JOURNAL OF VIROLOGY, Feb. 2002, p. 1548–1558 Vol. 76, No. 4
0022-538X/02/$04.00+0 DOI: 10.1128/JVI.76.4.1548–1558.2002

The Initiator Element in a Herpes Simplex Virus Type 1 Late-Gene Promoter Enhances Activation by ICP4, Resulting in Abundant Late-Gene Expression

Dool-Bboon Kim, Susan Zabierowski, and Neal A. DeLuca*

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

Received 9 August 2001/Accepted 5 November 2001

The start site regions of late genes of herpes simplex virus type 1 are similar to the eukaryotic initiator sequence (Inr), have been shown to affect the levels of expression, and may also play a role in transcription activation by the viral activator ICP4. A series of linker-scanning mutations spanning the start site of transcription and several downstream mutations in the true late gC promoter were analyzed in reconstituted in vitro transcription reactions with and without ICP4, as well as in the context of the viral genome during infection. The nucleotide contacts previously found to be important for Inr function were also found to be important for optimal induction by ICP4. While the Inr had a substantial effect on the accumulation of gC RNA during infection, no other sequence downstream of the TATA box to +124 had a significant effect on levels of expression during infection. Therefore, these studies suggest that TATA box and the Inr are the only *cis*-acting elements required to achieve optimal expression of gC, and that the high levels of late-gene transcription may be largely due to the induction by ICP4, functioning through the Inr element.

During productive infection by herpes simplex virus type 1 (HSV-1), approximately 80 genes are sequentially expressed in three major kinetic groups: immediate-early (IE, α), early (E, β), and late (L, γ) genes (36, 37, 46, 60, 61). Each HSV-1 gene has its own promoter with an obvious TATA box homology that is recognized by the host RNA polymerase II transcription machinery (1, 11, 89). A viral protein VP16, carried in with the virion, activates transcription of the five IE genes in a complex with two cellular proteins, Oct-1 and HCF (5, 55, 66). Once expressed, the IE gene products regulate expression of the early and late genes by a variety of mechanisms in the course of infection (75).

One of the IE proteins, ICP4 (infected-cell polypeptide 4), is essential for viral growth as it is required for optimal expression of most early and late genes (15, 19, 22, 25, 67, 68). ICP4 is a nuclear phosphoprotein that acts as a homodimer to activate or repress transcription depending upon the promoter (12, 15, 26, 54, 63, 67, 71). Genetic and biochemical studies have shown that ICP4 binds to DNA and interacts with a component(s) of general RNA polymerase II transcription machinery to activate or repress transcription (2, 8, 17, 29, 42, 64, 65, 84). ICP4 forms a tripartite complex with TFIIB and TATA binding protein (TBP) on DNA, interacts with TBP-associated factor 250 (8, 84), and promotes the formation of transcription preinitiation complexes on promoters (28).

The transcriptional regulation of HSV genes during infection is mostly determined by the structure of a given promoter. The promoters of IE and E genes contain a number of *cis*-acting elements upstream of the TATA box (89, 90). Viral and cellular proteins (Oct-1/VP16, Sp1, CTF or NF1, etc.) bind to these *cis*-acting elements to activate transcription (21, 24, 25, 45, 89). An important characteristic of late genes is that their expression is significantly influenced by viral DNA replication. Depending on their DNA replication requirement for expression, late genes are further subdivided into leaky-late ($\beta\gamma$, γ_1) and true-late (γ, γ_2) genes. Expression of true-late genes is strictly dependent on viral DNA replication (10, 34, 36, 37, 46). The mechanism underlying the requirement for DNA synthesis has yet to be determined (53, 76, 82). In contrast to IE and E gene promoters, true-late gene promoters lack upstream *cis*-acting elements, and the region downstream of the TATA box may play a more important role in regulating their transcription and temporal expression (23, 32, 35, 38, 44, 50, 70, 79, 85).

The TATA box is common to all HSV genes and plays a major role in determining the absolute level of expression. However, the sequence of the TATA box itself does not greatly affect the kinetics of expression (36, 37, 46, 60, 61). The core promoters of several late genes also possess a sequence resembling a eukaryotic initiator (Inr) element at the start site region (32, 39, 78, 86). The Inr can be found on various cellular and viral promoters (52, 92) and has a loose consensus sequence, $YYA_{+1}N(T/A)YY$ (where Y is pyrimidine and N is any nucleotide). It can direct the specific initiation of transcription by itself or synergistically in conjunction with other promoter elements (3, 20, 57, 69, 83, 92). While the activity of the late core promoter elements has been examined (32, 39, 78, 85), the specific roles of the start site region or Inr elements have not been studied in detail. In addition to the promoter, it has been reported that the leader region is also involved in the expression of late genes. A *cis*-acting element called downstream activation sequence (DAS) in the leader region of U_L38 was found to be important for U_L38 expression (31, 32).

The HSV-1 glycoprotein C sequence (gC, or U_L44) is a late gene whose promoter contains a fairly strong TATA box centered at −25 relative to the start site of transcription (35).

* Corresponding author. Mailing address: E1257 Biomedical Science Tower, Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261. Phone: (412) 648-9947. Fax: (412) 624-0298. E-mail: ndeluca@pitt.edu.

Transcription of gC is activated by ICP4, which facilitates the formation of the transcription initiation complex on the promoter (26, 27, 30, 79). Mutations in the core promoter region decrease the expression of gC (28, 85). Like other late genes, the sequence at the start site of gC transcription is similar to the consensus Inr sequences (28, 85). When the thymidine kinase gene (*tk*) start site region was replaced by that of gC, this chimeric promoter was activated by ICP4 better than wild-type (wt) *tk* promoter (28). Furthermore, while a mutation at the −1 site decreased transcription in the absence and presence of ICP4, transcription in the presence of ICP4 was affected to a greater degree (28). These results suggest that activation of the gC promoter by ICP4 may be mediated in part through the Inr. The presence of an Inr sequence in several late promoters (32, 39, 78, 85) also raises the possibility that the Inr is involved in the temporal regulation of their expression.

In this study, a series of linker-scanning (LS) mutations was introduced into the gC promoter region around the start site. The effects of the mutations on basal and ICP4-activated transcription were studied both in vitro and in vivo in the *tk* locus. Expression levels at different times after infection and the effect of DNA replication on expression were also determined. We also investigated contribution of the gC leader and start site regions at the natural gC locus during viral infection. The results suggest that (i) the gC start site contains a functional Inr, (ii) the same nucleotide contacts that are important for Inr function are also important for ICP4-activated transcription, (iii) Inr function is crucial for the accumulation of high levels of gC mRNA late after infection, and (iv) the TATA box and the Inr are the only *cis*-acting elements required for full expression of gC during viral infection.

MATERIALS AND METHODS

Viruses and cells. wt (KOS) and the ICP4-deficient n12 viruses were propagated and their titers were determined on Vero cells and on the ICP4-complementing cell line E5, respectively, as previously described (14). The gCLS/KOS, gC-gfp, gC-Inr*, gC-DPE*, gC-DAS*, and gCΔ+14/+118 viruses were grown on Vero cells. The gCLS/n12 viruses were grown on E5 cells. The n12 mutant contains a nonsense mutation that truncates ICP4 at amino acid 251 (16).

Plasmid constructions. Two parental plasmids, pgCHB (41) and pLSWT (28), were used to construct the series of LS pgCLS mutants. The LS mutations were introduced into the gC promoter region spanning −17 to +13 relative to the gC transcription initiation site by introducing new *Sac*I restriction sites using oligonucleotides. The mutagenic oligonucleotides contained the gC sequences from −22 (*Bsp*EI) to +14 (*Bgl*II) with the *Sac*I site at different positions (Fig. 1B). The oligonucleotides were inserted into pgCHB, replacing sequence −22 (*Bsp*EI) to +124 (*Bgl*II). The *Hin*dIII-*Bgl*II fragments from these plasmids were isolated and inserted into pLSWT, replacing the *tk* sequence from −500 (*Hin*dIII) to +55 (*Bgl*II). The resulting series of pgCLS plasmids contain the *tk* sequence from −500 to +2844 with the *tk* promoter sequence from −111 to +55 replaced by the gC sequence from −35 to +13. These plasmids also contain a temperature-sensitive (*ts*) mutation in the *tk* coding region (9, 40). All of the plasmids were sequenced to confirm the integrity of mutation.

The plasmids pSXgC, pUC18gC, pUC20gC, and pSXgC-gfp were used to generate the gC-gfp, gC-Inr*, gC-DPE*, and gC-DAS* viruses. pSXgC was constructed by inserting the *Sac*I-*Xba*I fragment containing the gC region (from −2057 to +1495) from pXbaE (courtesy of C. Smith, University of Pittsburgh) into pGEM-Zf11(+) (Promega). A *Pst*I-*Eco*RI fragment of pSXgC was subcloned into pUC18 to generate pUC18gC. pUC20gC was generated by subcloning the *Hin*cII-*Nru*I fragment of pUC18gC into pUC20. pUC20 is derived from pUC18 by inserting the oligonucleotides 5′-CGAGCTAGCACGCGTCGCG-3′ and 5′-AATTCGCGACGCGTGCTAGCTCGAGCT-3′ into the *Sac*I site of pUC18. To construct pSXgC-gfp, an *Ase*I-*Mlu*I fragment containing the green fluorescent protein (GFP) gene driven by the HCMV promoter from pGFPX (77) was isolated, modified with *Eco*RI linkers (New England Biolabs), and inserted into the *Eco*RI site of pSXgC. pSXgC-Inr*, pUC18gC-Inr*, and pUC20gC-Inr* have the same sequence as pSXgC, pUC18gC, and pUC20gC, respectively, except that they possess the LS8 *Sac*I mutation (Fig. 1B) at the gC start site. The *Sac*I mutation (LS8) was introduced into pUC20gC to generate pUC20gC-Inr* by PCR site-directed mutagenesis using the mutagenic primers 5′-GGCCTACCCGAGCTCCCGAGGGGG-3′ and 5′-CGCCCGCGGGAGCT CGGGTAGCCC-3′ and the M13/pUC sequencing primers (New England Biolabs). Using 50 ng of pUC20gC as a template, a PCR was carried out in the presence of 25 mM KCl, 10 mM Tris-HCl (pH 8.83), 5 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.2 mM deoxynucleoside triphosphates, 15 pmol of each primer, and 1.25 U of high-fidelity *Pwo* DNA polymerase (Boehringer Mannheim Biochemicals) in a final volume of 50 μl. The PCR ran for 25 cycles of 30 s at 94°C, 30 s at 45°C, and 45 s at 72°C (Perkin-Elmer Gene Amp PCR system 2400). The resulting PCR products were digested with *Hin*cII-*Nhe*I and substituted in place of the corresponding fragment in pUC18gC. The integrity of the mutation was confirmed by sequencing the plasmid in both directions (with an ABI Prism sequencer). Plasmids pSXgC-DPE* and pSXgC-DAS* were also constructed in the same way as pSXgC-Inr* except that different sets of mutagenic primers were used and the PCR products for subcloning were digested with *Pst*I and *Nhe*I. To construct pSXgC-DPE*, the oligonucleotide primers 5′-GAGGGCGCTGATC AAGGCCG-3′ and 5′-CGGCCTTGATCACAAGCGCCCTC-3′ were used. The primers 5′-CGCTTGGTCGGTTAACCGCATCG-3′ and 5′-CGATGCGG TTAACCGACCAAGCG-3′ were used to construct pSXgC-DAS*.

The plasmid pSXgCΔ+14/+118 was constructed to generate gCΔ+14/+118 virus. An *Eco*RI-*Pst*I fragment of pSXgC plasmid was replaced by an *Eco*RI-*Pst*I fragment of pUC18gCΔ+14/+118. The pUC18gCΔ+14/+118 was derived from pUC18gC by replacing *Bsp*EII-*Nru*I (−25 to +118) with the following synthetic oligonucleotides: 5′-CCGGAAGGGGACACGGGCTACCCTCACTACCGAG GGCG-3′ and 5′-CGCCCTCGGTAGTGAGGGTAGCCCGTGTCCCCTT-3′.

Nuclear extract and ICP4. The nuclear extract was prepared from HeLa cells as described previously (18, 30, 56, 74). ICP4 was purified by sequential column fractionation of infected Vero cell nuclei as previously described (80, 81).

In vitro transcription and primer extension. Supercoiled DNA templates (50 ng) were incubated in the presence or absence of ICP4 and nuclear extract (6 μl). The final concentrations of components in reaction buffer were 40 mM HEPES-KOH (pH 7.9), 60 mM KCl, 12% glycerol, 8.3 mM $MgCl_2$, 0.6 mM ribonucleoside triphosphates, 0.3 mM dithiothreitol, and 12 U of RNasin (Promega). The reaction was carried out in final volume of 30 μl for 1 h at 30°C and stopped by adding 70 μl of stop solution (150 mM sodium acetate [pH 5.3], 15 mM EDTA [pH 8.0], 150 μg/ml of tRNA). The reaction mixture was then phenol extracted, and the transcription products were ethanol-precipitated.

For the primer extension analysis, the RNA synthesized in vitro was annealed to 3 ng of ^{32}P-labeled primer 5′-AGGGGTACGAAGCCATACGCGCTTCTA CAAGGCGCT-3′ complementary to *tk* sequence from +90 to +125, with a buffer containing 10 mM Tris-HCl (pH 7.5), 250 mM KCl, and 1 mM EDTA (pH 8.0). Annealing was carried out in a final reaction volume of 10 μl. The annealed products were reverse transcribed by using 300 U of Moloney murine leukemia virus reverse transcriptase (Gibco/BRL) in a reaction mixture containing 50 mM Tris-HCl (pH 7.5), 75 mM KCl, 3 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM deoxynucleoside triphosphates, 12 U of RNasin, and 50 μg of actinomycin D. Reverse transcription was carried out in a final volume of 40 μl for 1 h at 42°C and stopped by addition of 60 μl of stop solution (1 M ammonium acetate, 20 mM EDTA). After phenol extraction and ethanol precipitation, the final products were analyzed on 6% denaturing polyacrylamide gels, dried, and exposed to XAR film (Kodak) for autoradiography. The radioactive signals were also quantified using a radioanalytic imaging system (Ambis Inc., San Diego, Calif.) or a PhosphorImager (Storm 840; Molecular Dynamics).

Primer extension reactions with RNA isolated from infected cells were carried out using the same conditions as above except that 6 μg of isolated total RNA was used in each reaction mixture. The same primer was used to detect the *tk* message and the primer, 5′-GTTCCCGGCAAAGCGAGACCGGGCATGAA AAC-3′, complementary to the gC region from +70 to +102 was used to detect the gC message. In the gCΔ+15/+118 experiment, an oligonucleotide spanning from +170 to +200, 5′-GGACACCCCCGCCCCGAGCCACAACAGGCTCC-3′, was used to detect gC message. As an internal control for some of the experiments, the primer 5′-CAACCCCCCGGGTAACCACGGGGTGC-3′, complementary to the ICP8 sequence from +81 to +106, was also used.

Recombinant viruses. All the gCLS/KOS mutants that carry the gCLS mutations in the *tk* locus of KOS were generated from the corresponding pgCLS plasmid by cotransfection of 10^6 Vero cells with 3 μg of KOS viral DNA and 3 μg of linearized pgCLS plasmid as previously described (14). The recombinant viruses, which have a *ts* Tk phenotype, were isolated in the presence of 120 mM

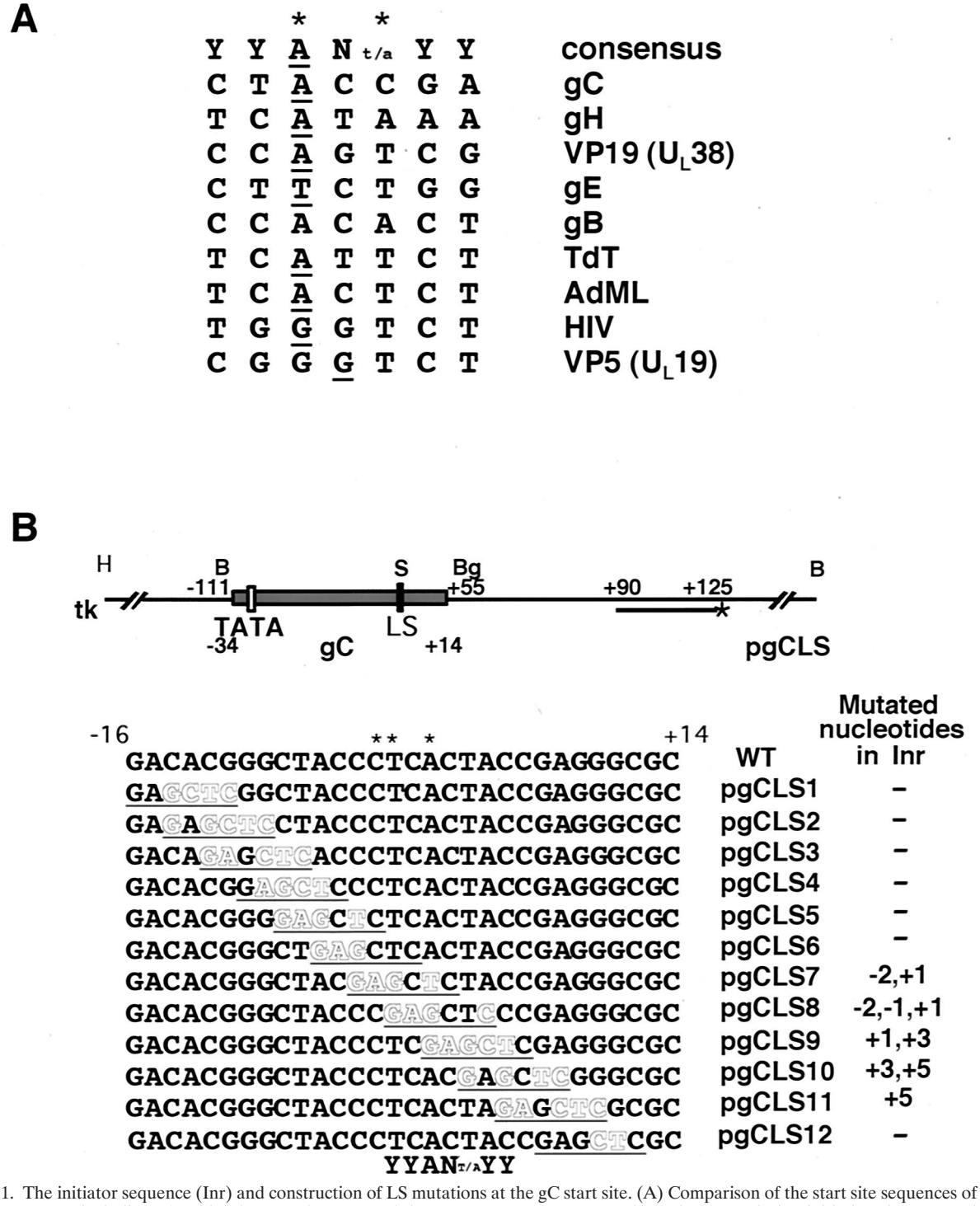

FIG. 1. The initiator sequence (Inr) and construction of LS mutations at the gC start site. (A) Comparison of the start site sequences of several HSV-1 late genes, including gC, with other Inr elements and the Inr consensus sequence. The major transcription initiation sites are underlined. The stars represent the nucleotides most deleterious to Inr activity when singly mutated (43, 57). (B) Schematic representation of the templates (pgCLS plasmids) used in the in vitro transcription assay. The pgCLS plasmids contain the gC promoter region from −34 to +14 (shaded box) replacing the *tk* promoter region from −111 to +55. The gC TATA box (white box) and the *Sac*I restriction sites specifying the LS mutations (black box) are shown. The line represents the approximately 3 kb of *tk* DNA sequences (between −500 and + 2860) flanking the gC promoter. The mutated gC region from −16 to +14 is also shown in more detail, with the new *Sac*I restriction sites underlined. The mutated nucleotides are in outlined letters, and the major start sites of gC are indicated by stars. The restriction enzyme cleavage sites shown are as follows: S, *Sac*I; H, *Hin*dIII; B, *Bam*HI; Bg, *Bgl*II.

acyclovir (Sigma) at 39.5°C. Acyclovir-resistant plaques were amplified and screened by Southern blot analysis. To confirm the transfer of LS mutations to the *tk* locus, *Sac*I-digested viral DNAs were fractionated on agarose gels, transferred to nitrocellulose membranes, and hybridized to ^{32}P-labeled pLSWT (28).

The gCLS/n12 mutants carrying both the LS mutations and a nonsense mutation in ICP4 were generated by crossing the n12 and corresponding gCLS/KOS viruses. 10^6 E5 cells were coinfected with the gCLS/KOS and n12 viruses (multiplicity of infection [MOI] = 10), and the resulting recombinants were isolated by selecting for the *ts* Tk phenotype. The LS mutation was confirmed as described above. To confirm the presence of the ICP4 nonsense mutations, viral DNA was isolated, digested with *Bam*HI and *Hpa*I, and used for Southern blot analysis using ^{32}P-labeled *Bam*HI-Y fragment from wt ICP4 as a probe (14).

The gC-gfp virus was generated by cotransfecting Vero cells with *Sac*I-linearized pSXgC-gfp and KOS DNA. Green fluorescent plaques were isolated (Nikon FXA photomicroscope) and plaque purified twice. The insertion of the GFP gene into the gC coding region was confirmed by Southern blot analysis. The gC-Inr*, gC-DPE*, gC-DAS*, and gCΔ+14/+118 viruses were generated by cotransfecting *Xba*I-linearized pSXgC-Inr*, pSXgC-DPE*, pSXgC-DAS*, and pSXgCΔ+14/+118 with the gC-gfp viral DNA into Vero cells, respectively. Nonfluorescent plaques were isolated and plaque purified. The introduction of the mutation and loss of the GFP gene were confirmed by Southern blot analysis.

Infected-cell RNA and Northern blots. To isolate total RNA, 5×10^6 Vero or E5 cells were infected at an MOI of 10. Total cellular RNA was prepared from the infected cells using Ultraspec (Biotecx Laboratories, Inc.) at the indicated time after infection. For experiments done in the presence of phosphonoacetic acid (PAA) (Lancaster Synthesis, Eastgate, England), the medium used during and after infection was supplemented with PAA (400 μg/ml).

For Northern blot analysis, the indicated amount of total RNA was separated on a 1.3% agarose gel containing formaldehyde and transferred to nitrocellulose membrane as previously described (40). The blots were initially hybridized to a probe for the *tk* gene, then stripped and hybridized to a probe for ICP8. The nick-translated probes used included the *Sac*I-*Sma*I fragment (+555 to +1217) for *tk* and the *Sal*I-*Bam*HI fragment (+585 to +1481) for ICP8. The blots were exposed for autoradiography, and the radioactive signals were quantified as described above.

RESULTS

The start site regions of many late genes, including the gC promoter, have similarities to the consensus initiator (Inr) sequence (26, 39, 78, 85, 86) (Fig. 1A) and affect transcriptional activity (26, 39, 78, 85). Mutational analyses on initiators in basal gene expression systems have shown that single changes in the +1 and +3 positions have the greatest effect on Inr activity (43, 57). We have previously found that the sequence at the start site of gC transcription affects the ability of ICP4 to activate transcription of the promoter (28). One explanation for these results is that ICP4 helps TFIID to utilize the Inr. A genetic test of this hypothesis would be to determine if the same changes that affect Inr activity also affect the magnitude of ICP4 induction.

To address this, a series of plasmids containing LS mutations around the gC start site region of the core promoter (−35 to +14) was constructed and substituted in place of the *tk* promoter in a plasmid containing the *tk* gene. The site of the LS mutation is marked by a new *Sac*I restriction site (Fig. 1B).

In vitro analysis of LS mutants. These plasmids containing the LS mutations were used directly in in vitro transcription reactions in the presence and absence of ICP4. The plasmids pgCLS1- to -5, -11, and -12 had no significant effect on transcription in vitro with or without ICP4 (data not shown; Fig. 2). The transcription of templates pgCLS6 to -10 (LS6 to LS10) and two unaffected templates, pgCLS1 and -11 (LS1 and LS11), were investigated in more detail. Consistent with previous results (28), ICP4 activated transcription from the wt gC promoter in a concentration-dependent manner (Fig. 2A). A number of the templates containing mutations around the start site resulted in reduced basal and ICP4-activated transcription. For example, relative to LS1, transcription from LS8 was reduced in the absence of ICP4, but was more impaired in the presence of ICP4, resulting in a lower fold induction (Fig. 2A). Additionally, the start site changes were evident using the template LS8.

The experiment shown in Fig. 2A was independently conducted using templates LS6 to -10 and LS11. Therefore, each template was separately compared to LS1. The fold induction by ICP4 of each template, and at each concentration of ICP4, was normalized to the fold induction of LS1 by ICP4 at the corresponding ICP4 concentration (Fig. 2B). This induction ratio is a measure of how well ICP4 can activate transcription of a given template relative to LS1. The activation of LS6 and -11 by ICP4 were not significantly affected relative to LS1. By contrast, the ability of ICP4 to activate LS7 to -10 was significantly reduced. Therefore, the most affected templates (LS7, -8, -9, and -10) with respect to ICP4 activation are those containing mutations affecting the contacts previously shown to be the most important for Inr function (43, 57).

Analysis of Inr and leader mutants in virus infection. To assess the contribution of this region to ICP4-activated transcription levels in the context of viral infection, recombinant viruses were generated containing the pgCLS1, pgCLS6 to -10, and pgCLS11 constructs in the *tk* locus of wt (strain KOS) virus. The expression of *tk* from these viruses again demonstrated that the pgCLS7, -8, -9, and -10 viruses were the most impaired, and with the possible exception of pgCLS8 all were expressed to a higher degree at early compared to late times postinfection (p.i.) (Fig. 3A). The addition of PAA did not eliminate expression from these constructs (data not shown). Similar results were obtained by analyzing the abundance of RNA by primer extension (Fig. 3B). Moreover, the primer extension products of LS7, -8, and -9 noticeably differed relative to the other constructs, indicating a change in start site. The common change in these constructs that is not present in the others is a change from the invariant A residue at position +1.

To ascertain if the reduction in expression in virus infection may in part be due to a reduction in the ability of ICP4 to activate Inr promoter mutants by ICP4, the viruses gCLS1, -6, -8, and -9 were genetically crossed with the virus n12 to put these promoters in the n12 background. Because n12 does not express functional ICP4 (16), the expression of the *tk* gene in Vero cells reflects expression in the absence of ICP4, and infection of E5 cells reflects expression in the presence of ICP4. Accordingly, Vero and E5 cells were infected with the gCLS/n12 recombinant viruses, and total RNA was isolated at 4 h p.i., which was subsequently analyzed by Northern blotting (Fig. 4A). The levels of ICP8 mRNA in the samples were determined to serve as a loading control. As expected, the levels of *tk* mRNA in gCLS/n12-infected Vero cells were very low in the absence of ICP4, but the differences in expression levels among the Inr mutants could still be discerned. In E5 cells, the most noticeably impaired promoter was the LS8 mutation. To quantify these signals, the amount of *tk* was normalized to the amount of ICP8 for a given cell type. The fact that eightfold more RNA prepared from Vero cells was used in the analysis was also considered. From these considerations it was

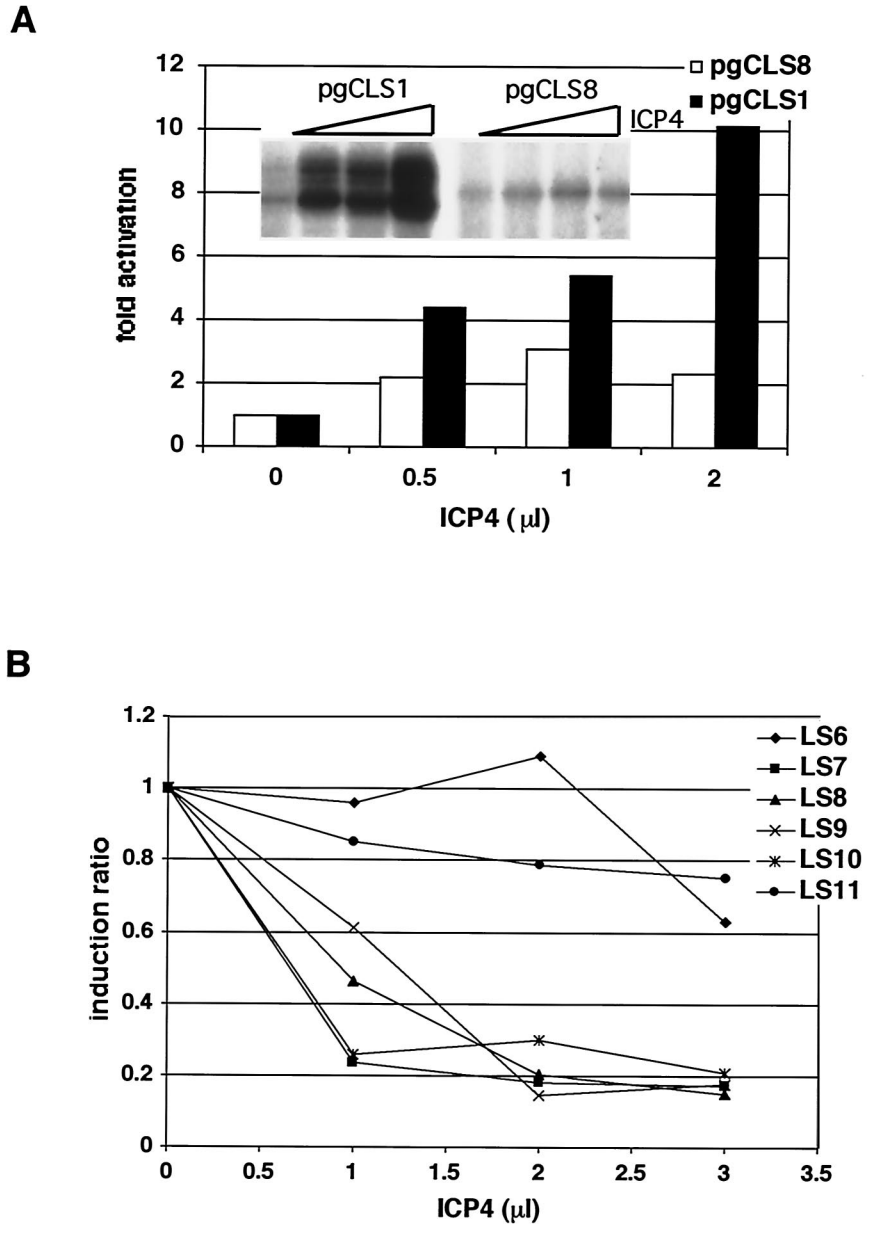

FIG. 2. In vitro transcription of the LS templates. (A) A hundred nanograms of each template (pgCLS1 or pgCLS8) was added to the in vitro transcription reactions with increasing amounts of ICP4. The primer used in this assay is complementary to the *tk* region from +90 to +125 (Fig. 1B). The final products were separated on a denaturing 6% polyacrylamide gel. For each set of reactions for each template, the first lane represents transcription in the absence of ICP4. The second, third, and fourth lanes are products in the presence of 50, 100, and 200 ng of ICP4, respectively. (B) Quantitative representation of induction ratio by ICP4 with each template. Fold induction of each template was normalized with fold induction of pgCLS1 at the different concentrations of ICP4 as defined in the text.

calculated that there was about 400-fold more *tk* RNA in gCLS1-infected cells in the presence of ICP4 than in its absence (Fig. 4B). In contrast, the abundance of *tk* RNA in gCLS8-infected cells was only 100-fold greater in the presence of ICP4 than in its absence. The amount of induction by ICP4 in gCLS9-infected cells was also reduced. Therefore, consistent with the results of the in vitro data, ICP4 was less efficient in inducing the Inr-mutated promoters.

To more directly address the contribution of the Inr and putative *cis*-acting sites downstream of the Inr to gC expression during infection, four mutants were constructed in the gC locus of KOS. One of the mutants, gC-Inr*, contains the LS8 mutation and is otherwise intact with respect to the remainder of the gC locus. The viruses gC-DPE* and gC-DAS* also have the intact gC locus but have specific sequences changed in the leader region. In gC-DPE*, the sequence between +17 and +23 was changed from GGTCGGG to TGATCAG (DPE consensus, [G/A]G[T/A]CGTG). In gC-DAS* the sequence between +22 and +28 was changed from GGAGGCC to GTTAACC. This corresponds to a change in the previously

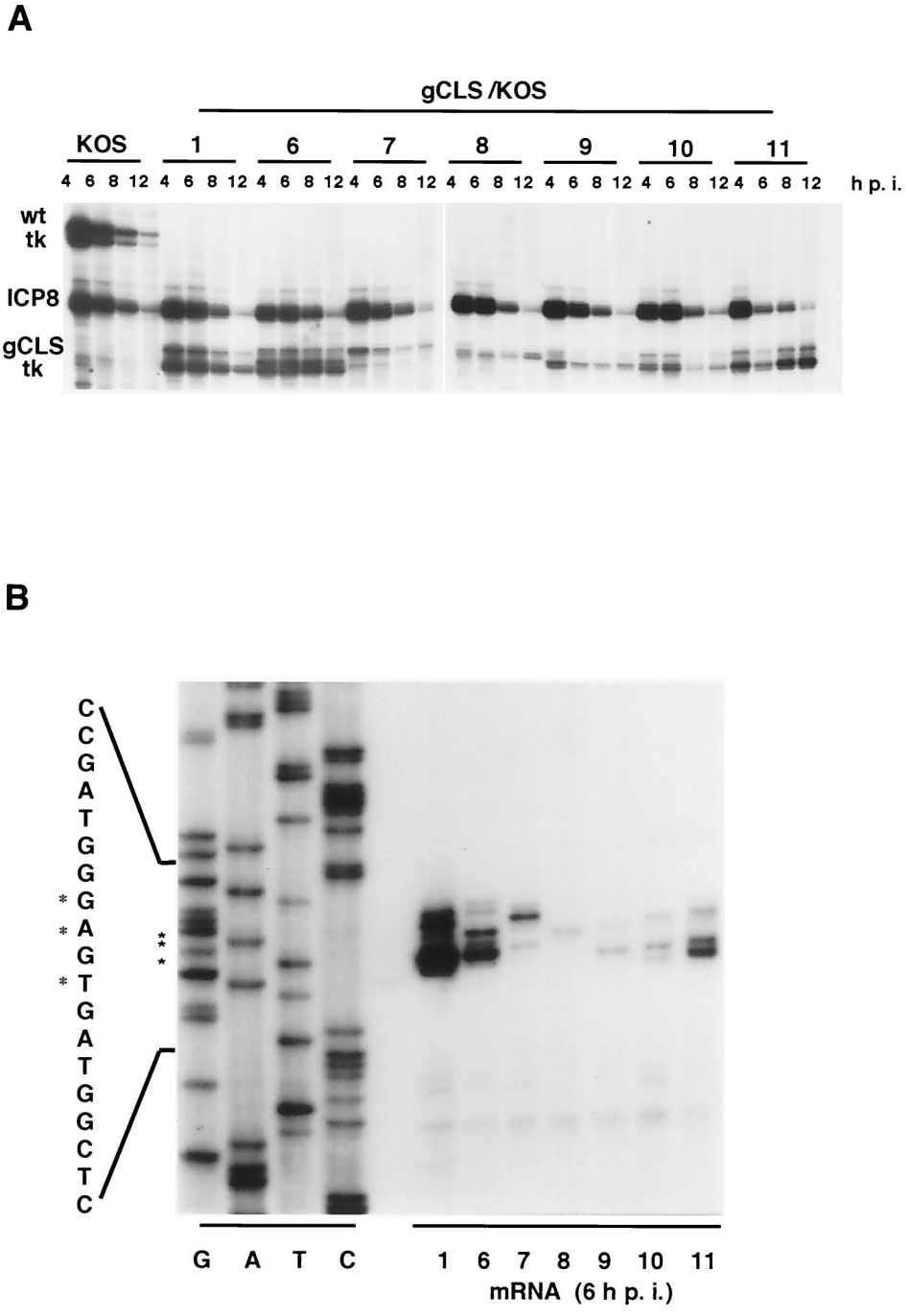

FIG. 3. Expression kinetics of the gCLS/KOS viruses and the effect of the LS mutations on the transcription start site during virus infection. (A) Primer extension analysis of RNA isolated at different times p.i. from Vero cells infected (MOI = 10) with KOS and gCLS/KOS1, -6, -7, -8, -9, -10, and -11 viruses. The reactions were carried out in the presence of both the *tk* and ICP8 primers. ICP8 was used as a control, and products were separated on a denaturing 6% polyacrylamide gel. The virus designations are shown at the top. (B) Vero cells were infected (MOI = 10) with the gCLS/KOS viruses, and total RNA was isolated at 6 h p.i. The first four lanes represent a sequencing ladder of the gC start site region using the same primer that was used in the primer extension reactions. The major start sites are marked with asterisks and stars. The gCLS/KOS virus designations are shown at the bottom.

described element, which was found to partially substitute for a DAS element (31). The final mutant, gCΔ+15/+118, deletes the region between +15 and +118 of the gC leader at the gC locus.

Vero cells were infected with the four mutants and wt virus in the presence and absence of PAA, and RNA was extracted at different times p.i. Primer extension analysis was carried out simultaneously using primers for an early gene (*tk*) and for gC (Fig. 5). A different gC primer had to be used to compare KOS and gCΔ+15/+118 (Fig. 5B), because of differences in the

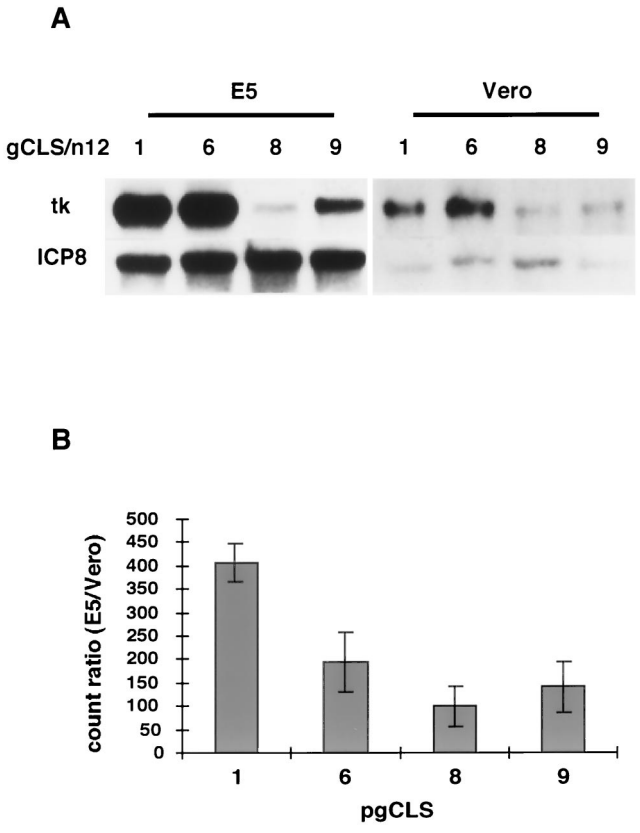

FIG. 4. Effect of the start site region on ICP4-activated transcription in virus infection. (A) Northern blot analysis of *tk* expression in Vero and E5 cells infected at a MOI of 10 PFU/cell. Total RNA (5 μg from E5 cells and 40 μg from Vero cells) from infected cells was separated on a formaldehyde gel. The blot was first probed for *tk* and then stripped and reprobed for ICP8. (B) Quantitative representation of results shown in panel A. The signals from the blot were quantified and normalized to the ICP8 signal, and ratios of fold activation in the presence (E5) and absence (Vero) of ICO_4 were calculated. Each value represents the mean and standard deviation (error bars) from three independent experiments.

leader regions. The results of the experiment using the gC-Inr* mutant (Fig. 5A and 6) reveal that (i) the start site of gC transcription from gC-Inr* was affected as anticipated from previous results, (ii) the kinetics of gC transcription was not significantly affected by the gC-Inr* mutation, and (iii) the levels of gC transcription were dramatically reduced as a function of the gC-Inr* mutation. Figure 6 shows a quantitative representation of the data in Fig. 5A for the internal controls (*tk*) and the gC genes expressed in KOS and gC-Inr*-infected cells. It is clear that the Inr sequence of gC has a profound effect on the accumulation of this abundantly transcribed late message.

In contrast to the effects seen with the mutant initiator, the effects seen in gC-DPE*, gC-DAS*, and gCΔ+15/+118-infected cells were relatively minor, if present at all. While the level of gC RNA in gC-DAS*-infected cells may be slightly reduced relative to that in KOS-infected cells, this reduction is not significant relative to that seen in gC-Inr*-infected cells (Fig. 5A and 6). Furthermore, deletion of the entire leader between +15 and +118 had no significant effect on the accumulation of gC mRNA (Fig. 5B). Lastly, the accumulation of gC mRNA in all the infections was completely dependent on DNA synthesis (data not shown and Fig. 5B). Thus, it appears that the sequence having the greatest effect on the accumulation of gC mRNA is the Inr region and that downstream sequences to +118 have little effect on gC gene expression.

DISCUSSION

gC start site region as an Inr and facilitator of ICP4 activation. Mutagenesis of the terminal deoxynucleotidyl transferase (TdT) and adenovirus major late (AdML) Inrs has shown that Inr activity cannot be abrogated by single nucleotide changes. Single mutations at positions +1 and +3 are the most consequential for Inr function, and the presence of a pyrimidine at position −1 is also important both in cells and in vitro (43, 57). Simultaneous mutation of these sites has more substantial effects. The gC Inr spans from −4 to +5 with one nucleotide mismatch (+4, Y to U) compared to the Inr consensus sequence (Fig. 1A). The LS mutant most affected (LS8) for transcription level, both in vitro and in virus infection, has mutations at −2, −1, and +1. The second-most-affected LS mutant (LS9) has mutations at positions +1 and +3 (Fig. 1B). Additionally, LS mutants 7 and 10 were also substantially affected in all the assays, and these have mutations at −2 and +1 and +3 and +5, respectively (Fig. 1B and 2). Therefore, the residues previously shown to be important for Inr function in other systems correlate with the residues important for gC expression in the virus and in vitro. Furthermore, the results also suggest that there are no other *cis*-acting sites affecting expression during infection in this region (−17 to +13).

Although mutations in the Inr sequence resulted in decreased basal transcription, ICP4-activated transcription was significantly more affected (Fig. 2 and 3). ICP4 still activated the LS8 template, albeit less efficiently than the templates with an intact Inr. The activation of the promoter by ICP4 in the absence of Inr function may reflect the ability of ICP4 to promote transcription preinitiation complexes (PICs) on the promoter by facilitating TFIID binding to the TATA box (27). The additional presence of a functional Inr and an activity of ICP4 may enhance PIC formation or trigger some postrecruitment step that results in full activation. While further studies are necessary to distinguish between these two possibilities, some reported data are relevant to this question.

TFIID binds to the TATA box via the TBP (for a review, see reference 33). However, TBP-associated factors (TAFs) in TFIID are involved in recognizing the Inr (6, 48, 58, 72, 87, 91, 93). In the *Drosophila melanogaster* system, a complex consisting of TBP-TAF150-TAF250 is sufficient for Inr activity. Evidence suggests that TAF250 recognizes the Inr and that dTAF150 (or the mammalian homologue CIF150) is required to support Inr activity (47, 49, 59, 88,). TAFs also bridge transactivators and general transcription factors by protein-protein interactions, thus affecting initiation complex formation at the transcription start site, resulting in activation (57, 86). ICP4 requires TAFs for activation (28), and it interacts with TFIID in solution through TAF250 (8). The ability of ICP4 to interact with TAF250 is mediated by the carboxy-terminal 524 amino acids (8). Mutants that lack this region of

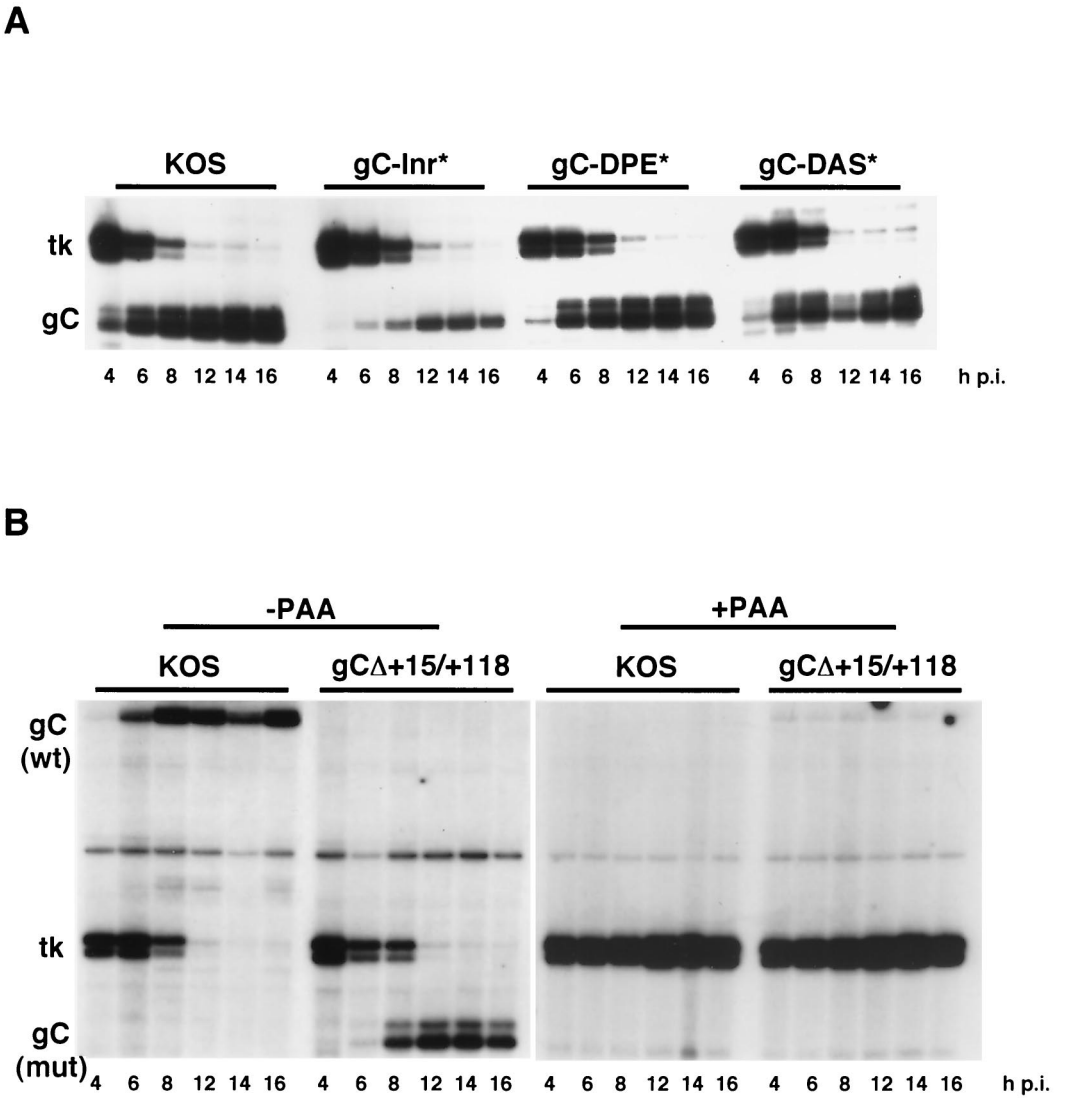

FIG. 5. Effect of the initiator and the leader region mutations at the gC locus. (A) Primer extension analysis of RNA isolated from Vero cells infected (MOI = 10) with KOS, gC-Inr*, gC-DPE*, and gC-DAS* mutants. RNA isolated from the indicated viruses at different time points was used in primer extension analysis in the presence of both the *tk* and gC primers. To detect the gC message, an oligonucleotide spanning the sequence from +70 to +102 was used as the primer. To detect *tk* message, the same primer described in Fig. 2 was used. (B) Primer extension analysis of RNA isolated from Vero cells infected with gCΔ+15/+118 mutant in the presence (+) and absence (−) of PAA. To detect gC message in this experiment, an oligonucleotide spanning the sequence from +170 to +200 was used.

ICP4 are significantly impaired in activation (80). The residual activation is a function of the amino-terminal 774 amino acids of ICP4 (16). This fragment of ICP4 is sufficient to promote PIC formation on the gC promoter (27).

Therefore, the amino terminus of ICP4 promotes PIC formation independent of an interaction with TAF250. The interaction with TAF250 as a consequence of some function in the carboxy-terminal region may allow TFIID to more efficiently function through the Inr. Thus, the interaction between ICP4 and TAF250 may enable the recognition of the Inr by TAF250 in a manner analogous to dTAF150 or CIF150, which may additionally involve a conformational change in the complex. Alternatively, the additive effects of additional multiple protein-protein and protein-DNA interactions as a function of ICP4, TAF250, the TATA box, and the Inr at the start site might drive the formation of an activated transcription complex. At present, we favor the former hypothesis. The latter hypothesis, while seemingly inconsistent with the PIC formation studies (27), warrants further testing considering the limitations of such studies.

Role of the Inr and leader region during viral infection. The two important features of late genes that are in part functions of the activities of their promoters are their requirement for DNA synthesis and their abundant expression late after infection. The former question is a complex problem that has yet to be fully explained. The latter question is interesting in that the function of true late promoters is not apparently augmented by upstream activators, as is the case for IE and E

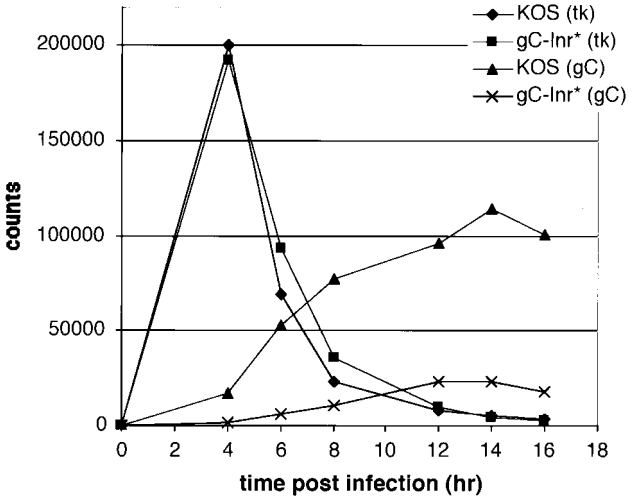

FIG. 6. Effect of Inr mutation in vivo. Transcription levels of gC from the wt and gC-Inr* mutant from Fig. 5A were quantitatively compared using a phosphorimager.

genes. Therefore, the biological significance of the function of the Inr present in many late-gene promoters is an important question.

Several mechanisms have been suggested to explain late-gene expression during viral replication. These include differences in promoter strength and architecture, the presence of inhibitors, the conformational change of the genome after viral replication, and differences in polyadenylation sites (53, 62, 76, 82). It has been suggested that *cis*-acting elements in late genes play a role in temporal regulation. Furthermore, it was previously reported that the gC region from −35 to +124 in the *tk* locus results in late-expression kinetics and required DNA replication for expression to occur (35). In this study we found that the core gC promoter (−35 to +14) inserted in the *tk* locus did not result in true late-gene expression. However, the presence of a wt Inr in these constructs resulted in prolonged gene expression. Mutation of the Inr generally reduced expression and resulted in a gene expression profile more closely resembling that of early genes. Therefore, while the Inr does not mediate the requirement for DNA synthesis, it may contribute to true late-gene expression by allowing higher levels of expression late after infection.

When the core promoter was examined in the natural gC locus, expression was maintained late after infection. However, in contrast to the results obtained with the core promoter in the *tk* locus, expression was completely dependent on DNA synthesis. Moreover, the level of expression was very similar to that observed with the intact gC locus. Therefore, the region between +15 and +124 does not greatly affect the expression of the gC gene at the natural gC locus. This suggests that the context in which the promoter is situated and subsequently analyzed is very important. It may be that the requirement for DNA synthesis is not mediated by a definable set of *cis*-acting elements but is rather a function of both the promoter structure and the context of the promoter in the viral genome.

When the initiator element at the natural gC locus was mutated, expression kinetics were qualitatively unaltered relative to wt virus. However, expression was dramatically reduced. By comparison, the expression from the viruses containing the leader region deletion, or mutations in the potential DAS and DPE elements, were relatively unaffected compared to wt virus. In terms of sequence and location, DAS is highly similar to a downstream promoter element, DPE. DPE is a core promoter element mostly found in *Drosophila* TATA-less promoters (4, 7), and frequently functions in a combination with the Inr (3, 4). Having these additional core promoter elements (Inr and DPE) along with TATA box would be advantageous for expression of late genes. Our studies argue that while these elements may function in some late genes, they apparently contribute little to gC expression.

Our results suggest that true late-gene promoters greatly rely on the function of the Inr. Importantly, it is not simply the basal activity of the Inr that is important for high-level late-gene expression, but rather the ability of ICP4 to more efficiently activate transcription with a functional Inr. Late-gene expression and viral DNA replication occur in globular nuclear structures called replication compartments (13, 73), where specific viral proteins, including ICP4 (51), have been shown to localize. Thus, the compartmentalization of the viral genome and ICP4 and the ability of ICP4 to activate transcription more efficiently as a function of interactions with TFIID in the presence of an Inr, may collectively result in the high levels of late mRNA during infection. These high levels of expression are necessary to produce the large quantities of structural proteins required to assemble thousands of progeny virions per cell.

ACKNOWLEDGMENTS

We thank Michael Carrozza and Lorna Samaniego for helpful discussions and comments on the manuscript.

This work was supported by NIH grant AI30612.

REFERENCES

1. **Alwine, J. C., W. L. Steinhart, and C. W. Hill.** 1974. Transcription of herpes simplex type 1 DNA in nuclei isolated from infected HEp-2 and KB cells. Virology **60:**302–307.
2. **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.
3. **Burke, T. W., and J. T. Kadonaga.** 1997. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. Genes Dev. **11:**3020–3031.
4. **Burke, T. W., and J. T. Kadonaga.** 1996. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. **10:**711–724.
5. **Campbell, M. E., J. W. Palfreyman, and C. M. Preston.** 1984. Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. **180:**1–19.
6. **Carcamo, J., L. Buckbinder, and D. Reinberg.** 1991. The initiator directs the assembly of a transcription factor IID-dependent transcription complex. Proc. Natl. Acad. Sci. USA **88:**8052–8056.
7. **Carcamo, J., E. Maldonado, P. Cortes, M. H. Ahn, I. Ha, Y. Kasai, J. Flint, and D. Reinberg.** 1990. A TATA-like sequence located downstream of the transcription initiation site is required for expression of an RNA polymerase II transcribed gene. Genes Dev. **4:**1611–1622.
8. **Carrozza, M. J., and N. A. DeLuca.** 1996. Interaction of the viral activator protein ICP4 with TFIID through TAF250. Mol. Cell. Biol. **16:**3085–3093.
9. **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.
10. **Conley, A. J., D. M. Knipe, P. C. Jones, and B. Roizman.** 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of gamma polypeptides. J. Virol. **37:**191–206.
11. **Costanzo, F., G. Campadelli-Fiume, L. Foa-Tomasi, and E. Cassai.** 1977. Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. J. Virol. **21:**996–1001.

12. **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.
13. **de Bruyn Kops, A., and D. M. Knipe.** 1988. Formation of DNA replication structures in herpes virus-infected cells requires a viral DNA binding protein. Cell **55:**857–868.
14. **DeLuca, N. A., A. M. 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.
15. **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:**1997–2208.
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 initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res.**11:**1475–1489.
19. **Dixon, R. A., 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. **Emami, K. H., A. Jain, and S. T. Smale.** 1997. Mechanism of synergy between TATA and initiator: synergistic binding of TFIID following a putative TFIIA-induced isomerization. Genes Dev. **11:**3007–3019.
21. **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.
22. **Everett, R. D.** 1984. Trans activation of transcription by herpes virus products: requirement for two HSV-1 immediate-early polypeptides for maximum activity. EMBO J. **3:**3135–3141.
23. **Flanagan, W. M., A. G. Papavassiliou, M. Rice, L. B. Hecht, S. Silverstein, and E. K. Wagner.** 1991. Analysis of the herpes simplex virus type 1 promoter controlling the expression of UL38, a true late gene involved in capsid assembly. J. Virol. **65:**769–786.
24. **Gelman, I. H., and S. Silverstein.** 1987. Dissection of immediate-early gene promoters from herpes simplex virus: sequences that respond to the virus transcriptional activators. J. Virol. **61:**3167–3172.
25. **Gelman, I. H., and S. Silverstein.** 1985. Identification of immediate early genes from herpes simplex virus that transactivate the virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **82:**5265–5269.
26. **Godowski, P. J., and D. M. Knipe.** 1986. Transcriptional control of herpesvirus gene expression: gene functions required for positive and negative regulation. Proc. Natl. Acad. Sci. USA **83:**256–260.
27. **Grondin, B., and N. DeLuca.** 2000. Herpes simplex virus type 1 ICP4 promotes transcription preinitiation complex formation by enhancing the binding of TFIID to DNA. J. Virol. **74:**11504–11510.
28. **Gu, B., and N. DeLuca.** 1994. Requirements for activation of the herpes simplex virus glycoprotein C promoter in vitro by the viral regulatory protein ICP4. J. Virol. **68:**7953–7965.
29. **Gu, B., R. Kuddus, and N. A. DeLuca.** 1995. Repression of activator-mediated transcription by herpes simplex virus ICP4 via a mechanism involving interactions with the basal transcription factors TATA-binding protein and TFIIB. Mol. Cell. Biol. **15:**3618–3626.
30. **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.
31. **Guzowski, J. F., J. Singh, and E. K. Wagner.** 1994. Transcriptional activation of the herpes simplex virus type 1 UL38 promoter conferred by the *cis*-acting downstream activation sequence is mediated by a cellular transcription factor. J. Virol. **68:**7774–7789.
32. **Guzowski, J. F., and E. K. Wagner.** 1993. Mutational analysis of the herpes simplex virus type 1 strict late UL38 promoter/leader reveals two regions critical in transcriptional regulation. J. Virol. **67:**5098–5108.
33. **Hernandez, N.** 1993. TBP, a universal eukaryotic transcription factor? Genes Dev. **7:**1291–1308.
34. **Holland, L. E., K. P. Anderson, C. Shipman, Jr., and E. K. Wagner.** 1980. Viral DNA synthesis is required for the efficient expression of specific herpes simplex virus type 1 mRNA species. Virology **101:**10–24.
35. **Homa, F. L., T. M. Otal, J. C. Glorioso, and M. Levine.** 1986. Transcriptional control signals of a herpes simplex virus type 1 late ($\gamma 2$) gene lie within bases -34 to $+124$ relative to the 5' terminus of the mRNA. Mol. Cell. Biol. **6:**3652–3666.
36. **Honess, R. W., and B. Roizman.** 1974. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. **14:**8–19.
37. **Honess, R. W., and B. Roizman.** 1975. Regulation of herpesvirus macromolecular synthesis: sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA **72:**1276–1280.
38. **Huang, C. J., S. A. Goodart, M. K. Rice, J. F. Guzowski, and E. K. Wagner.** 1993. Mutational analysis of sequences downstream of the TATA box of the herpes simplex virus type 1 major capsid protein (VP5/UL19) promoter. J. Virol. **67:**5109–5116.
39. **Huang, C. J., and E. K. Wagner.** 1994. The herpes simplex virus type 1 major capsid protein (VP5-UL19) promoter contains two *cis*-acting elements influencing late expression. J. Virol. **68:**5738–5747.
40. **Imbalzano, A. N., D. M. 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.
41. **Imbalzano, A. N., and N. A. 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.
42. **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.
43. **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.
44. **Johnson, P. A., and R. D. Everett.** 1986. The control of herpes simplex virus type-1 late gene transcription: a 'TATA-box'/cap site region is sufficient for fully efficient regulated activity. Nucleic Acids Res.**14:**8247–8264.
45. **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.
46. **Jones, P. C., and B. Roizman.** 1979. Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. J. Virol. **31:**299–314.
47. **Kaufmann, J., K. Ahrens, R. Koop, S. T. Smale, and R. Muller.** 1998. CIF150, a human cofactor for transcription factor IID-dependent initiator function. Mol. Cell. Biol. **18:**233–239.
48. **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.
49. **Kaufmann, J., C. P. Verrijzer, J. Shao, and S. T. Smale.** 1996. CIF, an essential cofactor for TFIID-dependent initiator function. Genes Dev. **10:**873–886.
50. **Kibler, P. K., J. Duncan, B. D. Keith, T. Hupel, and J. R. Smiley.** 1991. Regulation of herpes simplex virus true late-gene expression: sequences downstream from the US11 TATA box inhibit expression from an unreplicated template. J. Virol. **65:**6749–6760.
51. **Knipe, D. M., D. Senechek, S. A. Rice, and J. L. Smith.** 1987. Stages in the nuclear association of the herpes simplex virus transcriptional activator protein ICP4. J. Virol. **61:**276–284.
52. **Kollmar, R., and P. J. Farnham.** 1993. Site-specific initiation of transcription by RNA polymerase II. Proc. Soc. Exp. Biol. Med. **203:**127–139.
53. **Koop, K. E., 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.
54. **Kristie, T. M., and B. Roizman.** 1986. DNA-binding site of major regulatory protein alpha 4 specifically associated with promoter-regulatory domains of alpha genes of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83:**4700–4704.
55. **Kristie, T. M., and P. A. Sharp.** 1993. Purification of the cellular C1 factor required for the stable recognition of the Oct-1 homeodomain by the herpes simplex virus alpha-trans-induction factor (VP16). J. Biol. Chem. **268:**6525–6534.
56. **Lin, Y. S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64:**971–981.
57. **Lo, K., and S. T. Smale.** 1996. Generality of a functional initiator consensus sequence. Gene **182:**13–22.
58. **Martinez, E., C. M. Chiang, H. Ge, 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.
59. **Martinez, E., H. Ge, Y. Tao, C. X. Yuan, V. Palhan, and R. G. Roeder.** 1998. Novel cofactors and TFIIA mediate functional core promoter selectivity by the human TAFII150-containing TFIID complex. Mol. Cell. Biol. **18:**6571–6583.
60. **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.
61. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res.**14:**1727–1745.
62. **McGregor, F., A. Phelan, J. Dunlop, and J. B. Clements.** 1996. Regulation of

herpes simplex virus poly(A) site usage and the action of immediate-early protein IE63 in the early-late switch. J. Virol. **70:**1931–1940.
63. **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.
64. **Michael, N., and B. Roizman.** 1989. Binding of the herpes simplex virus major regulatory protein to viral DNA. Proc. Natl. Acad. Sci. USA **86:**9808–9812.
65. **Michael, N., D. Spector, P. Mavromara-Nazos, T. M. Kristie, and B. Roizman.** 1988. The DNA-binding properties of the major regulatory protein alpha 4 of herpes simplex viruses. Science **239:**1531–1534.
66. **O'Hare, P., C. R. Goding, and A. Haigh.** 1988. Direct combinatorial interaction between a herpes simplex virus regulatory protein and a cellular octamer-binding factor mediates specific induction of virus immediate-early gene expression. EMBO J. **7:**4231–4238.
67. **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.
68. **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.
69. **O'Shea-Greenfield, A., and S. T. Smale.** 1992. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. J. Biol. Chem. **267:**1391–1402.
70. **Pederson, N. E., S. Person, and F. L. Homa.** 1992. Analysis of the gB promoter of herpes simplex virus type 1: high-level expression requires both an 89-base-pair promoter fragment and a nontranslated leader sequence. J. Virol. **66:**6226–6232.
71. **Pereira, L., M. H. Wolff, M. Fenwick, and B. Roizman.** 1977. Regulation of herpesvirus macromolecular synthesis. V. Properties of alpha polypeptides made in HSV-1 and HSV-2 infected cells. Virology **77:**733–749.
72. **Pugh, B. F., and R. Tjian.** 1991. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. Genes Dev. **5:**1935–1945.
73. **Quinlan, M. P., L. B. Chen, and D. M. Knipe.** 1984. The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. Cell **36:**857–868.
74. **Reinberg, D., and R. G. Roeder.** 1987. Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. J. Biol. Chem. **262:**3331–3337.
75. **Roizman, B., M. Kozak, R. W. Honess, and G. Hayward.** 1975. Regulation of herpesvirus macromolecular synthesis: evidence for multilevel regulation of herpes simplex 1 RNA and protein synthesis. Cold Spring Harbor Symp. Quant. Biol. **39:**687–701.
76. **Romanelli, M. G., P. Mavromara-Nazos, D. Spector, and B. Roizman.** 1992. Mutational analysis of the ICP4 binding sites in the 5′ transcribed noncoding domains of the herpes simplex virus 1 UL 49.5 gamma 2 gene. J. Virol. **66:**4855–4863.
77. **Samaniego, L. A., L. Neiderhiser, and N. A. DeLuca.** 1998. Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins. J. Virol. **72:**3307–3320.
78. **Sethna, M., and J. P. Weir.** 1993. Mutational analysis of the herpes simplex virus type 1 glycoprotein E promoter. Virology **196:**532–540.
79. **Shapira, M., F. L. Homa, J. C. Glorioso, and M. Levine.** 1987. Regulation of the herpes simplex virus type 1 late (gamma 2) glycoprotein C gene: sequences between base pairs −34 to +29 control transient expression and responsiveness to transactivation by the products of the immediate early (alpha) 4 and 0 genes. Nucleic Acids Res.**15:**3097–3111.
80. **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.
81. **Shepard, A. A., P. Tolentino, and N. A. DeLuca.** 1990. *trans*-dominant inhibition of herpes simplex virus transcriptional regulatory protein ICP4 by heterodimer formation. J. Virol. **64:**3916–3926.
82. **Silver, S., and B. Roizman.** 1985. γ2-Thymidine kinase chimeras are identically transcribed but regulated as γ2 genes in herpes simplex virus genomes and as β genes in cell genomes. Mol. Cell. Biol. **5:**518–528.
83. **Smale, S. T.** 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta **1351:**73–88.
84. **Smith, C. A., P. Bates, R. Rivera-Gonzalez, B. Gu, and N. A. DeLuca.** 1993. ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIB. J. Virol. **67:**4676–4687.
85. **Steffy, K. R., and J. P. Weir.** 1991. Mutational analysis of two herpes simplex virus type 1 late promoters. J. Virol. **65:**6454–6460.
86. **Steffy, K. R., and J. P. Weir.** 1991. Upstream promoter elements of the herpes simplex virus type 1 glycoprotein H gene. J. Virol. **65:**972–975.
87. **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.
88. **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.
89. **Wagner, E. K., J. F. Guzowski, and J. Singh.** 1995. Transcription of the herpes simplex virus genome during productive and latent infection. Prog. Nucleic Acid Res. Mol. Biol. **51:**123–165.
90. **Wagner, E. K., M. D. Petroski, N. T. Pande, P. T. Lieu, and M. Rice.** 1998. Analysis of factors influencing kinetics of herpes simplex virus transcription utilizing recombinant virus. Methods **16:**105–116.
91. **Wang, J. C., and M. W. Van Dyke.** 1993. Initiator sequences direct downstream promoter binding by human transcription factor IID. Biochim. Biophys. Acta **1216:**73–80.
92. **Weis, L., and D. Reinberg.** 1992. Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. FASEB J. **6:**3300–3309.
93. **Zenzie-Gregory, B., A. O'Shea-Greenfield, and S. T. Smale.** 1992. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J. Biol. Chem. **267:**2823–2830.