JOURNAL OF VIROLOGY, Dec. 1994, p. 7953–7965
0022-538X/94/$04.00+0

Vol. 68, No. 12

# Requirements for Activation of the Herpes Simplex Virus Glycoprotein C Promoter In Vitro by the Viral Regulatory Protein ICP4

BAOHUA GU AND NEAL DELUCA*

*Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261*

Received 14 July 1994/Accepted 30 August 1994

**During infection with herpes simplex virus, infected-cell polypeptide 4 (ICP4) activates transcription of most herpes simplex virus genes. In the present study, the mechanism of activation of transcription by ICP4 was investigated by using a reconstituted in vitro system with fractionated and purified general transcription factors, coupled with DNA-binding assays. The templates used in the reactions included regions of the gC and thymidine kinase (tk) promoters in plasmids, and on isolated fragments, allowing for the evaluation of the potential function of naturally occurring and inserted ICP4-binding sites and elements of the core promoter. ICP4 efficiently activated transcription of the gC promoter by facilitating the formation of transcription initiation complexes. ICP4 could not substitute for any of the basal transcription factors. Moreover, TATA-binding protein (TBP) could not substitute for TFIID in activation, suggesting a requirement for TBP-associated factors. Interactions between ICP4 and DNA 3′ to the start site was necessary for activation of the gC promoter. The requirement for DNA-protein contacts could be met either by the presence of an ICP4-binding site in the gC leader, by the presence of a site more than 150 nucleotides further downstream, by an inserted site that normally acts to repress transcription, or by the addition of sufficient non-site-containing DNA. The gC TATA box and start site, or initiator element (inr), were individually sufficient for activation by ICP4 and together contributed to optimal activation. In contrast to gC, the tk promoter was poorly activated in the reconstituted system. However, the tk TATA box was efficiently activated when the tk start site region was replaced with the gC inr, suggesting that activation was mediated through the inr and inr-binding proteins. In addition, mutation of the inr core resulted in a gC promoter that was very poorly activated by ICP4. The results of this and previous studies demonstrate that ICP4 activates transcription in a complex manner involving contacts with DNA 3′ to the start site, TBP, TFIIB, TBP-associated factors, and possibly proteins functioning at the start site of transcription.**

Herpes simplex virus (HSV) contains a 152-kb double-stranded DNA genome encoding approximately 75 genes (45, 46). Early in infection, five of these genes, termed immediate-early (IE) or α genes (32), are efficiently transcribed (82) partly as a result of the transcriptional activation function of VP16 (7), which is brought in with the infecting virion (2, 59). The products of the IE genes function at different levels and at different times in the viral life cycle to achieve the efficient and regulated expression of the remainder of the viral genome (9, 33, 81). These are infected-cell polypeptide 4 (ICP4), ICP0, ICP27, ICP22, and ICP47 (9, 32, 57, 82). Of these, ICP4 is essential for the growth of herpes simplex virus because it activates transcription of most of the essential genes of the virus (13, 14, 19, 21, 24, 52, 60, 62, 81). In the absence of ICP4, the accumulation of viral early and late mRNA ranges from undetectable to several percent of that seen in the presence of ICP4 (16, 35, 36). ICP4 can also function as a transcriptional repressor, most notably acting to repress transcription of its own promoter (14, 26, 49, 53, 64, 65).

ICP4 is a 175-kDa phosphoprotein (11, 57) that localizes mostly in the nucleus (11) and exists in solution as a homodimer (48, 72). Genetic and biochemical studies of ICP4 have demonstrated that ICP4 is composed of discrete functional domains including DNA-binding, dimerization, nuclear localization, and transcription activation regions (16, 54, 55, 71), which collectively function to produce the transcriptional profile seen in infected cells. ICP4 has a specific DNA-binding activity, recognizing a very degenerate consensus sequence (17, 23, 41, 50, 51). The specific DNA-binding activity is important for repression of transcription in vivo and in vitro (26, 49, 65) and may also be involved in the temporal regulation of some viral genes, repressing transcription prior to viral DNA synthesis (40, 64). With few exceptions (37, 56, 70), most mutations in ICP4 that reduce its ability to specifically bind to DNA also reduce its ability to repress and activate transcription (55, 71). However, several previous studies have been unable to demonstrate that any specific sequence or ICP4-binding site(s) is necessary for activation in HSV-infected or transfected cells (10, 20, 22, 28).

ICP4 and the pseudorabies virus 180-kDa IE protein are functional analogs that possess regions of amino acid conservation (8), and both have been shown to activate transcription in vitro (1, 3). Pizer and colleagues (78) have proposed that ICP4-binding sites in the vicinity of the gD promoter facilitate activation by ICP4; however, the same binding sites are completely dispensable in the context of the viral genome (75). This led to the hypothesis that ICP4 can activate transcription through any one of the numerous ICP4-binding sites present in the HSV genome, despite the relatively low affinity of ICP4 for

* Corresponding author. Mailing address: E1257 Biomedical Science Tower, Department of Molecular Genetics and Biochemistry, University of Pittsburgh, Pittsburgh, PA 15261. Phone: (412) 648-9947. Fax: (412) 624-1401.

most of these sites. The results presented in this study test the above hypothesis as well as the absolute need for specific ICP4-binding sites in transcriptional activation.

Each of the HSV genes has its own promoter that is recognized by cellular RNA polymerase II, and almost all the promoters contain a TATA box, which is recognized by TATA-binding protein (TBP) (30, 39, 58). For the specific initiation of transcription, RNA polymerase II requires a number of other proteins termed general, or basal, transcription factors (44, 67). These include TFIID, TFIIB, TFIIE/F, TFIIA, TFIIH, and TFFIIJ (5). TFIID exists in cells as a complex containing TBP and TBP-associated factors (TAFs) (85), which have been shown to be important for the function of many activator proteins (4, 79). The sequential interaction and assembly of the general transcription factors to form an initiation complex on the promoter is believed to start with the binding of TBP to the TATA box (6). It has previously been shown that ICP4 can form a tripartite complex with TBP and TFIIB on DNA and that this interaction is mediated through a region of ICP4 that is important for activation (76). In addition to the TATA box, many core promoters contain an important element located around the start site of transcription, termed an initiator (inr) element (74, 84). This element can function in the absence of the TATA box to promote the site-specific initiation of transcription. The best-studied initiator is that for TATA-less terminal deoxynucleotidyl transferase (TdT) promoter, having the core sequence −3 CTCA +1 (74). The experiments in this study address the general requirements of TAFs for activation by ICP4 and the importance of the start site region, or inr element.

To investigate the mechanism of ICP4 activation and address the questions raised above, we reconstituted in vitro transcription with fractionated general transcription factors from HeLa cells and purified TFIIB, TBP, and ICP4 and templates containing the gC and thymidine kinase (tk) promoters. The promoters for gC (31, 83) and tk (10, 38, 47) have been extensively studied in intact cells, and while gC and tk differ with respect to the presence of functional *cis* sites for cellular factors upstream of the TATA box, both can be activated by ICP4 as a sole consequence of the TATA box (31, 35, 36). This study investigates the mechanism of and parameters for ICP4 to activate these core promoters in vitro.

## MATERIALS AND METHODS

**Construction of template plasmids.** Plasmid pgCL contains the gC promoter sequences from −35 to +124 relative to the transcription start site. The construction of pgCL is described below in greater detail, and other promoter-containing plasmids used in this study were made by the same strategy. The *Sma*I (−35)–*Nru*I (+124) fragment of the gC promoter region was cloned into the *Bam*HI-*Bgl*II sites of p−111/−32HB after the ends were made blunt with the Klenow fragment of *Escherichia coli* polymerase and converted with synthetic linkers to *Bam*HI and *Bgl*II sites, respectively. p−111/−32HB is a pUC18 derivative consisting of a *Hin*dIII (−500)–*Bgl*II (+55) fragment of HSV-1 tk promoter region that also has a *Bam*HI restriction site to replace sequences between −111 and −32 (36). Thus, the *Bam*HI-*Bgl*II-digested vector portion of p−111/−32HB lacks tk promoter sequences and was used to clone in the gC promoter and its derivatives. The *Hin*dIII-*Bgl*II fragment contained in the tk-gC hybrid HB plasmid was then used to replace the *Hin*dIII and *Bgl*II fragment in pLSWT. pLSWT contains the HSV sequence from −500 relative to the start of tk transcription to the *Bam*HI site 2.86 kb downstream of the tk start site (36). The resulting plasmid contains the gC promoter region situated within the tk gene, replacing the promoter region of tk from −111 to +55. We routinely construct promoter-tk gene chimeras in this way to test the constructs in the context of the viral genome. In this way, the same test genes can be readily analyzed both in the viral genome and in in vitro transcription. This study focuses mostly on in vitro studies.

The following gC derivative plasmids were made by using synthesized oligonucleotides. 5′GATCCGGGGTATAAATT CCGGAAGGGGACACGGGCTACCCTCACTA3′ and 5′G ATCTAGTGAGGGTAGCCCGTGTCCCCTTCCGGAATT TATACCCCG3′ were annealed to generate a *Bam*HI-*Bgl*II 46-mer double-stranded DNA consisting of the gC promoter from −35 to +5. The 46-mer was cloned into p−111/−32HB and then into pLSWT to give rise to p46gC. p46gCm was made with the same oligonucleotides except that there is a single C-to-G change at the indicated position. In the same way, a similar construct containing the tk TATA box and the gC start site region ptk-gC was constructed with the oligonucleotides 5′GATCCGGTTCGCATATTAAGGTGAAGGGGACACG GGCTACCCTCACTA3′ and 5′GATCTAGTGAGGGTAG CCCGTGTCCCCTTCACCTTAATATGCGAACCG3′. To make a template containing the gC TATA box and the tk start site region, the oligonucleotides 5′GATCCGGGGTATAAA TTCCGGA3′ and 5′CGCGTCCGGAATTTATACCCCG3′ were annealed and cloned into the *Bam*HI (−111) and *Mlu*I (−11) sites of p−111/−32HB and then into pLSWT as described above.

Several plasmids with 3′ deletions of the gC leader were generated, from which the templates in Fig. 5 were derived. DNA of pgCL was linearized at *Bgl*II and then treated briefly with *Bal* 31 exonuclease. *Bgl*II linkers were then added to the ends created by *Bal* 31, and the DNA was digested with *Hin*dIII and *Bgl*II. The resulting promoter-containing fragments were isolated and cloned into the vector portion of *Hin*dIII- and *Bgl*II-digested pLSWT. Individual deletion mutant isolates were sequenced to determine the deletion end point in the leader sequence.

To make gC TATA box deletion plasmids, the HB derivatives of p46gC and p46gCm DNA were digested with *Bam*HI (−35) and *Bsp*EI (−20) and treated with the Klenow fragment of *E. coli* DNA polymerase to fill in the overhangs. After blunt-end ligation the *Bam*HI restriction site was regenerated, and the resulting plasmid had a net deletion of 13 bp including the gC TATA box (CGGGGTATAAATT). The tk promoter upstream sequences from −500 (*Hin*dIII) to −32 (*Bam*HI) were used to replace the *Hin*dIII-*Bam*HI fragment from the above plasmid to position the Sp1- and CAAT-binding sites normally upstream of the tk promoter, upstream of the gC start site region. The constructs were then put into pLSWT as described above.

To clone the high-affinity ICP4 binding-site sequence from the ICP4 promoter into gC-containing plasmids, the oligonucleotides 5′GATCCGCCCCGATCGTCCACACGGAGCGC G3′ and 5′GATCCGCGCTCCGTGTGGACGATCGGGGC G3′ (the ICP4-binding consensus sequence is underlined) were annealed and ligated into the unique *Bgl*II site downstream of the transcription start in the pLSWT derivatives (at +48 in templates 9 and 10 of Fig. 5). Mutant 27-mer oligonucleotides, lacking ATC in the above sequences, were also used to generate control plasmids that gave rise to templates 8 and 12 in Fig. 5. DNA fragments used as templates in the transcription studies were generated with restriction enzymes and purified from 4% native polyacrylamide gels by electroelution and Sephadex G-50 chromatography.

**Cells and viruses.** HeLa cells were grown in suspension in minimal essential medium (Gibco) supplemented with 10% fetal bovine serum and used to make the general transcription factor fractions. The procedures for growth and maintenance of HSV in Vero or E5 cells have been described previously (13). HSV KOS was used as the initial source of ICP4. E5 cells are a transformed Vero cell line that express complementing levels of ICP4 (15). n12 is a mutant HSV strain that contains a nonsense mutation in ICP4 that specifies no ICP4 activity and can be propagated on E5 cells (16). The n12 derivative virus harboring the gCL-tk chimera at the tk locus on which the primer extension analysis was carried out (see Fig. 1C) was provided by Ramon Rivera-Gonzalez and will be described elsewhere.

**Preparation of general transcription factors, TBP, TFIIB, and ICP4.** HeLa cell suspension (80 liters) was used to make nuclear extracts by the method of Dignam et al. (18). Fractionation of the general transcription factors from the nuclear extract was performed by the methods of Reinberg and Roeder (63) and Lin and Green (43) as modified by Gu et al. (26). The designations of the phosphocellulose and DEAE-Sephacel fractions were as previously described (26). The bacterial expression and purification of human rTBP (39) and rTFIIB (27) were performed as described by the respective authors with minor modifications (36). Wild-type ICP4 and the deletion mutant X25 were purified from KOS- and X25-infected cells, respectively, as previously described (69, 72).

**In vitro transcription reactions.** Final conditions for the in vitro transcription reactions were 40 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES; pH 7.9); 60 mM KCl; 12% glycerol; 8.3 mM $MgCl_2$; 0.6 mM (each) ATP, CTP, GTP, and UTP; 0.3 mM dithiothreitol; and 12 U of RNasin; they have been described previously (26). The concentration of template DNA was 0.5 pmol/ml (0.08 μg of a 7.5-kb plasmid in a 30-μl reaction volume), except for the experiment in Fig. 2, for which the indicated amounts of DNA were used. The amounts of the individual transcription factors used in the reactions were 0.5 μl of AB (TFIIA), 6 μl of CB (TFIIE/F), 4 μl of CC (polymerase II), 0.5 μl of rTFIIB, and 5 μl of DB (TFIID). A 1-μl portion of ICP4 (0.1 μg/μl) was added to the transcription reaction mixtures after the mixing of the general factors and before the addition of the remaining components described above. After incubation of the transcription reaction mixtures at 30°C for 80 min, the synthesized RNAs were extracted twice with phenol and precipitated with ethanol. The conditions for primer extension reactions were as previously described (35). After phenol extraction and ethanol precipitation, the reverse transcription products were separated on 5% denaturing polyacrylamide gels. Autoradiographic images of the gels were established by exposing the dried gel to XAR films, and the signals were quantified by using a GS 300 scanning densitometer.

**Gel mobility shift assays.** DNA-binding reactions with end-labeled DNA probes and purified ICP4 were carried out in a total volume of 30 μl. The buffer consisted of 10 mM HEPES (pH 7.9), 5 mM ammonium sulfate, 8% (vol/vol) glycerol, 2% (wt/vol) polyethylene glycol 6000, 50 mM KCl, 5 mM β-mercaptoethanol, 0.2 mM EDTA, and 25 μg of poly(dGC) per ml. When monoclonal antibody against ICP4 was used, 1 μl of a 1/10 dilution of 58S (73) was added simultaneously to the reaction mixture. The reaction mixtures were incubated at 30°C for 30 min, and the products were separated at 200 V on a native 4% polyacrylamide gel containing 0.5× Tris-borate-EDTA (TBE). The gel was then dried and exposed to XAR 5 film.

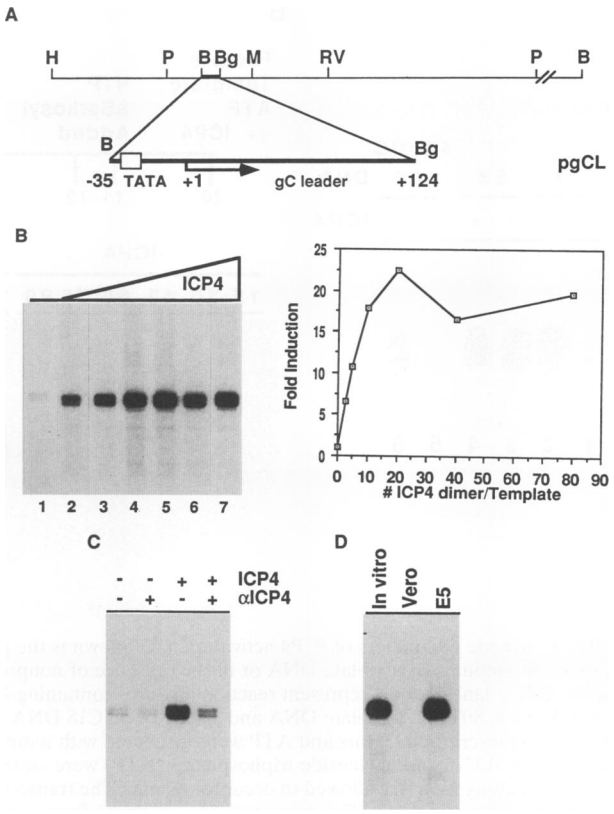


FIG. 1. Diagram and activation of the gC promoter templates by ICP4 in vitro. (A) The position of the gC promoter relative to sequences of tk is shown, as well as relevant restriction sites. The gC promoter region (bold line) from −35 to +124 relative to the gC transcription start was used to replace the tk sequences from −111 to +55 in pLSWT (see Materials and Methods). Thin lines represent the tk DNA upstream and downstream of the gC promoter. Abbreviations: H, *Hin*dIII; P, *Pst*I; B, *Bam*HI; M, *Mlu*I; RV, *Eco*RV; Bg, *Bgl*II. (B) Standard transcription reactions were carried out with 80 ng of pgCL DNA (0.02 pmol) and increasing amounts of ICP4. The in vitro synthesized RNA was then subjected to primer extension, and the products were separated on a 5% sequencing gel. The autoradiograph of the primer extension products is shown on the left and the quantitative representation of the data is shown on the right. The amounts of ICP4 used in the reactions are as follows: lane 1, no ICP4; lane 2, 0.04 pmol; lane 3, 0.08 pmol; lane 4, 0.16 pmol; lane 5, 0.32 pmol, lane 6, 0.64 pmol; and lane 7, 1.2 pmol. In the graphic representation this is represented as the number of ICP4 dimers per template. (C) The presence and absence of 1 μl of ICP4 (0.1 μg = 0.3 pmol) and 1 μl of serum in standard transcription reactions with pgCL are indicated on top of the autoradiograph. (D) RNA (10 μg) was isolated from Vero and E5 cells infected with the viruses bearing the gC construct at 12 h postinfection, subjected to primer extension analysis, and compared with the extension products from RNA obtained from a standard in vitro reaction with pgCL as template in the presence of ICP4.

## RESULTS

The transcriptional activation function of ICP4 in a reconstituted in vitro transcription system was studied with the HSV gC gene promoter as target. The promoter constructs used were inserted into a plasmid encoding the tk gene such that the native tk promoter was replaced by the gC test promoter (Fig. 1A). This approach was taken to allow for the easy insertion of the test promoter into the HSV genome for the purpose of

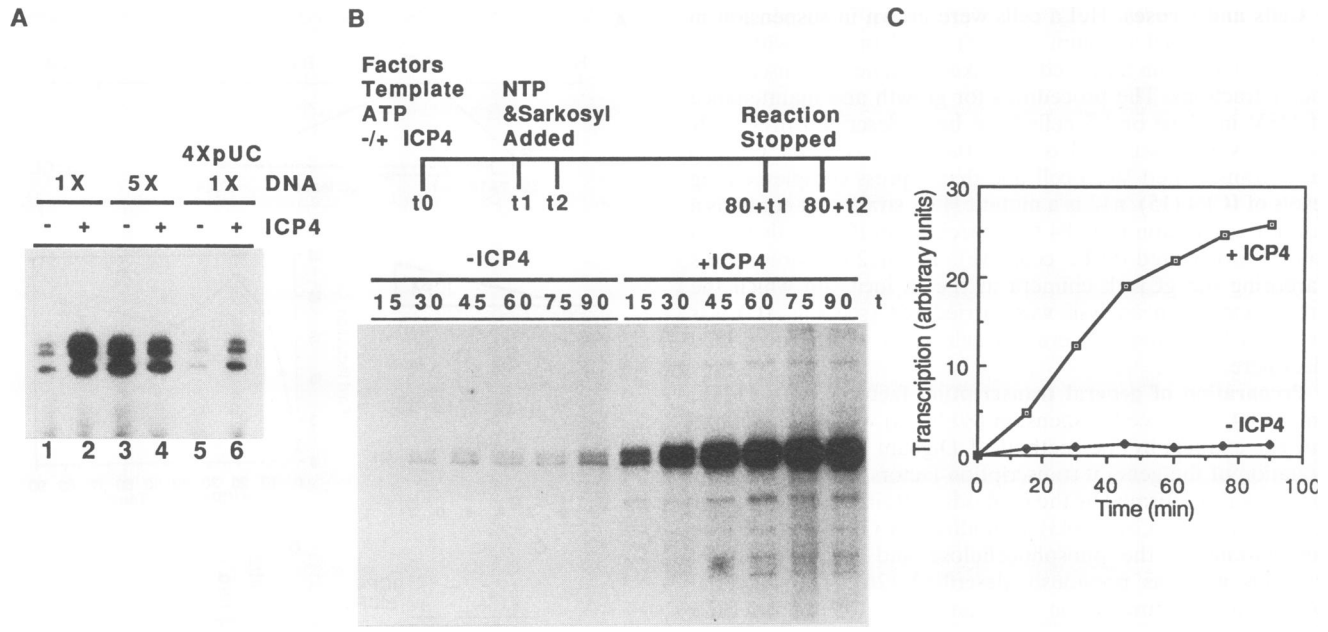


FIG. 2. Kinetic parameters of ICP4 activation. (A) Shown is the primer extension analysis of RNA from standard transcription reactions with two different amounts of template DNA or in the presence of nonpromoter DNA. Lanes 1 and 2 represent reaction mixtures containing 80 ng of template DNA; lanes 3 and 4 represent reaction mixtures containing 400 ng of template DNA, as indicated by 5×. In reaction mixtures represented in lanes 5 and 6, 80 ng of template DNA and 320 ng of pUC18 DNA (indicated by 4×) were present. (B) The scheme for the experiment is shown on the top. Transcription factors and ATP were incubated with template in the presence or absence of ICP4. At time t, 3 μl of 1% Sarkosyl (final concentration, 0.1%) and nucleoside triphosphates (NTP) were added to prevent the further formation of transcription initiation complexes, and transcription elongation was allowed to occur for 80 min. The transcripts were then analyzed by primer extension, and the resulting autoradiograph is shown at the bottom. (C) The primer extension signal from panel B was quantitated by a GS 300 Hoefer Scientific Instruments scanning densitometer. The data are plotted with the intensity of the signal band on the *y* axis (relative transcription) and the preincubation time on the *x* axis (minutes).

conducting comparative in vivo experiments and also to permit analysis of gC-tk chimeric promoters. The experiments described below address the transcription factor requirements and kinetic parameters for activation by ICP4, the role of ICP4 DNA binding, and the contributions of the TATA box and inr to activation in vitro.

**ICP4 activates transcription from the gC promoter in vitro.** Fractionated HeLa cell nuclear extracts were used as the source of basal transcription factors along with purified recombinant human TFIIB made in *E. coli*. After basal transcription had been reconstituted with the cellular factors by using an intact plasmid containing the gC promoter as template (pgCL), purified ICP4 from wild-type-virus-infected cells was added to the transcription mixture. ICP4 strongly activated transcription from the gC promoter in a concentration-dependent manner (Fig. 1B). Activation was observed on the addition of as few as two dimer molecules of ICP4 per template. There was a sharp increase in transcription, up to 20-fold, with the addition of between 2 and 10 ICP4 dimers per template, with saturation occurring at higher levels of ICP4. To confirm that the activation of transcription from the gC promoter was specifically a function of the activity of ICP4, polyclonal serum against ICP4 was included in the transcription reactions (Fig. 1C). In the absence of ICP4, the amount of serum used in the reactions had little effect on basal transcription; however, activation was almost completely abrogated by inclusion of the antibody. Lastly, the transcription start sites of the gCL template in vitro and in the context of viral infection were compared. The gC-tk chimeric construct gCL was inserted into the n12 HSV genome, which does not express a functional ICP4 (16). As a consequence, primer extension signals were obtained from RNA isolated only from infected ICP4-producing cells (E5) but not from its parent cell (Vero), demonstrating that ICP4 activated transcription in vivo as it did in the in vitro transcription experiments (Fig. 1D). Moreover, the primer extension products obtained from in vitro transcribed RNA were the same as those obtained from the RNA isolated from infected cells. These results document the authenticity of the in vitro experiments with respect to induction by ICP4 as well as for the start of transcription.

**Kinetic parameters for activation of the gC promoter by ICP4.** The effect of template concentration and nonspecific competitor DNA on the ability of ICP4 to activate transcription was assessed first. Increasing the amount of template DNA could potentially drive the formation of preinitiation complexes or titrate out molecules that act to repress transcription, thereby bypassing the requirement for ICP4 to achieve high levels of transcription and operationally reducing its effectiveness. Consistent with this hypothesis, significantly higher levels of transcription were achieved when a fivefold-higher concentration of gC template DNA was used (Fig. 2A). Furthermore, also consistent with results obtained with the pseudorabies IE protein (1), the amount of ICP4 that activated transcription with low concentrations of template DNA did not further increase the already high levels of transcription in the presence of high concentrations of template DNA. The pseudorabies virus IE protein is the functional analog to the ICP4 protein of HSV. However, unlike the previous results obtained with the pseudorabies virus IE protein, the addition of a fourfold excess of nontemplate DNA (pUC) did not result in

elevated levels of transcription and did not reduce the effectiveness of, in this case, the ICP4 protein. The difference between the two results may not reflect a difference in the function of the proteins but, rather, may reflect the use of nuclear extracts as opposed to fractionated factors and the possible involvement of chromatin or inhibitors in the previous study. Therefore, while the viral factors may be capable of relieving the effects of inhibitory proteins in a manner previously documented for other factors (12), these experiments demonstrate that the ICP4 protein can directly activate transcription, possibly by facilitating the formation of initiation complexes.

To test the hypothesis that ICP4 acts to promote the formation of initiation complexes, the experiment outlined in Fig. 2B was performed. Template and basal transcription factors were incubated for different lengths of time in the absence and presence of ICP4. At the different times, the formation of additional initiation complexes was inhibited by the addition of Sarkosyl, and elongation was then allowed to proceed for an additional 80 min. It has been shown that low concentrations of Sarkosyl inhibit the formation of transcription initiation complexes but that elongation of a committed transcription complex is unaffected (29). Therefore, the quantity of primer extension product is a measurement of the number of transcription initiation complexes formed in the time interval prior to the addition of Sarkosyl. Figure 2B and C shows the primer extension products observed and their quantitative representation, respectively. In the absence of ICP4, the level of transcription increased very slowly with time and remained relatively low even after prolonged incubation, indicating that transcription initiation complexes formed slowly. In the presence of ICP4, the transcription signal was easily detected at the early time points and the level of signal increased at a much greater rate than in the absence of ICP4. These results indicate that ICP4 induces transcription by increasing the rate of transcription initiation complex formation and hence the number of complexes formed at a given time. This conclusion is consistent with that previously made for the action of ICP4 in nuclear extracts on the gD promoter (3).

**Transcription factor and ICP4 domain requirements for activation.** If ICP4 acts at a stage in the formation of initiation complexes, it is conceivable that it may substitute for one or more of the general factors. To address this possibility, transcription reactions were conducted in the presence and absence of ICP4, with each of the different general transcription fractions omitted (Fig. 3A). Similar to what we have previously observed with the ICP4 promoter (26), transcription from the gC promoter in the absence of the TFIIA fraction yielded similar levels of transcripts to those seen in the presence of TFIIA fraction, indicating that the TFIIA fraction (AB) was not absolutely necessary in our in vitro transcription reactions. When omitted, the CC fraction, which contains RNA polymerase II, resulted in only a partial decrease in transcription, possibly because of its presence in the CB fraction as well. When the TFIIE/F fraction (CB) was omitted, a significant decrease in the transcription signal was observed. CC fraction contains trace amounts of TFIIE/F, which would account for the low-level transcription seen in the absence of CB. When either the DB (TFIID) fraction or rTFIIB was omitted, no transcription signals were detected. The same factor dependence was also seen in the absence of ICP4, although the absolute level of transcription was lower (sixth to eleventh lanes from left). The data indicated that the TFIID, TFIIE/F, and TFIIB fractions were all required to achieve full ICP4-activated transcription.

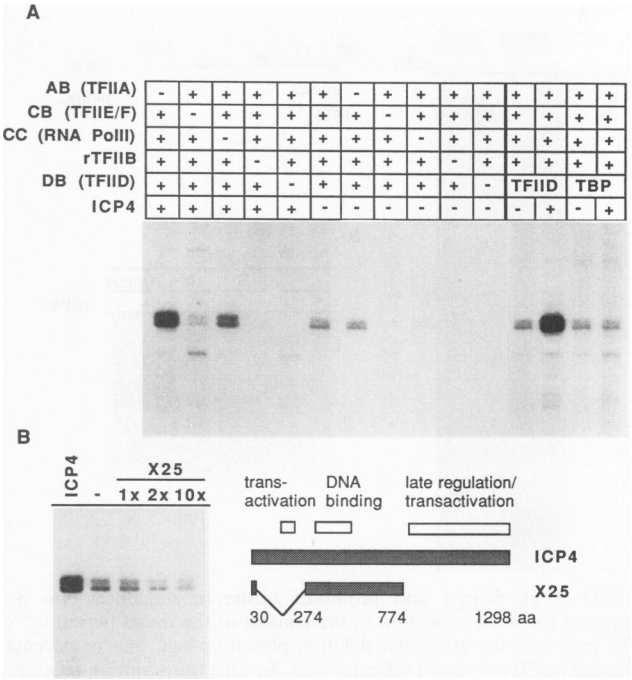


FIG. 3. Factor and domain requirements for ICP4 function. (A) Primer extension analysis of products from transcription reactions was conducted with the template pgCL and the indicated fractions and protein factors. The first five lanes represent reactions in the presence of ICP4, each of which has one protein fraction omitted, as indicated by −. The sixth to eleventh lanes from the left represent the same reaction conditions above but in the absence of ICP4. The last two lanes represent primer extension results from reactions in which TBP was used in place of TFIID. (B) Schematic representations of the wild-type ICP4 and mutant X25 proteins are shown on the right. Wild-type ICP4 is 1,298 amino acids (aa) long, and the positions of the DNA-binding, *trans*-activation, and late regulation/*trans*-activation domains are shown. The mutant X25 protein has two deletions, Δ30–274, and Δ774–1298, which are indicated. The primer extension analysis of transcription reactions with X25 is shown on the left. The amounts of ICP4 and X25 used are as follows (from left to right): 0.3 pmol of ICP4, no ICP4, 0.3 pmol of X25, 0.6 pmol of X25, and 3 pmol of X25.

The DB fraction contains TFIID, which is a multiprotein complex composed of TBP and 5 to 10 TAFs. Many transcriptional activators require TAFs in TFIID for transcriptional activation, as operationally defined by the inability of TBP alone to support activated transcription. Figure 3A shows that ICP4 also requires TAFs, since activation did not occur when rTBP was substituted for TFIID.

Transcriptional activator proteins often contain domains that communicate with basal factors and/or TAFs to activate transcription. We have previously shown that a mutant ICP4 protein, termed X25 (72), which retains the ability to bind to DNA but lacks the domains required for activation in the context of viral infection (71, 72), also does not form complexes with TBP and TFIIB on DNA (76). Consistent with in vivo results, X25 failed to activate transcription (Fig. 3B). Even when 10 times more X25 protein was included in the transcription reactions, there was no increase in transcription above basal levels. Therefore, DNA binding is not sufficient for activation, and domains that are required for activation in vivo and also for the formation of complexes with TFIIB and TBP are also required for activation in vitro.

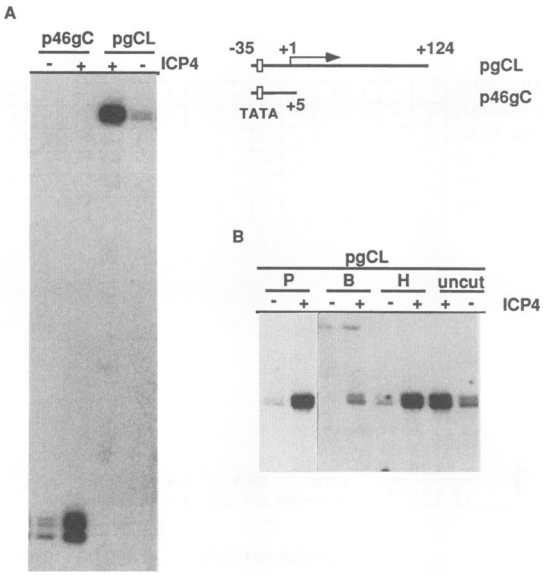


FIG. 4. gC leader and promoter upstream sequences are not required for ICP4 activation in the context of the intact plasmid. (A) The inserts in the gC leader deletion plasmid p46gC and its parental plasmid pgCL are shown schematically. In vitro transcription reactions were conducted with equal amounts of p46gC (lanes 1 and 2) and pgCL (lanes 3 and 4) as template in the presence and absence of ICP4, and the resulting transcripts were analyzed by primer extension. Since p46gC lacks the gC leader sequence, its transcripts gave primer extension products that are 119 bases shorter than those from pgCL. (B) Primer extension analysis of transcripts from in vitro reactions with restriction enzyme-digested pgCL as the template. The restriction enzymes used to digest pgCL are indicated above the gel (P, *Pst*I; B, *Bam*HI; H, *Hin*dIII). The positions of the restriction sites are indicated in Fig. 1A. The presence and absence of ICP4 are indicated by + and −, respectively.

**Contribution of ICP4 DNA binding to activation of the gC promoter.** The question of the role of ICP4-binding sites in activation has been extremely difficult to address by genetic studies in the context of viral infection and in transfection studies as well. From such studies it is generally agreed that no single site or any collection of sites that have been examined affect ICP4 activation. This is possibly a function of the degeneracy of the ICP4-binding-site consensus, ensuring that all the templates tested in the past always contained binding sites that have an affinity for ICP4 that is greater than their affinity for nonspecific DNA. The in vitro studies described below address these issues.

The above results demonstrated that gC DNA sequences, including the TATA box and the 5′ untranslated leader from −35 to +124, when located between −111 and +55 of tk, were sufficient for ICP4-mediated activation of transcription. The potential specific role of the gC leader sequence in activation was examined by constructing a template plasmid that lack the leader sequences and performing comparative in vitro transcription reactions in the presence and absence of ICP4 (Fig. 4A). The plasmid, p46gC, contains gC promoter sequences from −35 to +5, including the gC TATA box and the transcription initiation site to +5, but lacks sequences from +6 to +124. The levels of basal and activated transcription from the gC promoter in p46gC were similar to those from pgCL, and the resulting transcripts were generated from the same initiation site(s). Therefore, in the context of the intact plasmid template, a gC promoter consisting of 39 bp of DNA is sufficient to direct both basal and ICP4-activated transcription in vitro, and specific sites within the leader are not required.

To consider whether there were any *cis* DNA elements upstream of the gC promoter in the plasmid contributing to basal and activated transcription, the plasmid pgCL was digested with restriction enzymes that cut at various locations upstream of the TATA box (Fig. 1A) and used as templates in transcription reactions in the presence and absence of ICP4 (Fig. 4B). The linear DNA templates produced by *Hin*dIII or *Pst*I digestion resulted in similar levels of basal and ICP4-induced transcription to that seen with uncut templates. The result indicates that the sequences upstream of the *Pst*I site are not required for basal or ICP4-activated transcription. Previous studies involving gel mobility shift assays have demonstrated that there are no ICP4-binding sites in the region between the *Pst*I and the 5′ side of the gC TATA box (37). To further assess any potential role of these sequences, the template was digested with *Bam*HI, which cuts 4 bp upstream of the gC TATA box. It was found that ICP4 activated transcription even when the template was truncated just 5′ to the TATA box, confirming that upstream sequences are not necessary. However, in these reactions, basal and ICP4-activated transcription were generally decreased most probably because of the short distance between the end of the DNA and the TATA box, reducing the ability of a general transcription complex to form.

Similar experiments to those described above were conducted to examine the involvement of sequences downstream of the transcription initiation site. The templates were generated by digesting gC plasmid constructs with *Pst*I and *Mlu*I (Fig. 5) and purifying the resulting promoter-containing fragments from polyacrylamide gels. Transcription from the fragment containing the whole gC leader sequence (template 1) was efficiently activated by ICP4, as was transcription from template 2, albeit to a lesser extent (Fig. 5). Template 2 is deleted for sequences between +85 and +124. However, templates 3 and 4, which are deleted for sequences between +48 and +124 and between +5 and +124, respectively, were not activated by ICP4. These data indicate that there is an important *cis* site for activation located between +48 and +85 and that deletion of this region in the leader is compensated for by other sites provided in *cis* in the intact plasmid template (Fig. 4A). To address this possibility, templates that are deleted between +5 and +124 and that have increasing extents of DNA downstream of the *Mlu*I site were prepared and tested in vitro. When DNA downstream of the *Mlu*I site was appended to the template, activation was restored (templates 5 and 6). Moreover, the level of activation increased with increasing extents of DNA downstream of the *Mlu*I site. Therefore, the deletion of DNA possibly containing specific sites for activation can be compensated for by the presence of distally located sites, or ICP4 may just require a certain linear extent of DNA to function, independent of specific binding sites.

ICP4 has a remarkably degenerate binding-site consensus that has been summarized to be, roughly, RTCGTCNNY NYSG (R = purine; N = A, T, C, or G; Y = pyrimidine; and S = G or C) (17). Therefore, it is possible that such sites exist in the regions found to be important for activation in the above experiments. Indeed, a site having the sequence GTGGTC CGTGTGG has been previously described in the region between +48 and +85 (17). To investigate the existence and possible function of specific ICP4-binding sites in templates that gave activation, DNA gel mobility shift experiments were carried out with purified ICP4 and DNA fragments derived from the templates used in the above experiments. Initial

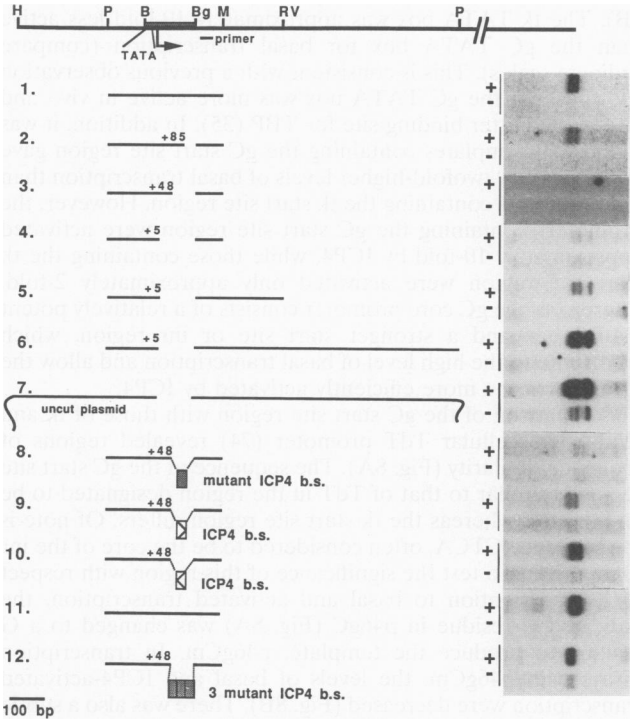


FIG. 5. Contribution of *cis* sites to activation by ICP4. Shown are primer extension data from transcription reactions conducted with DNA fragments as templates. The DNA fragments were generated by digesting gC promoter-containing plasmids with restriction enzymes, and the indicated fragments were purified from polyacrylamide gels as described in Materials and Methods. The templates used are as follows: 1, *Pst*I-*Mlu*I fragment from pgCL; 2, *Pst*I-*Mlu*I fragment from pgCΔ(85–124); 3, *Pst*I-*Mlu*I fragment from pgCΔ(48–124); 4, *Pst*I-*Mlu*I fragment from p46gC; 5, *Pst*I-*Eco*RV fragment from p46gC; 6, *Pst*I-*Pst*I fragment from p46gC; 7, intact p46gC; 8, same as template 3 except containing a 27-bp mutant ICP4-binding site (b.s.); 9, same as template 3 except containing a 30-bp ICP4-binding site; 10, same as template 3 except containing a 30-bp ICP4-binding site in the reverse orientation; 11, same as template 1; 12, same as template 3 except containing three tandem 27-bp mutant ICP4-binding sites. The intact ICP4 binding site has the sequence from the cap site of the ICP4 gene. The ΔATC-binding-site mutant does not bind ICP4 and is nonfunctional as a repressor site (16, 64). Since the primer used for reverse transcription locates 3′ to the gC leader, the primer extension products obtained from some of the deletion templates are different in size (Fig. 4). For ease of representation, the regions of signal for templates 1 to 3 were cut from the same gel and pasted alongside each other. The presence and absence of ICP4 in the reactions are indicated by + and −, respectively. The restriction sites are those indicated in Fig. 1.

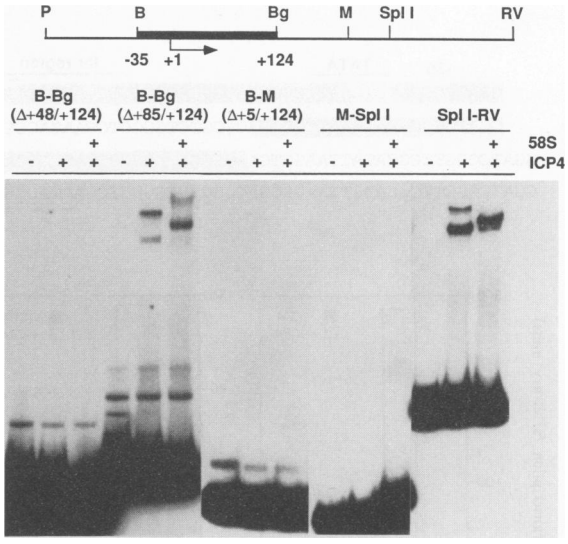


FIG. 6. ICP4-binding sites in the gC leader and in the tk sequences 3′ to the gC promoter. At the top of the figure, the gC-tk fusion region is schematically shown with the relevant restriction sites indicated as described in the legend to Fig. 1. gC sequences are shown by the bold line, and tk sequences are shown by the thin line. The end-labeled DNA fragments are designated by their restriction ends and their respective deletions in the gC leader region. The DNA fragments used as probes are indicated above each set of mobility shifts. The presence of ICP4 (0.1 μg) and 58S anti-ICP4 monoclonal antibody (1 μl of a 1/10 dilution) in the binding-reaction mixtures used to generate the mobility shifts is appropriately indicated. Abbreviations: P, *Pst*I; B, *Bam*HI; Bg, *Bgl*II; M, *Mlu*I; RV, *Eco*RV.

studies showed that a probe containing the whole gC promoter and leader sequence from −35 to +124 was bound by ICP4 (data not shown). An analogous probe with the sequences between +85 and +124 deleted was also bound by ICP4, whereas the probe deleted for the sequence between +48 and +124 failed to bind ICP4 (Fig. 6). The existence of this site is consistent with the previous report mentioned above and also represents a positive correlation with the functional data for these templates. Since it was found that sequences downstream of the *Mlu*I site could compensate for the deletion of the leader sequence for activation, the existence of ICP4-binding sites in this region was investigated. The results in Fig. 6 show that there is a specific ICP4-binding site present between the *Spl*I and *Eco*RV sites, reaffirming an apparent correlation between the presence of an ICP4-binding site in a template and the ability of the template to be activated by ICP4.

Given the apparent correlation between the ability of ICP4 to activate a template and the presence of an ICP4-binding site, we investigated the possibility that the addition of a strong ICP4-binding site to a template that was not activated by ICP4 would result in activation. The relatively strong ICP4-binding sequence from the start site of the ICP4 promoter region (23, 41, 51) was introduced into template 3 at the *Bgl*II site in both orientations as a 30-mer oligonucleotide (Fig. 5, templates 9 and 10). This site is normally responsible for autorepression of the ICP4 promoter in transient expression assays, in the viral genome, and in vitro (26, 49, 64, 65). Templates 9 and 10 were activated by five- and ninefold, respectively, while a similar template (template 8) containing a 27-mer representing a nonfunctional ICP4-binding site (ΔATC) was not activated. It is interesting that the ICP4-binding site conferred activation when inserted in either direction. Similar results were obtained when the ICP4-binding site was inserted in template 4 (data not shown). These results demonstrate that an ICP4-binding sequence which normally acts to repress transcription could also serve to enhance transcription.

Finally, as a test of whether specific ICP4-binding sites were absolutely necessary for activation and to address the possibility that greater extents of nonspecific DNA compensate for the lack of a specific binding site, three copies of the 27-mer mutant binding-site oligonucleotide were inserted at the *Bgl*II site in template 3, to generate template 12. It should be noted that it has been previously shown that a single copy of the mutated ICP4-binding site will not specifically bind to ICP4

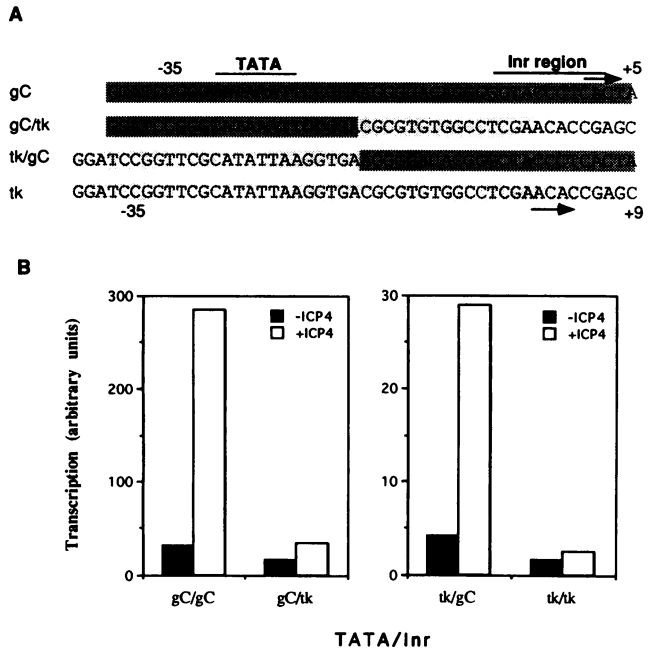


FIG. 7. ICP4 activates the gC and tk promoters differently. (A) Sequences of the promoter region of gC, tk, and gC-tk hybrids are aligned around their TATA box sequence to show the junction of the TATA box–start site hybrid promoters. The dark- and light-shaded areas represent gC and tk sequences, respectively. The transcription start sites are shown for tk and gC by the arrows. (B) Quantitation of in vitro transcription and primer extension assays with gC/gC, gC/tk, tk/gC, and tk/tk templates are represented by bar graphs. The data for the gC/gC and gC/tk templates are shown at a 10-fold-larger scale than those for the tk/gC and tk/tk templates, because of the difference in the activities of the gC and tk TATA boxes.

and will not repress transcription in transient expression assays, in the viral genome, and in vitro (16, 26, 65). It was found that templates 8 (one mutant binding site) and 12 (three mutant binding sites) were as deficient in specific binding to ICP4, so that the trimerization of the mutant site did not fortuitously create a new site (data not shown). However, in contrast to template 8, which was not induced, transcription from template 12 was induced in the presence of ICP4 (Fig. 5). Template 11 is identical to template 1, representing an experiment conducted simultaneously with template 12. The reason for this representation is that templates 11 and 12 are approximately the same size, as indicated by the mobility of the extension products, but template 11 contains the specific gC leader ICP4-binding site. In repeated experiments, template 11 was activated 10-fold while template 12 was activated by approximately 4-fold.

**Role of the start site region and TATA box in basal and ICP4-activated transcription.** Preliminary in vitro transcription results showed that the gC promoter was much stronger than the tk TATA box promoter for basal transcription and was activated to a greater extent by ICP4. The basis for this difference was investigated by conducting transcription experiments with several chimeric promoters between the gC and tk sequences. The chimeric promoters paired the start site region and TATA box of the two promoters. The templates are designated for the TATA box/start site region they contain and are shown in Fig. 7A. The indicated templates were used in transcription reactions, and the primer extension signals obtained in these reactions were quantified by densitometry (Fig. 7B). The tk TATA box was approximately 10-fold less active than the gC TATA box for basal transcription (compare ordinate scales). This is consistent with a previous observation showing that the gC TATA box was more active in vivo and also was a better binding site for TBP (35). In addition, it was found that templates containing the gC start site region gave approximately twofold-higher levels of basal transcription than did promoters containing the tk start site region. However, the promoters containing the gC start site region were activated approximately 10-fold by ICP4, while those containing the tk start site region were activated only approximately 2-fold. Therefore, the gC core promoter consists of a relatively potent TATA box and a stronger start site or inr region, which contribute to the high level of basal transcription and allow the promoter to be more efficiently activated by ICP4.

Comparison of the gC start site region with those of tk and TATA-less cellular TdT promoter (74) revealed regions of sequence similarity (Fig. 8A). The sequence of the gC start site region is similar to that of TdT in the region designated to be an initiator, whereas the tk start site region differs. Of note is the sequence CTCA, often considered to be the core of the inr region (84). To test the significance of this region with respect to its contribution to basal and activated transcription, the indicated C residue in p46gC (Fig. 8A) was changed to a G residue to produce the template, p46gCm. In transcription assays with p46gCm, the levels of basal and ICP4-activated transcription were decreased (Fig. 8B). There was also a subtle effect on the position of the measured transcription start site. The net level of activation by ICP4 of p46gCm was approximately 2.5-fold, a situation similar to that for tk (Fig. 7B). This experiment implicates the gC start site region as an important determinant in basal and, more importantly, ICP4-activated transcription and also raises the possibility that this region functions as an initiator.

The inr regions can direct accurate transcription in the absence of a TATA box (74). To determine if the gC start site region can indeed act as an inr element and to determine if ICP4 can activate inr-mediated transcription, the TATA box sequences from p46gC and p46gCm were deleted by eliminating 13 bp between the *Bam*HI and *Bsp*EI (−20) sites. In preliminary experiments, a very low level of transcription was observed with the TATA box deletions and fractionated factors (data not shown). Accordingly, the tk upstream region was appended to the promoters and nuclear extracts were used as a source of factors. The intact gC promoter consisting of the TATA box and the inr gave measurable levels of basal transcription and substantial levels of ICP4-activated transcription; however, the TATA-less template also produced detectable signal in the presence of ICP4 (Fig. 8C). Transcription in the presence of ICP4 was not observed from the TATA-less promoter when the initiator was mutated by the single C-to-G change. Therefore, given the sequence similarity of the gC start site region of the TdT inr, the ability of the gC start site region to allow accurate transcription in the absence of the TATA box, and the observation that a single-base-pair change within the start site region can abrogate transcriptional activity in the absence of the TATA box (Fig. 8C), this region of gC may be considered an initiator. Moreover, ICP4 can activate transcription as a function of the inr, independently of the TATA box.

## DISCUSSION

The experiments in this study document the authentic reconstruction of the activation of the gC promoter by ICP4 in vitro with fractionated HeLa nuclear factors. ICP4 efficiently

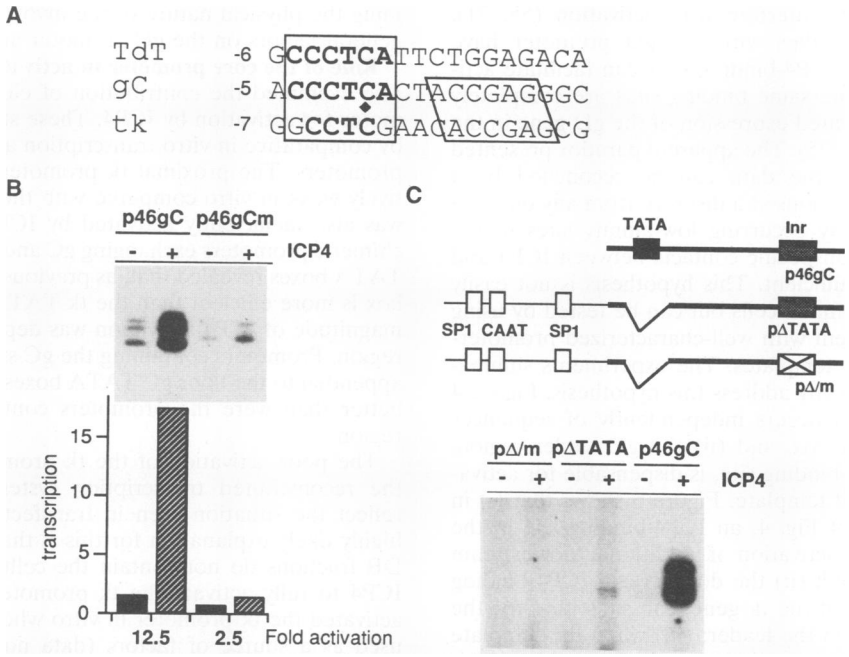


FIG. 8. The gC promoter contains an initiator that contributes to ICP4 activation. (A) Sequences of the transcription start site region of gC, tk, and TdT promoters aligned as indicated. The homologous bases among gC, tk, and TdT are boxed, with the conserved **CCCTCA** in boldface letters. tk differs from TdT and gC in two positions, including the highly conserved **A** found in most initiators. Another region of 6 bases between tk and gC that are identical are also boxed. The indicated nucleotide **C** in the gC sequence was changed to a **G** residue in p46gCm and used in transcription reactions in panel B. (B) in vitro transcription reactions were carried out with p46gCm and p46gC, and primer extension results are shown in the autoradiograph. Quantitation of the gel and the calculated fold activation of transcription by ICP4 for each template are also shown below the gel. p46gCm has a single C-to-G change at the position indicated in panel A. (C) The two templates containing 13-bp deletions, including the TATA box, pΔTATA, and pΔ/m, were derived from p46gC and p46gCm, respectively (see Materials and Methods). pΔTATA, pΔ/m, and p46gC are schematically shown at the top. Primer extension results from RNA obtained with these templates are shown at the bottom. Transcription reactions for data presented in this figure were carried out with nuclear extract to provide Sp1. The presence and absence of ICP4 are indicated by + and −, respectively.

activated the gC promoter, giving rise to start sites that were indistinguishable from those observed in the context of viral infection. Activation occurred by facilitating the formation of initiation complexes on the gC promoter and/or by relieving repression. This was inferred by allowing complexes to form in the presence and absence of ICP4 without allowing elongation and then inhibiting further complex formation and allowing the elongation from the preformed initiation complexes. While the approach and quantitative aspects differ, these results are in agreement with previously reported data (1, 3). However, in contrast to previous results (1), the inclusion of nonpromoter DNA in the reaction did not bypass the requirement for ICP4. This implies that ICP4 not only can relieve the negative effects of nonspecific repressor molecules but also can directly activate transcription by facilitating the formation of initiation complexes. Lastly, an ICP4 mutant protein that retains its ability to dimerize and bind to DNA but does not activate transcription in infected cells (71, 72) was also unable to activate transcription in vitro. Therefore, ICP4 was able to truly activate transcription in a manner consistent with its activity inferred from genetic experiments in intact cells.

By using fractionated factors, it was possible to infer that ICP4 does not eliminate the requirement for any individual general transcription factor. Therefore, the function of ICP4 is to augment the activity of these factors and not to functionally replace them. Additionally, when TBP was substituted for TFIID in the reactions, basal transcription occurred, which was not activated by ICP4. This result suggests that like many other activator proteins, such as Sp1 (61) and VP16 (86), ICP4 requires the TAFs for activation. It has also been previously shown that ICP4 will not function in its repressor mode when TBP is substituted for TFIID (26), again indicating the importance of TAFs for ICP4 function. Given that ICP4 has been shown to form tripartite complexes on DNA with TBP and TFIIB and that the presence of ICP4 will lower the amount of TBP needed to detect occupancy of the TATA box (76), these results present an apparent paradox that is not unique to ICP4. E1A (34, 42) and VP16 (43, 77) have been shown to interact with TBP and with TBP and TFIIB, respectively, yet, as mentioned above, both require TAFs for activation (25, 86). It is possible that the interactions observed in vitro are not strong enough to be functionally significant on their own and that additional interactions with TAFs are necessary or that initiation complexes formed in the absence of TAFs are not in the proper conformation to allow activation to occur. It has been shown that VP16 interacts in a functionally significant way with a 40-kDa TAF, suggesting that activation is mediated through interactions with multiple proteins in the initiation complex (25, 79). Studies aimed at exploring the possible interactions between ICP4 and individual TAFs should shed further light on the mechanism of ICP4 function.

**Role of DNA binding in the activation of the gC promoter by ICP4.** Studies conducted with infected or transfected cells have shown that no single sequence or ICP4-binding site or any collection of sequences or sites confer activation (10, 20, 22, 28). On the other hand, most mutations in ICP4 that eliminate

its DNA-binding activity interfere with activation (55, 71). Additionally, in vitro studies with the gD promoter have concluded that multiple ICP4-binding sites can facilitate activation (78). However, the same binding sites are completely dispensable for full activated expression of the gD gene in the context of viral infection (75). The apparent paradox presented by consideration of all the data can be reconciled by a hypothesis that ICP4 functions at a distance from any one of a number of the frequently occurring low-affinity sites in the HSV genome or that nonspecific contacts between ICP4 and DNA can, in part, be sufficient. This hypothesis is not easily tested in the context of intact cells but can be tested by using an efficient in vitro system with well-characterized promoter-containing fragments as templates. The experiments summarized in Fig. 4 and 5 directly address this hypothesis. Figure 4 shows that (i) activation occurs independently of sequences upstream of the TATA box, and (ii) the gC leader region, which contains an ICP4-binding site, is dispensable for activation in an intact plasmid template. Figure 5 shows that (i) in contrast to the results of Fig. 4, an ICP4-binding site in the leader is required for activation if additional downstream binding sites are removed; (ii) the downstream ICP4-binding sites, which are located in the tk gene, can substitute for the lack of the binding site in the leader; (iii) when the template was sufficiently truncated to eliminate all detectable ICP4-binding sites, it could no longer be activated by ICP4; and (iv) the addition of a strong ICP4-binding site that normally acts to repress transcription to the above template restored activation independently of orientation. These results demonstrate that ICP4 binding sites downstream of the transcriptional start site can act to facilitate activation, independently of their position and orientation. The elimination of any given site can be compensated for by the presence of another site.

An additional important aspect of the experiments in Fig. 5 follows from the consideration of the level of activation obtained with templates 4, 5, and 6. Template 6 has multiple ICP4-binding sites downstream of the *Eco*RV site (data not shown) and, of the three, is activated to the greatest degree. Template 5 has a single binding site between the *Mlu*I and *Eco*RV sites and is activated to a lesser degree than template 6; template 4 has no binding sites and is not activated. Therefore, more binding sites may allow for more-efficient activation, or the extent of DNA downstream of the start site is important, implying a potential role for nonspecific DNA-protein contacts. When the nonfunctional mutant ICP4-binding site was trimerized and inserted into template 3 to give template 12, activation by ICP4 resulted. Therefore, if sufficient non-site-containing DNA is present downstream of the start site, ICP4 can activate transcription. This did not occur when even greater extents of DNA were appended upstream of the TATA box (data not shown).

In summary, our results support the hypothesis that ICP4-binding sites downstream of the transcriptional start site facilitate activation of the gC promoter. The absence of individual sites can be compensated for by the presence of other downstream sites. Because of the degeneracy of the ICP4-binding consensus (17), it is highly likely that this will be true for other HSV promoters as well. In addition, ICP4 can activate transcription if sufficient nonspecific DNA is present. However, ICP4 did not activate transcription from a template that had very little DNA flanking the promoter, suggesting that for the gC promoter, ICP4 is not just acting to bridge general transcription factors functioning at the core promoter but must make direct contact with or be tethered to the DNA via interaction with another DNA-binding protein. Studies examining the physical nature of the involvement of ICP4 with the general factors on the gC promoter are currently under way.

**Role of the core promoter in activation by ICP4.** This study also examined the contribution of elements in the core promoter for activation by ICP4. These studies were precipitated by comparative in vitro transcription analysis of the tk and gC promoters. The proximal tk promoter was found to be relatively weak in vitro compared with the core gC promoter and was also inefficiently activated by ICP4 (Fig. 7). The use of chimeric promoters exchanging gC and tk start site regions and TATA boxes revealed that, as previously shown, the gC TATA box is more efficient than the tk TATA box (36) and that the magnitude of ICP4 activation was dependent on the start site region. Promoters containing the gC start site region, whether appended to the tk or gC TATA boxes, were activated fivefold better than were the promoters containing the tk start site region.

The poor activation of the tk promoter demonstrates that the reconstituted transcription system does not accurately reflect the situation seen in transfected or infected cells. A highly likely explanation for this is that the AB, CB, CC, and DB fractions do not contain the cellular proteins that allow ICP4 to fully activate the tk promoter. ICP4 very efficiently activated the tk promoter in vitro when nuclear extracts were used as a source of factors (data not shown). It will be of interest to determine if other components of the nuclear extract can augment the fractionated system and allow ICP4 to fully activate tk. For the purpose of the present study, the different responses of the tk and gC promoters suggest that ICP4 may fully activate these two promoters by different specific mechanisms. The difference is not inherent in the TATA box but appears to be a function of the start site region. Inspection of this region reveals that the gC start site region has substantial similarities to the TdT initiator (Fig. 8A), including the core CTCA (84), and that the analogous tk region deviates from this. When the CTCA of gC was changed to CTGA, the ability of ICP4 to induce transcription was markedly reduced. The role of the start site region and its function as an initiator was implied by the ability of ICP4 to activate transcription of a promoter containing the start site region but lacking a TATA box. Together, the results of experiments with the hybrid gC-tk promoters, the mutant gC inr promoter, and the inr-only promoter imply a role for this element in activation by ICP4. It is possible that proteins binding at start sites, or inr elements, such as TFIIi (66), YY1 (68), or TAFII150 (80) are involved with ICP4 in activation. tk may employ a different set of factors functioning at the start site than does gC. Coen et al. have shown that a sequence between +5 and +15 is important for optimal transcription in virus infected cells (10). Perhaps the factor(s) that functions through this region is missing in the reconstituted system to achieve full activation.

For activation of the gC promoter in vitro, ICP4 requires contacts with DNA. It has been shown previously that ICP4 is capable of interacting in a complex with TBP, TFIIB, and DNA and that a region important for this interaction is also involved in transactivation (76). Additionally, TAFs are required for activation, suggesting that ICP4 interacts with one or more of these proteins. Lastly, given that the sequence of the start site region is also important for full activation by ICP4, it is reasonable to propose that inr-binding proteins may also functionally interact with ICP4. ICP4 is a large and genetically and structurally complex protein. It is likely that it functions via contacts with many of the proteins that make up the initiation complex. Additional studies examining the potential interactions between specific TAFs, inr-binding pro-

teins, and ICP4, as well as studies to elucidate the nature of the initiation complex in the presence of ICP4, are necessary before we fully understand how this complex protein activates transcription.

## ACKNOWLEDGMENTS

We thank Ramon Rivera-Gonzalez and Stanley Person for helpful discussions and comments on the manuscript.

This work was supported by NIH grants AI27431 and AI30612 to N.D.

## REFERENCES

1. **Abmayr, S. M., J. L. Workman, and R. G. Roeder.** 1988. The pseudorabies immediate early protein stimulates in vitro transcription by facilitating TFIID:promoter interactions. Genes Dev. **2**:542–553.
2. **Batterson, W., and B. Roizman.** 1983. Characterization of the HSV-associated factor responsible for the induction of α genes. J. Virol. **46**:371–377.
3. **Beard, P., S. Faber, K. W. Wilcox, and L. I. Pizer.** 1986. Herpes simplex virus immediate early infected-cell polypeptide 4 binds to DNA and promotes transcription. Proc. Natl. Acad. Sci. USA **83**:4016–4020.
4. **Berger, S. L., W. D. Cress, A. Cress, S. J. Triezenberg, and L. Guarente.** 1990. Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: evidence for transcriptional adaptors. Cell **61**:1199–1208.
5. **Buratowski, S.** 1994. The basics of basal transcription by RNA polymerase II. Cell **77**:1–3.
6. **Buratowski, S., S. Hahn, L. Guarente, and P. A. Sharp.** 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell **56**:549–561.
7. **Campbell, M. E. M., J. W. Palfreyman, and C. M. Preston.** 1984. Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for the stimulation of immediate early transcription. J. Mol. Biol. **180**:1–19.
8. **Cheung, A. K.** 1989. DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus. Nucleic Acids Res. **17**:4637–4646.
9. **Clements, J. B., R. J. Watson, and N. M. Wilkie.** 1977. Temporal regulation of herpes simplex virus type 1 transcription: localization of transcripts on the viral genome. Cell **12**:275–285.
10. **Coen, D. M., S. P. Weinheimer, and S. L. McKnight.** 1986. A genetic approach to promoter recognition during trans induction of viral gene expression. Science **234**:53–59.
11. **Courtney, R. J., and M. Benyesh-Melnick.** 1974. Isolation and characterization of a large molecular weight polypeptide of herpes simplex virus type 1. Virology **62**:539–551.
12. **Croston, G. E., L. A. Kerrigan, L. M. Lira, D. R. Marshak, and J. T. Kadonaga.** 1991. Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. Science **251**:643–649.
13. **DeLuca, N. A., A. McCarthy, and P. A. Schaffer.** 1985. Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. J. Virol. **56**:558–570.
14. **DeLuca, N. A., and P. A. Schaffer.** 1985. Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. Mol. Cell. Biol. **5**:1997–2008.
15. **DeLuca, N. A., and P. A. Schaffer.** 1987. Activities of herpes simplex virus type 1 (HSV-1) ICP4 genes specifying nonsense peptides. Nucleic Acids Res. **15**:4491–4511.
16. **DeLuca, N. A., and P. A. Schaffer.** 1988. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol. **62**:732–743.
17. **DiDonato, J. A., J. R. Spitzner, and M. T. Muller.** 1991. A predictive model for DNA recognition by the herpes simplex virus protein ICP4. J. Mol. Biol. **219**:451–470.
18. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription by RNA polymerase II in soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
19. **Dixon, R. A. F., and P. A. Schaffer.** 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein VP175. J. Virol. **36**:189–203.
20. **Eisenberg, S. P., D. M. Coen, and S. L. McKnight.** 1985. Promoter domains required for expression of plasmid-borne copies of the herpes simplex virus thymidine kinase gene in virus-infected mouse fibroblasts and microinjected frog oocytes. Mol. Cell. Biol. **5**:1940–1947.
21. **Everett, R. D.** 1984. Transactivation of transcription by herpes virus product: requirements for two HSV-1 immediate-early polypeptides for maximum activity. EMBO J. **3**:3135–3141.
22. **Everett, R. D.** 1984. A detailed analysis of an HSV-1 early promoter: sequences involved in trans-activation by viral immediate-early gene products are not early gene specific. Nucleic Acids Res. **12**:3037–3056.
23. **Faber, S. W., and K. W. Wilcox.** 1986. Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences. Nucleic Acids Res. **14**:6067–6083.
24. **Godowski, P. J., and D. M. Knipe.** 1986. Transcriptional control of herpes virus gene expression: gene functions required for positive and negative regulation. Proc. Natl. Acad. Sci. USA **83**:256–260.
25. **Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian.** 1993. Drosophila TAFii40 interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell **75**:519–530.
26. **Gu, B., R. Rivera-Gonzalez, C. A. Smith, and N. A. DeLuca.** 1993. Herpes simplex virus infected cell polypeptide 4 preferentially represses Sp1-activated over basal transcription from its own promoter. Proc. Natl. Acad. Sci. USA **90**:9528–9532.
27. **Ha, I., W. S. Lane, and D. Reinberg.** 1991. Cloning of a human gene encoding the general transcription initiation factor IIB. Nature (London) **352**:689–695.
28. **Halpern, M. E., and J. R. Smiley.** 1984. Effects of deletions on expression of the herpes simplex virus thymidine kinase gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. J. Virol. **50**:733–738.
29. **Hawley, D. K., and R. G. Roeder.** 1985. Separation and partial characterization of three functional steps in transcription initiation by human RNA polymerase II. J. Biol. Chem. **260**:8163–8172.
30. **Hernandez, N.** 1993. TBP, a universal eukaryotic transcription factor? Genes Dev. **7**:1291–1308.
31. **Homa, F. L., T. M. Otal, J. C. Glorioso, and M. Levine.** 1986. Transcriptional control signals of a herpes simplex virus type 1 late (γ2) gene lie within bases −34 to +124 relative to the 5′ terminus of the mRNA. Mol. Cell. Biol. **6**:3652–3666.
32. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpes virus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14**:8–19.
33. **Honess, R. W., and B. Roizman.** 1975. Regulation of herpes virus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA **72**:1276–1280.
34. **Horikoshi, N., K. Maguire, A. Kralli, E. Maldonado, D. Reinberg, and R. Weinmann.** 1991. Direct interaction between adenovirus E1A protein and the TATA box binding transcription factor IID. Proc. Natl. Acad. Sci. USA **88**:5124–5128.
35. **Imbalzano, A. N., D. Coen, and N. A. DeLuca.** 1991. Herpes simplex virus transactivator ICP4 operationally substitutes for the cellular transcription factor Sp1 for efficient expression of the viral thymidine kinase gene. J. Virol. **65**:565–574.
36. **Imbalzano, A. N., and N. DeLuca.** 1992. Substitution of a TATA box from a herpes simplex virus late gene in the viral thymidine kinase promoter alters ICP4 inducibility but not temporal expression. J. Virol. **66**:5453–5463.
37. **Imbalzano, A. N., A. A. Shepard, and N. A. DeLuca.** 1990. Functional relevance of specific interactions between herpes simplex virus type 1 ICP4 and sequences from the promoter-regulatory domain of the viral thymidine kinase gene. J. Virol. **64**:2620–2631.
38. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42**:559–572.
39. **Kao, C. C., P. M. Lieberman, M. C. Schmidt, Q. Zhou, R. Pei, and**

A. J. Berk. 1990. Cloning of a transcriptionally active human TATA binding factor. Science **248**:1646–1650.
40. **Koop, K., J. Duncan, and J. R. Smiley.** 1993. Binding sites for the herpes simplex virus immediate-early protein ICP4 impose an increased dependence on viral DNA replication on simple model promoters located in the viral genome. J. Virol. **67**:7254–7263.
41. **Kristie, T. M., and B. Roizman.** 1986. DNA-binding site of major regulatory protein α4 specifically associated with promoter-regulatory domains of α genes of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83**:4700–4704.
42. **Lee, W. S., C. C. Kao, G. O. Bryant, X. Liu, and A. J. Berk.** 1991. Adenovirus E1A activation domain binds the basic repeat in the TATA box transcription factor. Cell **67**:365–376.
43. **Lin, Y. S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64**:971–981.
44. **Matsui, T., J. Segall, P. A. Weil, and R. G. Roeder.** 1980. Multiple factors required for accurate initiation of transcription by RNA polymerase II. J. Biol. Chem. **255**:11992–11996.
45. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
46. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. K. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. **14**:1727–1744.
47. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**:316–324.
48. **Metzler, D. W., and K. W. Wilcox.** 1985. Isolation of herpes simplex virus regulatory protein ICP4 as a homodimeric complex. J. Virol. **55**:329–337.
49. **Michael, N., and B. Roizman.** 1993. Repression of the herpes simplex virus 1 α4 gene by its gene product occurs within the context of the viral genome and is associated with all three identified cognate sites. Proc. Natl. Acad. Sci. USA **90**:2286–2290.
50. **Michael, N., D. Spector, P. Mavromara-Nazos, T. M. Kristie, and B. Roizman.** 1988. The DNA-binding properties of the major regulatory protein α4 of herpes simplex virus. Science **239**:1531–1534.
51. **Muller, M. T.** 1987. Binding of the herpes simplex virus type 1 gene product ICP4 to its own transcription start site. J. Virol. **61**:858–865.
52. **O'Hare, P., and G. S. Hayward.** 1985. Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. J. Virol. **53**:751–760.
53. **O'Hare, P., and G. S. Hayward.** 1985. Three *trans*-acting regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. J. Virol. **56**:723–733.
54. **Paterson, T., and R. D. Everett.** 1988. Mutational dissection of the HSV-1 immediate-early protein vmw175 involved in transcriptional transactivation and repression. Virology **166**:186–196.
55. **Paterson, T., and R. D. Everett.** 1988. The regions of the herpes simplex virus type 1 immediate early protein Vmw175 required for site specific DNA binding closely correspond to those involved in transcriptional regulation. Nucleic Acids Res. **16**:11005–11025.
56. **Paterson, T., V. G. Preston, and R. D. Everett.** 1990. A mutant of herpes simplex virus type 1 immediate early polypeptide Vmw175 binds to the cap site of its own promoter in vitro but fails to autoregulate in vivo. J. Gen. Virol. **71**:851–861.
57. **Pereira, L., M. H. Wolff, M. Fenwick, and B. Roizman.** 1977. Regulation of herpes virus macromolecular synthesis. V. Properties of a polypeptides made in HSV-1 and HSV-2 infected cells. Virology **77**:733–749.
58. **Peterson, M. G., N. Tanese, B. F. Pugh, and R. Tjian.** 1990. Functional domains and upstream activator properties of cloned human TATA binding protein. Science **248**:1625–1630.
59. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of alpha genes of HSV: Expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters. Cell **24**:555–565.
60. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *ts*K. J. Virol. **29**:275–284.
61. **Pugh, F. B., and R. Tjian.** 1990. Mechanism of transcriptional activation by Sp1: evidence for cofactors. Cell **61**:1187–1197.
62. **Quinlan, M., and D. Knipe.** 1985. Stimulation of expression of a herpes simplex virus DNA-binding protein by two viral functions. Mol. Cell. Biol. **5**:957–963.
63. **Reinberg, D., and R. Roeder.** 1987. Factors involved in specific transcription by mammalian RNA polymerase II: purification and functional analysis of initiation factors IIB and IIE. J. Biol. Chem. **262**:3310–3321.
64. **Rivera-Gonzalez, R., A. N. Imbalzano, B. Gu, and N. A. DeLuca.** 1994. The role of ICP4 repressor activity in temporal expression of the IE-3 and latency associated promoter during HSV-1 infection. Virology **202**:550–564.
65. **Roberts, M. S., A. Boundy, P. O'Hare, M. C. Pizzorno, D. M. Ciufo, and G. S. Hayward.** 1988. Direct correlation between a negative autoregulatory response element at the cap site of the herpes simplex virus type 1 IE175 (α4) promoter and a specific binding site for the IE175 (ICP4) protein. J. Virol. **62**:4307–4320.
66. **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.
67. **Samuels, M., A. Fire, and P. A. Sharp.** 1982. Separation and characterization of factors mediating accurate transcription by RNA polymerase II. J. Biol. Chem. **257**:14419–14427.
68. **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.
69. **Shepard, A. A., and N. A. DeLuca.** 1991. Activities of heterodimers composed of DNA-binding- and transactivation-deficient subunits of the herpes simplex virus regulatory protein ICP4. J. Virol. **65**:299–307.
70. **Shepard, A. A., and N. A. DeLuca.** 1991. A second-site revertant of a defective herpes simplex virus ICP4 protein with restored regulatory activities and impaired DNA-binding properties. J. Virol. **65**:787–795.
71. **Shepard, A. A., A. N. Imbalzano, and N. A. DeLuca.** 1989. Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. J. Virol. **63**:3714–3728.
72. **Shepard, A. A., P. Tolentino, and N. A. DeLuca.** 1990. Transdominant inhibition of the herpes simplex virus transcriptional regulatory protein ICP4 by heterocomplex formation. J. Virol. **64**:3916–3926.
73. **Showalter, S. D., M. Zweig, and B. Hampar.** 1981. Monoclonal antibodies to herpes simplex virus type 1 proteins, including the immediate-early protein ICP4. Infect. Immun. **34**:684–692.
74. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57**:103–113.
75. **Smiley, J. R., D. C. Johnson, L. I. Pizer, and R. D. Everett.** 1992. The ICP4-binding sites in the herpes simplex virus type 1 glycoprotein (gD) promoter are not essential for efficient gD transcription during virus infection. J. Virol. **66**:623–631.
76. **Smith, C. A., P. Bates, R. Rivera-Gonzalez, B. Gu, and N. DeLuca.** 1993. ICP4, the major regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. J. Virol. **67**:4676–4687.
77. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. An acidic transcriptional activation domain binds directly and selectively to the TATA box factor TFIID. Nature (London) **345**:783–786.
78. **Tedder, D. G., R. D. Everett, K. W. Wilcox, P. Beard, and L. I. Pizer.** 1989. ICP4-binding sites in the promoter and coding regions of the herpes simplex virus gD gene contribute to activation of in vitro transcription by ICP4. J. Virol. **63**:2510–2520.
79. **Tjian, R., and T. Maniatis.** 1994. Transcriptional activation: a complex puzzle with few easy pieces. Cell **77**:5–8.
80. **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.
81. **Watson, R., and J. B. Clements.** 1980. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature (London) **285**:329–330.

82. **Watson, R. J., C. M. Preston, and J. B. Clements.** 1979. Separation and characterization of herpes simplex virus type 1 immediate-early mRNAs. J. Virol. **31:**42–52.
83. **Weir, J. P., and P. R. Narayanan.** 1990. Expression of the herpes simplex virus type 1 glycoprotein C gene requires sequences in the 5′-noncoding region of the gene. J. Virol. **64:**445–449.
84. **Weis, L., and D. Reinberg.** 1992. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. FASEB J. **6:**3300–3310.
85. **Zhou, Q., P. M. Lieberman, T. G. Boyer, and A. J. Berk.** 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. Genes Dev. **6:**1964–1974.
86. **Zhu, H., and R. Prywes.** 1992. Identification of a coactivator that increases activation of transcription by serum responsive factor and GAL4-VP16 in vitro. Proc. Natl. Acad. Sci. USA **89:**5291–5295.