis (PAGE). The common complex served as a control for cross-linking the heterogeneous protein extract to DNA. It would be expected that the two bands isolated by gel mobility shift assay were generated by DNA-binding proteins with characteristics distinguishable by SDS-PAGE. The specific complex yielded a discrete protein band which migrated with an apparent molecular mass of approximately 150 kDa (Fig. 3C, lane 1). UV irradiation dose-response curves did not detect smaller-size proteins binding to the RET7H probe DNA. The cross-linked protein of the common complex migrated at approximately 100 kDa (Fig. 3C, lane 2).
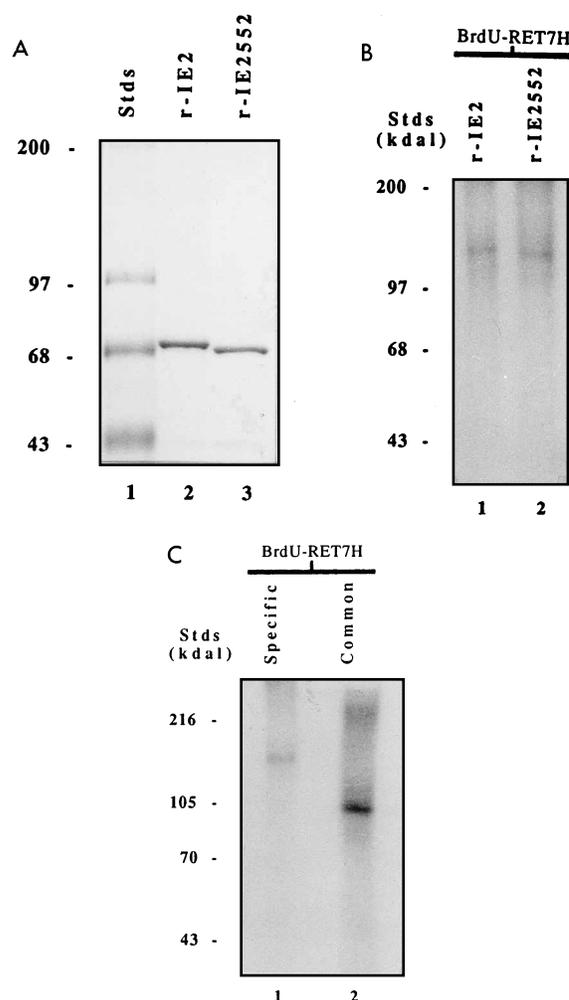
FIG. 3. Size determination of binding proteins cross-linked by ultraviolet (UV) irradiation to a BrdU-incorporated RET7H probe bearing 26 bp of wild-type CMV DNA. (A) Coomassie blue-stained SDS gel of viral r-IE2 proteins purified from *E. coli*. The carboxyl-terminal 290 amino acids of IE2 were cloned and expressed as a fusion protein designated r-IE2. r-IE2552 is similar to r-IE2 but has a 28-amino-acid deletion at the carboxyl terminus. Lanes: 1, molecular mass standards (Stds) in kilodaltons; 2, 2.0 μg of purified r-IE2 fusion protein; 3, 2.0 μg of purified r-IE2552 fusion protein. (B) SDS-PAGE of UV cross-linked r-IE2 and r-IE2552 protein to wild-type BrdU-RET7H probe. Lanes: 1, r-IE2 protein UV cross-linked to wild-type BrdU-RET7H DNA; 2, r-IE2552 protein UV cross-linked to wild-type BrdU-RET7H DNA. Migrations of prestained molecular mass standards are indicated on the left. (C) SDS-PAGE of proteins from HeLa nuclear extracts UV cross-linked to the wild-type BrdU-RET7H probe. The standard EMSA was performed as described in the legend to Fig. 2 but with approximately 5 × 10[4] cpm of BrdU-RET7H probe. The positions of the prestained protein molecular mass standards are indicated on the left. Lanes: 1, UV cross-linked protein derived from the CMV-specific complex; 2, cross-linked protein derived from the common complex.

Chemical cleavage DNA footprint analysis of the viral r-IE2 and cellular protein binding sites. Chemical cleavage footprinting was used to more precisely define the regions of the cloned CMV DNA that were in close contact with the viral or cellular binding proteins and to compare and contrast the binding of the two proteins. Gel mobility shift assays were performed with either purified r-IE2 protein or crude HeLa cell nuclear extract and the wild-type CMV DNA specifically end labeled on the 59-bp coding (RET7H) or the 61-bp noncoding (REBH) strand bearing 26 bp of wild-type CMV MIEP sequences from −18 to +8, relative to the transcription start
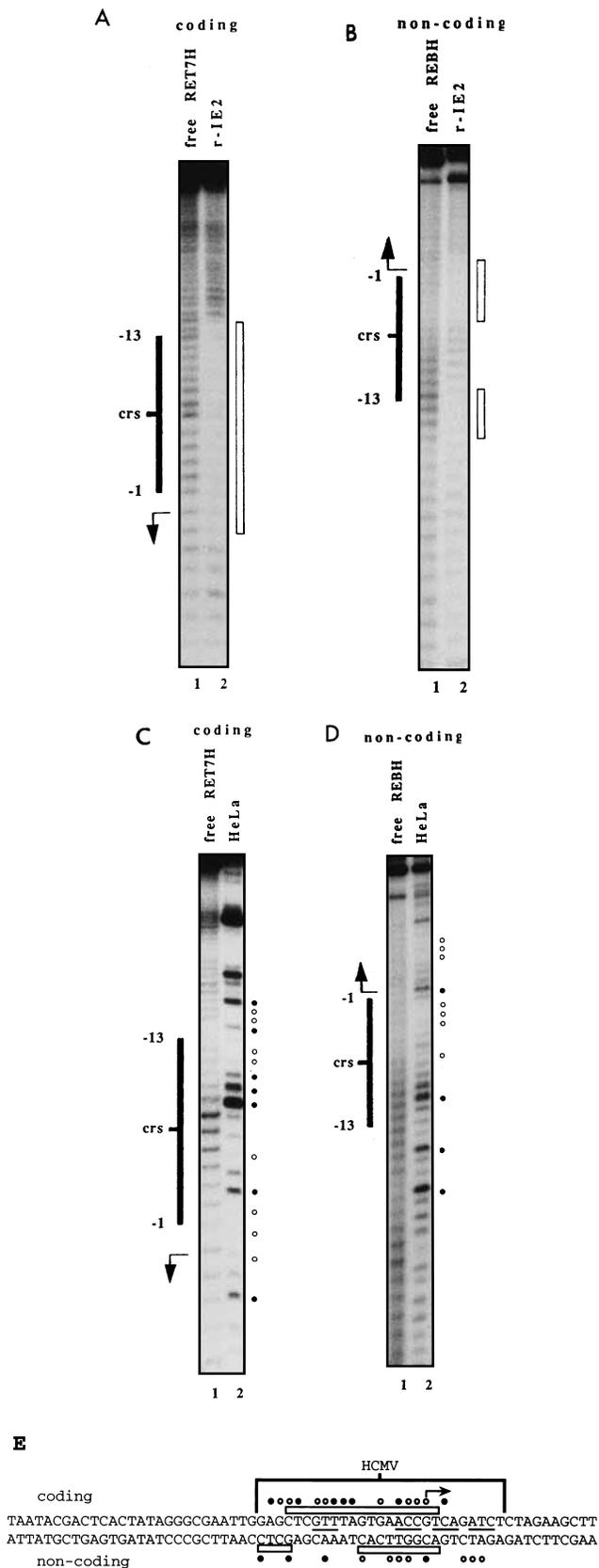
FIG. 4. Chemical footprint analyses to define sequence-specific protein-DNA interactions. Standard binding reaction mixtures for the r-IE2 protein and HeLa nuclear extract were scaled up 4- to 10-fold. Wet mobility shift gels were

site, as described in Materials and Methods. On the coding strand, r-IE2 protein protects a region from −15 to +2 (Fig. 4A). r-IE2 protein protects two discrete regions on the noncoding strand, from −16 to −13 and from −7 to +2 (Fig. 4B). These results differ slightly from those of Lang and Stamminger (17), who detected protection from −16 to −1 with purified histidine-tagged IE2 protein.

The bound cellular protein(s) protects the CMV DNA from cleavage at positions −16, −15, −12, −11, −5, −2, −1, and +1 on the coding strand (Fig. 4C) and at positions −7, −4, −3, −2, +5, +6, and +7 on the noncoding strand (Fig. 4D). Cellular binding protein(s) also induces hypersensitive sites on the flanking vector and CMV DNA sequences at positions −17, −14, −10, −9, −8, −3, and +3 on the coding strand (Fig. 4C) and at positions −18, −15, −11, and +1 on the noncoding strand (Fig. 4D). DNA fragments at the top of the gel did not fractionate well, and consequently, the 5′ and 3′ limits of the cellular protein binding site can be seen only at the bottom of the gel (Fig. 4C and D). The data derived from the chemical footprint analysis and the locations of mutations m1, m2, m3, and m4 are summarized in Fig. 4E. The r-IE2 protein footprint and the human cellular protein footprint are adjacent and overlapping. The footprint patterns of the viral and the cellular proteins binding to the DNA containing 26 bp of CMV MIEP DNA are very different and suggest different interactions with the DNA molecule.

In vitro transcription. We have shown previously that IE2 protein bound to the CMV DNA template represses in vitro transcription and that the presence of the crs is required for this effect (19). To further investigate the effects of the promoter mutations, we compared the efficiencies of the wild-type and mutated CMV promoters in in vitro runoff transcription assays. Wild-type and mutant template DNAs were quantitated as described in Materials and Methods. Approximately 150 ng of DNA was used for each in vitro transcription runoff assay. Wild-type and mutant promoter templates bear the wild-type CMV TATA motif, as well as the upstream enhancer sequences, and consequently, transcription initiated to generate runoff transcripts of approximately 268 nucleotides. Figure

5A shows a representative in vitro runoff transcription assay using wt and mutant promoter m1, m2, m3, or m4. The m3 mutant promoter was threefold weaker than the wild-type promoter and produced RNA transcripts with 34.5% of the efficiency of the wild-type promoter. Promoter mutations adjacent to m3 produced RNA transcripts with 51.3% (m2) and 75.4% (m4) of the efficiency of the wild-type promoter. The mutation in the m1 promoter had no detectable effect. To determine the effect of the entire region, mutation m5, in which the sequences both upstream and downstream of the transcription start site were replaced with heterologous DNA, was also analyzed by in vitro transcription. Figure 5B shows a representative in vitro runoff transcription assay. The efficiency of the m5 mutant promoter was approximately 45% of that of the wild-type CMV promoter. Because the CMV MIEP strongly competes with other promoters, another promoter as an internal standard was not feasible, and consequently, in vitro transcription from the m5 mutant promoter was repeated numerous times with different templates and nuclear extract preparations. The mean ($n = 11$) transcriptional efficiency level of the m5 promoter (43%) was consistently below that of the wild-type promoter (Fig. 5C). Our results indicate that a cellular protein(s), detected in both HeLa and HFF cell nuclear extracts, binds specifically to the wild-type CMV sequences and that the ability of the protein to bind to the CMV promoter correlates with transcription efficiency.

DISCUSSION

The CMV IE2 protein and human cellular protein(s) interact with the CMV MIEP at the transcription start site. Mutational analysis indicates that the core DNA sequences for IE2 protein and cellular protein binding sites are adjacent. The m1 and m2 mutations immediately upstream of the transcription start site eliminated or reduced IE2 protein binding, respectively. The m3 mutation immediately downstream of the transcription start site eliminated cellular protein binding. Chemical footprint analysis indicated that the binding sites overlap. These findings imply roles for the promoter binding proteins in determining the net activity of the CMV MIEP in the virus-infected cells. We (18) and others (2, 23) have proposed that the CMV IE2 protein is involved in down-regulation of the strong CMV MIEP during viral infection. In contrast, the cellular protein may have a positive effect on transcription efficiency. The 3-bp substitutions immediately downstream of the transcription start site of the CMV MIEP (m3), which eliminated binding by the cellular protein, also significantly reduced the level of transcription efficiency from the enhancer-containing promoter. These results were demonstrated by in vitro transcription assays. Transient transfection assays in HeLa or HFF cells with wild-type or mutant promoters upstream of the chloramphenicol acetyltransferase gene confirmed the in vitro transcription assay results (data not shown).

The position of the cellular DNA binding site is similar to those of initiator elements in that it overlaps the transcription initiation site. The site also contains the consensus sequence found in other initiator elements, Py Py A N T/A Py Py (11). A component of TFII-D specifically interacts with an initiator (13, 21, 41). In *Drosophila* cells a TAFII with a molecular mass of 150 kDa (dTAFII150) binds to the DNA sequence containing an initiator element and functions in the assembly of an initiation complex (41, 42). However, the dTAFII150 footprint on the major late promoter of adenovirus was different from the footprint we observed on the CMV MIEP. dTAFII150 gives a footprint between −1 and +38 relative to the transcription start site (41, 42), whereas the unknown human cellular
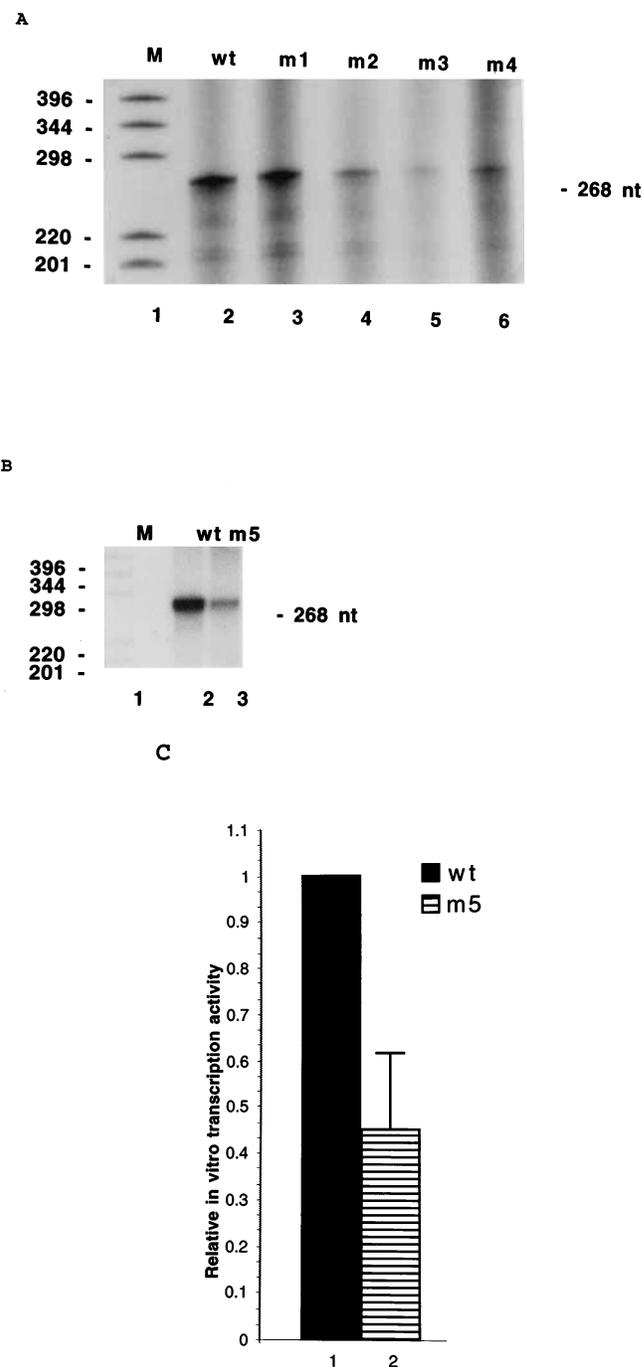

FIG. 5. Effects of the CMV MIEP mutations. (A) Representative in vitro transcription efficiencies from the wild-type and mutant CMV MIEP. Sixty-four micrograms of crude HeLa nuclear extracts was used to transcribe 150 ng (0.04 pmol) of the truncated templates, the wild-type template wt, and mutant templates m1, m2, m3, and m4 in the presence of 10 μCi of [^{32}P]UTP. Lanes: 1, molecular mass markers (M); 2, wt; 3, mutant m1; 4, mutant m2; 5, mutant m3; 6, mutant m4. Transcription products were analyzed as described in Materials and Methods. (A and B) The position of the 268-nucleotide (nt) runoff transcript from the wild type and mutant constructs is indicated on the right. (B) Representative in vitro transcription efficiencies from the wt (lane 2) and mutant m5 (lane 3) promoters. Lane 1 contains molecular mass markers (M). (C) Quantitation of the runoff transcripts generated by in vitro transcription of the mutant m5 promoter (by the AMBIS Image Acquisition system) yielded the data presented in the bar graph ($n = 11$). Transcription from the wild-type CMV MIEP was set as equal to 1, and mutant m5 promoter activity is presented relative to that of the wild type.

protein gives a footprint between −16 and +7 on the CMV MIEP. Both cellular DNA-binding proteins induced footprints demonstrating protected bases and hypersensitive cleavage sites. Whether dTAFII150 is related to the human cellular protein described in this report is not known. Antibody to dTAFII150 apparently does not cross-react with the human homolog (39a).

It had been reported that the positive effect of an initiator element was sensitive to the efficiency of the promoter's TATA box and the presence of upstream activating sequences (11). In this case, the CMV MIEP has a very strong upstream enhancer (39). The two- and threefold reductions at the transcriptional level with the mutant m5 and m3 promoters, respectively, are quantitatively significant given the potency of the complex enhancer system. The m4 mutation, which is at the 3′ end of the putative initiator sequence, had less of an effect on promoter efficiency, but the downstream pyrimidine residues may have facilitated cellular protein binding.

By UV cross-linking assays, the molecular mass of the cellular DNA-binding protein of our study appears to be approximately 150 kDa. This procedure does not exclude the possibility that the 150-kDa size is a result of the UV cross-linking of not one protein but of a stable protein complex with a total molecular mass of 150 kDa.

The chemical footprint patterns differ notably between the r-IE2 protein and the cellular protein. The r-IE2 protein footprints on both the coding and noncoding strands of the DNA targets exhibit broad regions of protection. In contrast, the cellular protein's binding appears to distort the DNA structure to produce a footprint pattern that reveals positions of protection as well as positions of greatly enhanced chemical cleavage. The reasons for these different footprints remain to be determined, but they may be related to how these two different proteins interact with the CMV MIEP. While the IE2 protein requires minor groove contacts (17), the cellular protein may require major groove contacts.

Of particular interest regarding the location of the binding site for the cellular factor was that it overlapped sequences known to be required for down-regulation of the strong CMV MIEP by the viral IE2 protein. IE2-mediated repression could not be explained as physical interference with the TATA-binding protein (TBP), since purified IE2 and purified TBP could be shown by gel shift assays to bind simultaneously to TATA-bearing CMV probes (12). To determine the relative affinities of the viral and cellular proteins for binding to the transcription start site requires purification of the cellular factor(s). The viral IE2 protein may have a higher binding affinity level than the unknown cellular protein. The cellular protein may be an abundant protein with a low-level binding affinity. It is possible that transcription levels from the CMV MIEP are regulated by the interplay of the cellular factor(s) and the viral IE2 protein. The IE2 protein could prevent the assembly of the transcription initiation complex and prevent the binding of the human cellular factor(s) to the transcription start site. Because the binding sites for IE2 and the cellular protein(s) overlap, it is possible that part of the mechanism of IE2-mediated repression is a result of interference with the normally positive activity of a cellular protein. This mechanism would be consistent with the recent report that showed IE2 represses transcription from the CMV MIEP during assembly of the transcription preinitiation complex and not during the elongation process (45). It is possible, but less likely, that the cellular factor increases the level of transcription efficiency by acting on elongation. An understanding of how cellular protein binding to the initiation site increases transcription efficiency levels and viral protein represses transcription will contribute towards our knowledge of preinitiation complex formation and transcription regulation. These events are linked to CMV gene expression and may have an influence on viral latency or reactivation in certain cell types.

REFERENCES

1. **Carcamo, J., L. Buckbinder, and D. Reinberg.** 1991. The initiator directs the assembly of a transcript factor IID-dependent transcription complex. Proc. Natl. Acad. Sci. USA **88:**8052–8056.
2. **Cherrington, J. M., E. L. Khoury, and E. S. Mocarski.** 1991. Human cytomegalovirus *ie*2 negatively regulates α gene expression via a short target sequence near the transcription start site. J. Virol. **65:**887–896.
3. **Chiou, C.-J., J. Zong, I. Waheed, and G. S. Hayward.** 1993. Identification and mapping of dimerization and DNA-binding domains in the C terminus of the IE2 regulatory protein of human cytomegalovirus. J. Virol. **67:**6201–6214.
4. **Du, H., A. L. Roy, and R. G. Roeder.** 1993. Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. EMBO J. **12:**501–511.
5. **Furnari, B. A., E. Poma, T. F. Kowalik, S.-M. Huong, and E.-S. Huang.** 1993. Human cytomegalovirus immediate-early gene 2 protein interacts with itself and with several novel cellular proteins. J. Virol. **67:**4981–4991.
6. **Ghazal, P., H. Lubon, B. Fleckenstein, and L. Henninghausen.** 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.
7. **Hansen, S. K., and R. Tjian.** 1995. TAFs and TFIIA mediate differential utilization of the tandem Adh promoters. Cell **82:**565–575.
8. **Hermiston, T. W., C. L. Malone, and M. F. Stinski.** 1990. Human cytomegalovirus immediate-early two protein region involved in negative regulation of the major immediate-early promoter. J. Virol. **64:**3532–3536.
9. **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.
10. **Ho, M.** 1990. Epidemiology of cytomegalovirus infections. Rev. Infect. Dis. **12:**5701–5710.
11. **Javahery, R., A. Khachi, K. Lo, B. Zenzie-Gregory, and S. T. Smale.** 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol. Cell. Biol. **14:**116–127.
12. **Jupp, R., S. Hoffmann, A. Depto, R. M. Stenberg, P. Ghazal, and J. A. Nelson.** 1993. Direct interaction of the human cytomegalovirus IE86 protein with the *cis* repression signal does not preclude TBP from binding to the TATA box. J. Virol. **67:**5595–5604.
13. **Kaufmann, J., and S. T. Smale.** 1994. Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev. **8:**821–829.
14. **Kondo, K., H. Kaneshima, and E. S. Mocarski.** 1994. Human cytomegalovirus latent infections of granulocyte-macrophage progenitors. Proc. Natl. Acad. Sci. USA **91:**11879–11883.
15. **Kuwabara, M. D., and D. S. Sigman.** 1987. Footprinting DNA-protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: *Escherichia coli* RNA polymerase-lac promoter complexes. Biochemistry **26:**7234–7238.
16. **Lang, D., and T. Stamminger.** 1993. The 86-kilodalton IE-2 protein of human cytomegalovirus is a sequence-specific DNA-binding protein that interacts directly with the negative autoregulatory response element located near the cap site of the IE-1/2 enhancer-promoter. J. Virol. **67:**323–331.
17. **Lang, D., and T. Stamminger.** 1994. Minor groove contacts are essential for an interaction of human cytomegalovirus IE2 protein with its DNA target. Nucleic Acids Res. **22:**3331–3338.
18. **Liu, B., T. W. Hermiston, and M. F. Stinski.** 1991. A *cis*-acting element in the major immediate-early (IE) promoter of human cytomegalovirus is required for negative regulation by IE2. J. Virol. **65:**897–903.
19. **Macias, M. P., and M. F. Stinski.** 1993. An *in vitro* system for human cytomegalovirus immediate early 2 protein (IE2)-mediated site-dependent repression of transcription and direct binding of IE2 to the major immediate early promoter. Proc. Natl. Acad. Sci. USA **90:**707–711.
19a. **Macias, M. P., and M. F. Stinski.** Unpublished data.
20. **Malone, C. L., D. H. Vesole, and M. F. Stinski.** 1990. Transactivation of a human cytomegalovirus early promoter by gene products from the immediate-early gene IE2 and augmentation by IE1: mutational analysis of the viral proteins. J. Virol. **64:**1498–1506.
21. **Martinez, E., C.-M. Chiang, H. Ga, and R. G. Roeder.** 1994. TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct

basal transcription from a TATA-less class II promoter. EMBO J. **13:**3115–3126.

22. **Means, A. L., J. E. Slansky, S. L. McMahon, M. W. Knuth, and P. J. Farnham.** 1992. The HIP-1 binding site is required for growth regulation of the dihydrofolate reductase gene promoter. Mol. Cell. Biol. **12:**1054–1063.

23. **Pizzorno, M. C., and G. S. Hayward.** 1990. The IE2 gene products of human cytomegalovirus specifically down-regulate expression from the major immediate-early promoter through a target sequence located near the cap site. J. Virol. **64:**6154–6165.

24. **Pizzorno, M. C., P. O'Hara, 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.

25. **Rawlinson, W. D., and B. G. Barrell.** 1993. Spliced transcripts of human cytomegalovirus. J. Virol. **67:**5502–5513.

26. **Roy, A. L., M. Meisterernst, P. Pognonec, and R. G. Roeder.** 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature (London) **354:**245–248.

27. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

28. **Santomenna, L. D., and P. A. Colberg.** 1990. Induction of cellular hsp70 expression by human cytomegalovirus. J. Virol. **64:**2033–2040.

29. **Seto, E., Y. Shi, and T. Shenk.** 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription *in vitro*. Nature (London) **354:**241–245.

30. **Smale, S. T., and D. Baltimore.** 1989. The initiator as a transcription control element. Cell **57:**103–113.

31. **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.

32. **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.

33. **Stenberg, R. M., J. Fortney, S. W. Barlow, B. P. Magrane, J. A. Nelson, and P. Ghazal.** 1990. Promoter-specific transactivation and repression of human cytomegalovirus immediate-early proteins involve common and unique protein domains. J. Virol. **64:**1556–1565.

34. **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.

35. **Stinski, M. F.** 1990. Cytomegalovirus and its replication, p. 1959–1980. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed. Raven Press, New York.

36. **Stinski, M. F., M. P. Macias, C. L. Malone, A. R. Thrower, and L. Huang.** 1993. Regulation of transcription from the major immediate early promoter by cellular and viral proteins, p. 3–12. *In* S. Michelson and S. A. Plotkin (ed.), Multidisciplinary approach to understanding cytomegalovirus disease. Elsevier Science Publishing Company, Inc., Amsterdam.

37. **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.

38. **Taylor-Wiedeman, J., J. G. P. Sissons, L. K. Borysiewicz, and J. H. Sinclair.** 1991. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. J. Gen. Virol. **72:**2059–2064.

39. **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.

39a.**Tjian, R.** Personal communication.

40. **Usheva, A., and T. Shenk.** 1994. TATA-binding protein independent initiation: YY1, TFIIB and RNA polymerase II direct basal transcription on supercoiled template DNA. Cell **76:**1115–1121.

41. **Verrijzer, C. P., J.-L. Chen, K. Yokomori, and R. Tjian.** 1995. Binding of TAFs to core elements directs promoter selectivity by RNA polymerase II. Cell **81:**1115–1125.

42. **Verrijzer, C. P., K. Yokomori, J.-L. Chen, and R. Tjian.** 1994. *Drosophila* TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. Science **264:**933–941.

43. **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.

44. **Williams, M., A. Brys, A. M. Weiner, and N. Maizels.** 1992. A rapid method for determining the molecular weight of a protein bound to nucleic acid in a mobility shift assay. Nucleic Acids Res. **20:**4935–4936.

45. **Wu, J., R. Jupp, R. M. Stenberg, J. A. Nelson, and P. Ghazal.** 1993. Site-specific inhibition of RNA polymerase II preinitiation complex assembly by human cytomegalovirus IE86 protein. J. Virol. **67:**7547–7555.