JOURNAL OF VIROLOGY, May 1992, p. 2973–2981
0022-538X/92/052973-09$02.00/0

Vol. 66, No. 5

Effect of Genomic Location on Expression of β-Galactosidase mRNA Controlled by the Herpes Simplex Virus Type 1 U_L38 Promoter

SHERYL A. GOODART, JOHN F. GUZOWSKI, MARCIA K. RICE, AND EDWARD K. WAGNER*
Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, California 92717

Received 14 November 1991/Accepted 5 February 1992

To examine the effect of genomic location on the details of expression of selected herpes simplex virus promoters, we have constructed recombination vectors for placing such promoters controlling the β-galactosidase reporter gene into two regions of the viral genome lacking any nearby promoter or regulatory elements. The first vector generates the promoter–β-galactosidase reporter gene inverted within the locus of the gC (U_L44) translational reading frame; the second replaces the LAT promoter and the first 600 bases of the primary transcript in both copies of the R_L region. These locations were chosen to obviate any possible influence of upstream but noncontiguous heterologous or homologous DNA sequence elements upon promoter activity. When the reporter gene controlled by the strict late (γ) U_L38 promoter was placed in the gC location, it was significantly less active than in its normal location; in contrast, promoter activity was comparable to wild-type values when the promoter was recombined into the R_L region. The low level of activity in the gC location could be partially alleviated by the incorporation of additional DNA sequences upstream of the U_L38 promoter. Despite the effect of genomic location upon the level of expression, the kinetics of expression in either location mirrors the wild-type U_L38 strict late kinetics of expression. Finally, we used deletional analysis to demonstrate that no more than 29 bases of DNA sequence 5′ of the mRNA cap site are required for promoter activity in either location; this result is consistent with earlier results of transient-expression assays and indicates that the U_L38 promoter shares general features with other strict late (γ) herpes simplex virus promoters.

Transcripts encoded by herpes simplex virus type 1 (HSV-1) are expressed during the lytic replication cycle in a sequential cascade pattern (14). Regulation of this expression pattern occurs primarily at the level of transcription (21, 26, 29, 31). Transcripts encoding the five immediate-early (α) genes are the first to be expressed; these gene products, especially ICP4 (α4), are required for expression of subsequent classes of HSV-1 genes during lytic expression. Early or delayed-early (β) genes are the second group of genes to be expressed. Generally the polypeptide products of these genes are involved in replication of viral DNA; their expression tends to reach maximal levels around the peak of DNA replication and then decrease (31). Late transcripts are maximally expressed only after the onset of viral DNA replication. They can be grouped into leaky-late (βγ or γ_1) genes and true or strict late (γ or γ_2) genes; the difference between these two subclasses is in the stringency of the requirement for DNA replication. Transcripts encoding βγ genes are expressed at low but readily detectable levels before DNA synthesis, whereas γ transcripts evidence a strict requirement for DNA replication for any significant expression (10).

The promoters controlling the γ or strict late genes have been the subject of some investigation, especially those controlling the gH (U_L22), U_L38, gC (U_L44), and U_S11 genes (11, 13, 16, 17, 22, 24). Although these promoters do not evidence a requirement for viral DNA replication for expression in transient-expression assays, the extent of the promoters delineated generally agrees with those determined by generation of recombinant viruses containing modified promoter sequences. Current data indicate that γ promoters are quite limited in length, extending from a recognizable TATA homology 28 to 30 bases 5′ of the mRNA cap site and including sequences near the transcript initiation site extending to 10 to 20 bases 3′ of this point.

The limited number of DNA sequence elements responsible for the control of γ transcription is of interest because of its potential utility in saturation mutagenesis studies and as probes for the identification of transcription factors important in late gene expression. As noted, to date the most clearly recognizable promoter element found in these promoters is the TATA homology. Since this element is such a notable feature of all γ promoters studied, it is reasonable to speculate that a protein factor(s) which recognizes this element might play a primary role in the expression of γ genes. Further, since γ genes require viral DNA replication for expression, it is plausible to speculate that such a factor might be altered in infection to enable it to better recognize γ promoters.

Despite the ubiquity of γ promoter TATA boxes, there is some controversy concerning the actual sequence requirements for function. Johnson and Everett reported that in transient-expression assays they could change the early gD (U_S6) promoter into an operationally strict late promoter on a plasmid containing the HSV ori_s by truncating it at −33, leaving the TATA element as the only recognizable promoter element (16). More recently, Steffy and Weir reported that three TATA elements derived from the α4, β thymidine kinase (TK; U_L23), and γ gC (U_L44) promoters could substitute for the γ gH (U_L22) TATA element in a recombinant virus, although the levels of expression differed significantly depending upon the exact element used (24). These results argue against a strict sequence requirement for a

* Corresponding author.

functional γ TATA box. However, Homa et al. concluded that the TK (U_L23) promoter-leader (−37 through +57) possessing the intact TATA box (ATATTAA) could not substitute for the gC promoter-leader (−146 through +124) in recombinant virus (11, 12). No gC mRNA was expressed from the TK promoter-leader–gC fusion until the TK TATA element was mutated to GTATAAA, which more closely resembles the gC sequence, TATAAATT. This mutated TATA box generated 10% of wild-type (wt) levels of gC mRNA.

Although results of studies with recombinant viruses bearing mutated γ promoter regions generally agree with results of transient-expression assays, these experiments have been carried out at the normal location of the promoter in question (11, 12, 24). This allows the possibility that noncontiguous DNA sequence elements further upstream of the TATA box will influence mRNA expression. Further, such experiments require that the promoter in question control expression of a gene which is nonessential for viral replication in cell culture. We have developed two recombination vectors which allow the insertion of any HSV promoter controlling the β-galactosidase indicator gene into two regions of the HSV-1 genome, one within the location of the nonessential gC (U_L44) gene and the other replacing a portion of the latency-associated transcription unit and promoter in the long repeat (R_L) region. These two locations of recombination will obviate the influence of any specific promoter elements occurring upstream of the "core" promoter and allow a rough estimate of the effect of genomic position on expression. In this system, the wt transcription unit and control sequences of the gene whose promoter is being studied are not altered and the expression of the wt mRNA provides an absolute internal standard for measurement.

We have chosen the promoter controlling the U_L38 gene as a useful model for the further analysis of γ promoters. This promoter is within a 280-bp region of DNA which also contains the promoter for the β U_L37 mRNA transcribed in the opposite direction (2, 31). These two promoters are completely separable in transient-expression assays. Sequences extending from −29 to +10 relative to the cap site provide essentially full expression upon viral superinfection, and DNA extending from −48 to +99 bases is sufficient for the expression of a marker gene (β-galactosidase) in a recombinant virus. Three T+A-rich regions occur in this latter region: AAAT at −38, TTTAAA at −29, and TATA at −17. Further, antibody-mediated DNA supershift assays indicate that the α4 protein interacts with DNA near the cap site; and, finally, the target for this interaction has sequence homology with promoter elements of the γ42 (U_L49) gene (6).

In the current report, we describe the details of expression of U_L38-controlled β-galactosidase in these recombinant virus systems. With the U_L38 promoter, we found that genomic position strongly influences the level but not kinetics of expression. Deletional analysis shows that in the virus, the sequence TTTAAA between −29 and −23 of the U_L38 mRNA cap site is absolutely required for transcription, but levels of expression can be readily increased by the addition of further sequence elements farther upstream. Thus, we can clearly differentiate sequence elements required for the kinetics of expression from those important in influencing levels of expression. The unusual nature of the U_L38 TATA homology supports the major body of experimental data indicating that a specific sequence of the γ TATA box is of minor importance in mediating transcription.

MATERIALS AND METHODS

Cells and virus. HSV-1 (17*syn*$^+$) was the parental strain used to generate the recombinant viruses described below. Parental and recombinant virus infection of rabbit skin fibroblasts at a multiplicity of infection of 5 PFU per cell was used to generate the RNA used for Northern (RNA) blots and RNase protection assays. Vero cells were used for plaque assays, and both cell types were used during recombinant virus screening and isolation. The cells were maintained at 37°C under 5% carbon dioxide in Eagle minimum essential medium containing 5% cadet calf serum, 100 U of penicillin per ml, and 100 μg of streptomycin per ml (3, 5, 7, 30).

Enzymes. Restriction enzymes were purchased from Boehringer Mannheim Biochemical Corp.; Bethesda Research Laboratories, Inc.; or New England Biolabs. Restriction buffers recommended by the suppliers were used for all digestions. T7 RNA polymerase and RNasin used in RNase protection were purchased from Promega Biotec, Madison, Wis. T4 polynucleotide kinase (Bethesda Research Laboratories) was used to 5′ end label appropriate markers for RNase protection. Finally, Klenow enzyme used for both labeling and blunting the ends of DNA fragments was purchased from Boehringer Mannheim.

Recombinant DNA. pBR322 and pGEM plasmid derivatives were grown in *Escherichia coli* DH5α. Growth and purification of plasmids were essentially as previously described (1, 7, 18). Oligonucleotide linkers were purchased from Collaborative Research, Inc., Waltham, Mass. DNA fragments for probe generation were isolated by digestion with the appropriate restriction enzymes followed by electrophoresis and either electroelution from agarose or overnight elution from acrylamide at 37°C in a solution of 500 mM ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% (wt/vol) sodium dodecyl sulfate, and 10 μg of *E. coli* tRNA (Sigma) per ml. Probes were uniformly labeled by random priming and synthesis by using Klenow enzyme with [α-^{32}P]dCTP (3 Ci/μmol; Amersham).

RNA isolation and fractionation. Methods for isolation of RNA by using guanidinium isothiocyanate-hot phenol have been previously described (27, 28). RNA generated from infected cells was isolated at specific times postinfection (p.i.) and either used directly as total RNA or first fractionated into poly(A)$^+$ and poly(A)$^-$ fractions by use of oligo(dT)-cellulose (Collaborative Research) chromatography (27).

Phosphonoacetic acid (PAA) or thymine 1-β-D-arabinofuranoside (ara-T) was used to inhibit viral DNA replication. Cells were incubated with PAA (300 μg/ml) for 30 min before infection or with ara-T (50 μg/ml) for 1 h before infection. Drugs were present during viral adsorption and for the times noted in the text.

Actinomycin D (10 μg/ml) was used to study the stability of the mRNAs of interest by its addition 6 h p.i. The infection was allowed to proceed for another 1 or 2 h before RNA isolation.

Recombinant virus. All sequence numbers are based on the complete sequence of the 17*syn*$^+$ strain of HSV-1 (19). A gC recombinant vector extending from the *Sal*I site at base 94853 in the gC region to the *Hin*dIII site at base 98823 was constructed by using appropriate clones from the 17*syn*$^+$ strain of HSV-1. This construct contains an intact gC promoter. The 3.9-kbp *Xba*I-*Bam*HI cassette of pCAL5Div17 (6) contains 48 bases of the U_L38 promoter 5′ of the cap site and 99 bases of U_L38 leader fused to the bacterial β-galac-

tosidase gene with a 4-base linker. It is terminated with the bidirectional simian virus 40 (SV40) termination-polyadenylation sequences. The *Bam*HI site was blunted with Klenow enzyme, and the cassette was ligated into the *Xba*I-*Eco*RV-digested gC recombinant vector, replacing 22 bp of gC sequences. This construct places the chimeric gene in the opposite orientation to and downstream of the gC cap and promoter. As constructed, both a truncated gC and a chimeric β-galactosidase transcript can terminate at the bifunctional SV40 termination-polyadenylation signal. This construct contains a 360-bp *Xba*I-*Kpn*I cassette containing U_L38 promoter-leader sequences and the 5′ portion of the β-galactosidase gene. Promoter constructs with less DNA 5′ of the cap site were constructed by replacement of this cassette with appropriate *Xba*I-*Kpn*I cassettes from the U_L38 promoter constructs described elsewhere (6). A larger U_L38 promoter–β-galactosidase recombinant vector including all the U_L37 and U_L38 promoter regions driving β-galactosidase was generated by first inserting an *Xba*I oligonucleotide linker at the *Sal*I site near the U_L37-CAT junction in pCAL5Div and then cloning this *Xba*I-*Kpn*I cassette into the recombinant vector.

The generation of the recombinant virus FLA5, which contains 48 bases of the U_L38 promoter and the 99-base leader fused to the β-galactosidase gene recombined into the R_L regions of the genome substituting for the latency-associated transcript (LAT) promoter and 5′ sequence, has been described previously (6). It was constructed by using a recombination vector plasmid in which 1,800 bp of HSV DNA of the R_L region containing the LAT promoter and 5′ sequences of the primary latency-associated transcript were replaced with the same *Xba*I-*Bam*HI fragment containing the indicator gene as used in the gC recombination vector plasmid described above. Other R_L recombinant virus constructs were generated by replacing either the *Xba*I-*Kpn*I cassette or the *Xba*I-*Bam*HI cassette of the recombinant vector with appropriate modifications or in reversed orientation in the R_L region.

Recombinant virus was generated by first releasing the HSV-1 sequences in the plasmid vectors by digestion with appropriate restriction enzymes and then cotransfecting cloned DNA sequences with infectious 17*syn*$^+$ DNA into rabbit skin fibroblasts. Single-virus plaques derived by limiting dilution were screened in 96-well plates by hybridization of transfection aliquots with ^{32}P-labeled β-galactosidase probe. Wells containing recombinant viral DNA were then plaque purified at least three more times by repetition of the following procedure. Isolated plaques were picked from agar overlays of limiting dilutions of virus infecting Vero cells into 96-well plates containing rabbit skin fibroblasts and screened by hybridization. The resulting isolates were then analyzed by Southern blot. Purified isolates containing the complete promoter–β-galactosidase construct inserted into the gC region for gC recombinant virus or into both the IR_L and TR_L regions for the R_L recombinant virus were used in this report. Plaque-purified isolates from different transfections were taken for each construct, and the level of expression of the chimeric β-galactosidase gene was the same as assayed by colorimetric methods. The equivalency of levels of RNA expression from different recombinants bearing differing lengths of upstream promoter sequence or in different orientations confirmed that the expression seen was a result of the specific recombinations characterized and was not due to some nonspecific mutation in another region of the viral genome.

RNase protection assays. RNase protection assays have been described elsewhere (4, 9). pGEM templates were used with T7 polymerase (Promega Biotec) by following instructions supplied to generate the RNA probes. The probe was labeled with 100 μCi of [^{32}P]UTP (800 Ci/mmol; Amersham) per μg of template DNA. The U_L38 protection probe was a *Hha*I fragment generated from the U_L38–β-galactosidase chimeric recombinant virus construct. The *Hha*I fragment spans the region of DNA from 5 bp 5′ of the U_L38 cap to 28 bp within the β-galactosidase leader. This probe protects 99 bases of wild-type U_L38 and 127 bases of U_L38–β-galactosidase recombinant mRNA. A control dUTPase protection probe was a *Bgl*II-*Kpn*I fragment extending 62 bp 5′ of the cap to 95 bp 3′ of the cap. This dUTPase probe protects 95 bases of 17*syn*$^+$ wild-type dUTPase mRNA.

More than 95% of total RNase protection probes synthesized with T7 polymerase were full length on the basis of results with analytical gels. A total of 1.5×10^6 cpm (Cerenkov) of probe was hybridized with 10 to 20 μg of appropriate RNA samples in 50 μl of 80% formamide–40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.4)–400mM NaCl–1mM EDTA at 60°C for 11 h. The mixture was then digested by addition of 400 μl of RNase T_1 (2 μg/ml; Sigma) in 10 mM Tris (pH 7.5)–300 mM NaCl–5mM EDTA for 1 h at 30°C. The digestion mix was next treated with 60 μg of proteinase K (Sigma) at 37°C for 15 min and phenol-chloroform extracted. RNase-resistant material was fractionated on an 8% acrylamide–8 M urea sequencing gel with DNA size standards.

Northern blot analysis. RNA transfer blots were hybridized and washed essentially as described previously (6, 31). An 8-ml volume containing 4×10^7 cpm (Cerenkov) of radiolabeled random-primed DNA in the presence of 50% formamide, 0.4 M Na^+, 0.1 M HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 8.0), 0.005 M EDTA, and Denhardt's solution containing 100 μg of denatured calf thymus DNA per ml was used at 50°C for 40 h.

RESULTS

Placement of U_L38 promoter-controlled expression marker in different locations in the HSV-1 genome. To eliminate the influence of transcription signals upstream of the U_L38 core promoter, we generated recombinant viruses in which the promoter-indicator gene cassette was placed in the inverted orientation in the gC locus. The general structure of such recombinants is shown in Fig. 1A; the 5′ terminus of the U_L38 promoter is bounded by an *Xba*I site at base 97669. This is 250 bases 5′ of the cap site of the 730-base γ mRNA encoding the U_L45 gene, which is coterminal with the 3′ end of gC (8). The genomic structure and purity of recombinants were confirmed by Southern blots of viral DNA. As shown in Fig. 1A, the 3,600-bp *Sal*I fragment was replaced by the 7,440-bp fragment containing the β-galactosidase gene, various portions of the U_L38 promoter, and the SV40 termination-polyadenylation sequences described in Materials and Methods.

The bidirectional activity of the SV40 polyadenylation signal was confirmed by analysis of RNA expressed by recombinant viruses containing the chimeric U_L38–β-galactosidase reporter gene in the gC region. Figure 2A shows an example in which Northern blots of poly(A)$^+$ RNA from cells infected with the U_L38 promoter construct from −48 to +99 expressed a truncated 1.5-kb portion of gC mRNA (track ii) compared with the 2.7-kb wt gC mRNA expressed in cells infected with the 17*syn*$^+$ strain (wt, track iii).

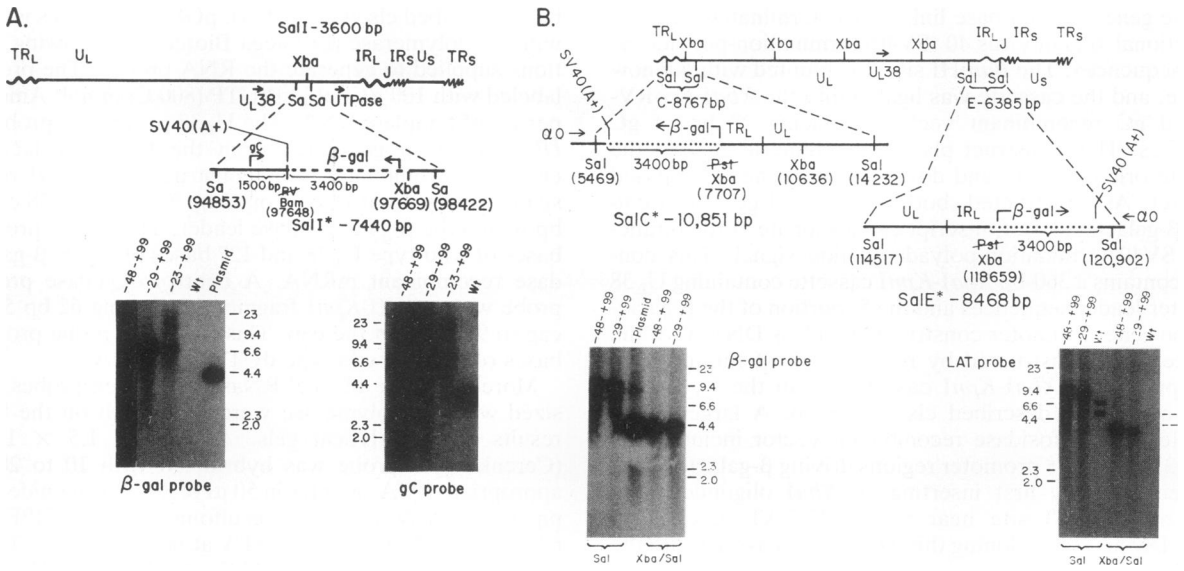

FIG. 1. HSV-1 recombinant viruses. The location and precise sequence numbers where the reporter gene-promoter constructs were inserted into the HSV genome are indicated in the maps. (A) Schematic representation of the expected structure of recombinants containing the U_L38 promoter-controlled β-galactosidase gene recombined into the gC region of the HSV-1 genome. Insertion of the promoter–β-galactosidase constructs generates a *Sal*I fragment (7,440 bp) which is larger than the wt fragment (3,600 bp). Below the map are shown diagnostic Southern blots demonstrating the insertion of the β-galactosidase fragment. Digestion with *Sal*I generates the expected-sized fragment for the recombinant viruses which hybridizes to both β-galactosidase-specific and HSV-1-specific probes. The latter probe was a 945-bp *Pst*I-*Sal*I fragment (corresponding to bases 97477 through 98422 of the complete 17*syn*$^+$ sequence [19]) cloned from the gC region of wt HSV-1 DNA. By contrast, wt viral DNA digested with *Sal*I generates the expected fragment which hybridized only to the gC-specific probe. (B) Schematic representation of the expected structure of the R_L recombinants with the U_L38 promoter-controlled β-galactosidase gene recombined into both the TR_L and the IR_L regions of the HSV-1 genome. Insertion into both regions generates two new fragments upon digestion with *Sal*I: *Sal*C*(10,851 bp) and *Sal*E* (8,468 bp). These fragments are larger than those generated by *Sal*I digestion of wt HSV-1 DNA: fragments *Sal*C (8,768 bp) and *Sal*E (6,385 bp), respectively. The recombinant fragments hybridize to both β-galactosidase-specific and wt HSV-1-specific probes. The latter probe was a 605-bp *Hpa*I-*Sal*I fragment (corresponding to bases 120297 through 120902) from the R_L region of wt DNA. Digestion with *Xba*I and *Sal*I generated a 3.9-kb fragment from the R_L recombinants. This fragment hybridizes with both β-galactosidase and HSV-1 probes. The same digest of wt DNA generated 6.3- and 5.2-kb fragments, which were detectable only with the HSV-1-specific probe.

Hybridization of blots with probes specific for U_L38 mRNA revealed both the normal 1.9-kb U_L38 mRNA and a 3.4-kb chimeric mRNA containing the 99 bases of U_L38 leader and β-galactosidase sequences, whereas RNA from wt-infected cells expressed only U_L38 mRNA (Fig. 2B, tracks i and ii). Hybridization of blots with β-galactosidase-specific probes reveals only the 3.4-kb mRNA (data not shown).

The U_L38 promoter–β-galactosidase reporter cassette was also recombined in both orientations into the R_L region of the HSV-1 genome by the substitution of the proximal promoter and 5′ regions of the latency-associated transcription unit. Examples of Southern blots of such recombinants showing recombination into both repeats are shown in Fig. 1B; recombination results in the generation of the 10.8-kbp *Sal*I fragment C* and the 8.5-kbp *Sal*I fragment E* derivatives containing the β-galactosidase reporter gene construct replacing the 8.8- and 6.4-kbp parental fragments.

Recombination of the reporter construct into the R_L region also yielded recombinant virus which exhibited significant expression of the chimeric U_L38–β-galactosidase mRNA promoter as well as normal U_L38 mRNA (Fig. 2B, track iii). These recombinants expressed normal-sized gC mRNA (Fig. 2A, track i). In other experiments (results not shown) we determined that recombination in either region of the HSV genome had no effect upon the size or the amount of α0 mRNA, which terminates just 3′ of the site of insertion of the reporter construct into the R_L region.

Amounts of reporter gene mRNA expressed are influenced by the genomic position of the promoter controlling them. Although U_L38 promoter-controlled β-galactosidase mRNA was efficiently expressed from recombinant viruses bearing the cassette either in the gC location or in the R_L regions, the levels of chimeric U_L38–β-galactosidase mRNA seen in Northern blots was always greater for the latter constructs. We confirmed and quantitated this difference in mRNA levels by use of RNase protection assays with an RNase probe extending from a *Hha*I site in the β-galactosidase sequences of the reporter cassette 28 bases 3′ of the junction of the U_L38 leader to a *Hha*I site 5′ of the mRNA cap site (see Materials and Methods). Chimeric U_L38–β-galactosidase mRNA protected 127 bases of this probe, whereas wt U_L38 mRNA protected 99 bases. Since the probe is uniformly labeled and since protected fragments have essentially the same size for both mRNAs, a direct comparison of the intensity of RNase-resistant hybrids results in a direct measure of the relative amounts of both mRNAs in infected cells in which the amount of wt U_L38 mRNA provides an internal standard. Further, since the structure of the chimeric mRNA was the same for all recombinants, any differences in levels of mRNA must reflect differences in levels of expression and not differential mRNA stability or other factors.

Our data were identical for U_L38 promoters containing 48 and 29 bases of DNA sequence upstream of the cap site. Results of a typical experiment are shown in Fig. 3; the ratio

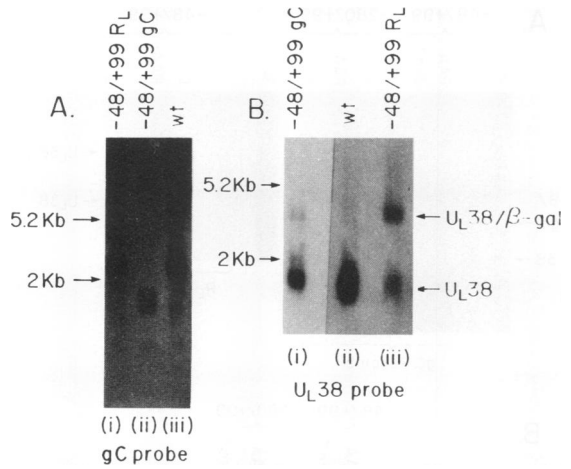

FIG. 2. Northern blot analysis of chimeric and gC mRNAs expressed during infection with recombinant viruses. Samples of poly(A)$^+$ RNA were isolated from rabbit skin fibroblasts 4 h after infection (hpi) with recombinant or wt virus, fractionated, blotted, and hybridized. (A) Representative Northern blot probed with an 896-bp *Eco*RI-*Eco*RV DNA fragment specific for gC (U_L44) mRNA (corresponding to bases 96524 through 95628). The 1.5-kb mRNA expressed by the recombinants (track ii) indicates that recombination into the gC region results in truncation of the gC mRNA. Both the R_L recombinant (track i) and parental wt (track iii) viruses express full-length gC mRNA (2.7 kb), as expected. (B) Representative Northern blot of gC and R_L recombinants along with their parental wt RNA probed with the *Xba*I-*Kpn*I fragment from the −48/+99–β-galactosidase construct. This fragment hybridizes to wt U_L38 mRNA and to β-galactosidase-containing mRNA. All three virus types express wt U_L38 mRNA (1.9 kb), whereas only the two recombinants express the U_L38–β-galactosidase mRNA (3.4 kb).

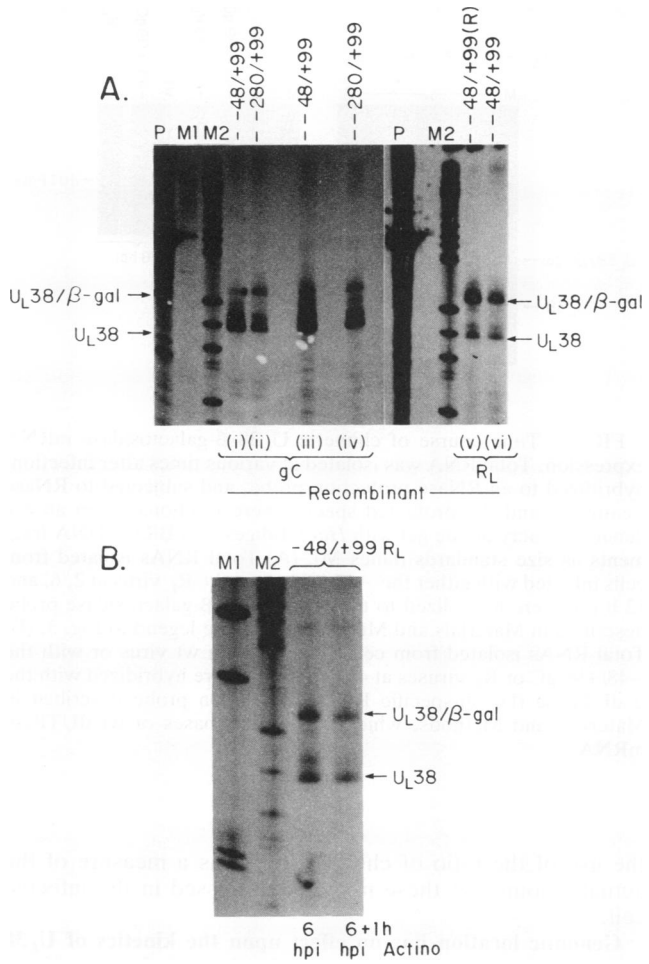

FIG. 3. Quantitation of the level of chimeric U_L38–β-galactosidase RNA expression in recombinant virus. Total RNA was isolated from rabbit skin fibroblasts and hybridized to the *Hha*I U_L38–β-galactosidase probe described in Materials and Methods. This probe protects 99 bases of wt U_L38 mRNA and 127 bases of chimeric U_L38–β-galactosidase mRNA. After hybridization and RNase digestion, the RNase-resistant material was fractionated on an 8% denaturing acrylamide gel with *Eco*RI-*Hin*fI- (M1) and *Hae*III (M2)-digested pBR322 DNA fragments as size markers. (A) Ten micrograms (tracks i, ii, v, and vi) or 20 μg (tracks iii and iv) of RNA isolated 6 h p.i. was used for each virus. Shown are RNA probe fragments protected by RNA isolated from cells infected with two gC recombinant viruses (−48/+99, tracks i and iii; −280/+99, tracks ii and iv) and two R_L recombinant viruses [−48/+99, track vi; −48/+99(R) in which the whole promoter-reporter gene construct is reversed, track v]. A lane containing 600,000 cpm (Cerenkov) of undigested probe (P) was included to check the labeling efficiency of the probe. The size of the undigested probe is 197 bp owing to the presence of the polylinker sequence. (B) RNA was isolated from cells infected with the −48/+99 R_L recombinant virus at 6 h p.i. and again after a 1-h treatment with 10 μg of actinomycin D per ml starting 6 h p.i. The markers M1 and M2 were the same as described above.

of chimeric to wt mRNA was determined densitometrically in two to four separate gels. Ratios varied ±10% between gels. Chimeric U_L38–β-galactosidase mRNA expressed from the U_L38 promoter in the gC region recombinants was reproducibly present at 20% of the level of wt U_L38 mRNA (tracks i and iii). In contrast, recombinant viruses with the same promoter construct recombined into either orientation in the R_L region expressed essentially twice as much chimeric mRNA as wt U_L38 mRNA (tracks v and vi). Since the recombinant promoter-reporter gene construct is present in two copies in the R_L constructs, our data indicate that the U_L38 promoter is as active in the R_L region as it is in its normal location.

The lower level of expression of the U_L38 promoter in the gC location could be partially alleviated by the incorporation of DNA sequences extending beyond 48 bases of the mRNA cap site. Thus, a construct containing 232 bases of DNA sequence inserted upstream of the 48-base U_L38 promoter resulted in a two- to threefold increase in chimeric mRNA levels (Fig. 3, tracks ii and iv). This additional sequence contained the DNA sequences extending from the cap site of the U_L37 mRNA, through its promoter, and including U_L38 sequences 5′ of −48. Once inserted, it reconstructed the wt U_L37-U_L38 juxtaposed promoters with the U_L38 promoter linked to β-galactosidase.

We confirmed the validity of our measurements of mRNA levels as an index of promoter activity by comparing the stability of chimeric U_L38–β-galactosidase mRNA with that of the wt transcript by using actinomycin chase experiments. The ratios of wt to chimeric mRNA do not change following 60 to 120 min of chase after 6 h of infection with recombinant virus; an example of a 60-min chase is shown in Fig. 3B. Therefore, the incorporation of bacterial β-galactosidase sequences in the reporter mRNA does not lead to any significant increase or decrease in the stability of chimeric as compared with wt U_L38 mRNA which would interfere with

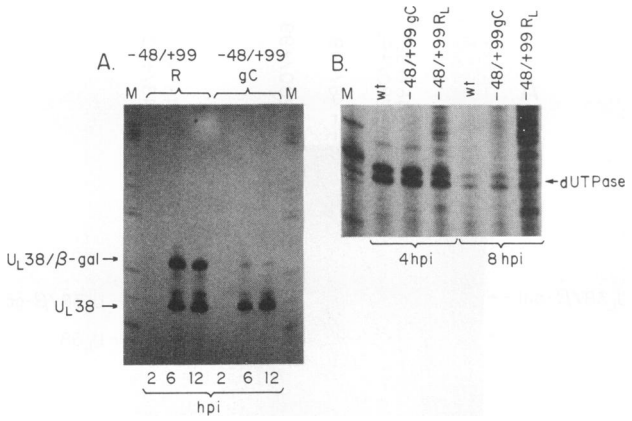

FIG. 4. Time course of chimeric U_L38–β-galactosidase mRNA expression. Total RNA was isolated at various times after infection, hybridized to an RNase protection probe, and subjected to RNase treatment, and the protected species were fractionated on an 8% denaturing acrylamide gel with *Hae*III-digested PBR322 DNA fragments as size standards (lanes M). (A) Total RNAs isolated from cells infected with either the −48/+99 gC or the R_L virus at 2, 6, and 12 h p.i. were hybridized to the *Hha*I U_L38–β-galactosidase probe described in Materials and Methods and in the legend to Fig. 3. (B) Total RNAs isolated from cells infected with wt virus or with the −48/+99 gC or R_L viruses at 4 and 6 h p.i. were hybridized with the β dUTPase (U_L50)-specific RNase protection probe described in Materials and Methods, which protects 95 bases of wt dUTPase mRNA.

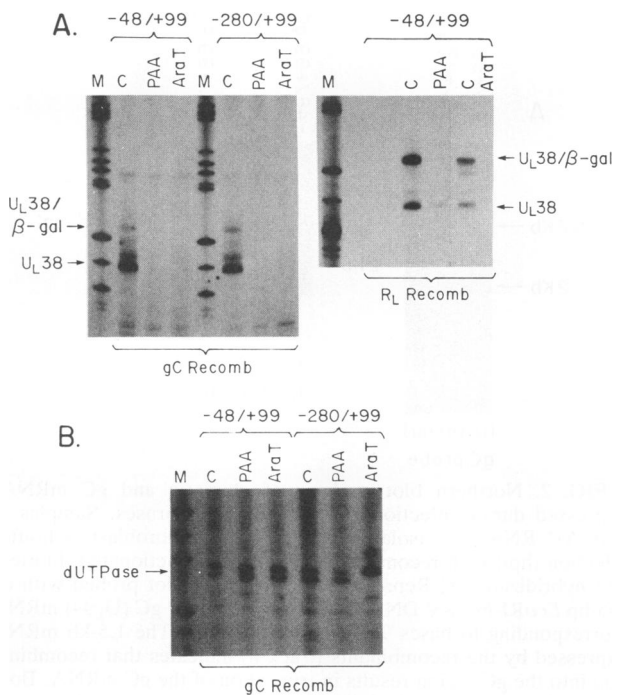

FIG. 5. Effect of DNA replication on the expression of chimeric U_L38–β-galactosidase mRNA. Total RNA was isolated 6 h p.i. from rabbit skin fibroblasts infected with recombinant virus in the presence or absence (control [lanes C]) of either 50 μg of ara-T per ml or 300 μg of PAA per ml. The RNA was hybridized to an RNase protection probe, treated with RNase, and fractionated on an 8% denaturing acrylamide gel with *Hae*III-digested pBR322 DNA fragments as size standards (lanes M). (A) RNA was hybridized to the *Hha*I U_L38–β-galactosidase probe described in Materials and Methods and in the legend to Fig. 3. The panel on the left shows an experiment with the −48/+99 and the −280/+99 gC recombinant viruses. The panel on the right shows an experiment with the −48/+99 R_L recombinant virus. (B) The same RNA samples from the gC recombinant virus-infected cells used in panel A were hybridized with the dUTPase probe described in Materials and Methods and the legend to Fig. 4.

the use of the ratio of chimeric to wt as a measure of the actual amounts of these mRNAs expressed in the infected cell.

Genomic location has no effect upon the kinetics of U_L38 promoter activity. We next examined the influence of genomic position on the kinetics of U_L38 promoter activity. The expression of wt U_L38 mRNA conforms to strict late (γ) kinetics; thus expression of this mRNA absolutely requires viral DNA replication (2, 31). Despite the variability in the levels of chimeric mRNA expressed from this promoter in the different recombinants, we found that its expression kinetics were essentially the same when it was recombined into either the gC or R_L location in the genome. Results of a typical experiment are shown in Fig. 4. RNA isolated at various times after infection with wt or recombinant viruses was analyzed for the presence of U_L38 and chimeric U_L38–β-galactosidase mRNA. No U_L38 or chimeric mRNA from recombinants in gC or in the R_L region could be detected at 2 h p.i. In contrast, these mRNAs were readily detectable by 6 to 12 h p.i. when DNA replication is occurring at high rates. As a control, RNase protection assays with a probe specific for the β class dUTPase (U_L50) mRNA yielded relatively high levels at 2 to 4 h p.i., with significantly less RNA recoverable from cells by 6 to 8 h p.i. Sample data for 4 and 8 h p.i. RNA are shown in Fig. 4B.

A strict requirement for DNA replication was also shown by using DNA synthesis inhibitors. In the presence of either 50 μg of ara-T per ml or 300 μg of PAA per ml, no wt or chimeric mRNA was expressed (Fig. 5A). By contrast, similar samples of RNA hybridized with the β dUTPase mRNA probe yielded levels as high as or higher than those in the untreated controls (Fig. 5B).

Basal U_L38 promoter activity requires sequences extending no farther than 29 bases 5′ of the transcript cap site. We next investigated the minimal sequence elements required for measurable U_L38 promoter activity with recombinants in both gC and the R_L regions. Reporter constructs with promoters extending to −29 bases expressed the same levels of chimeric U_L38–β-galactosidase mRNA as did those with −48 bases of promoter in both locations (Fig. 6A). A promoter construct extending to only −23 bases, however, expressed no measurable level of chimeric mRNA. This is shown for a gC recombinant in Fig. 6B. These results are somewhat at variance with previously reported results of transient-expression assays in which the −23 promoter still had about 30% of the activity of the more extensive ones (6). Extremely long exposures of similar RNase protection gels and examination of hybrids formed by using more than twice the amount of infected-cell RNA revealed no RNase protection product corresponding to the chimeric U_L38–β-galactosidase mRNA (data not shown). On the basis of control experiments with mixtures of chimeric and wt mRNAs, this places levels of expression from the −23 promoter constructs as being lower than 5% of wt levels.

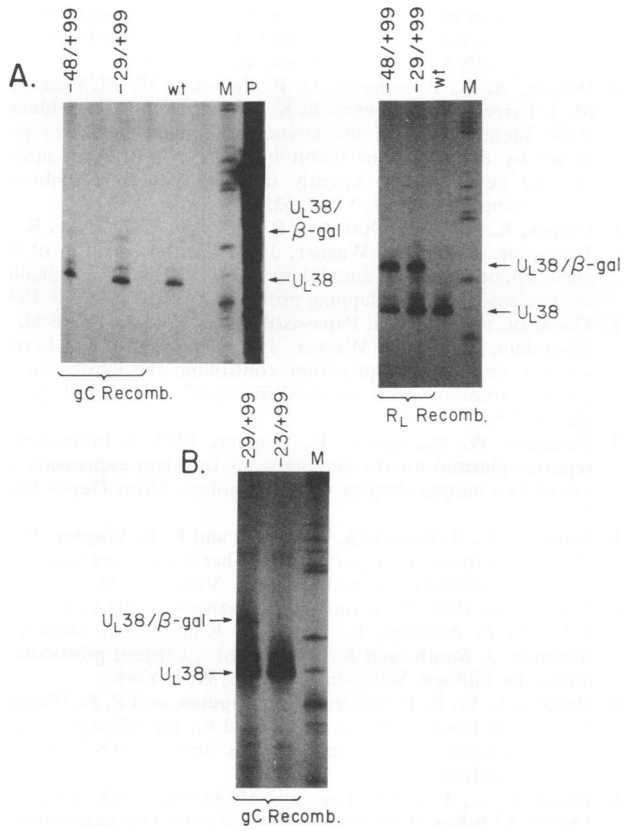

FIG. 6. U_L38 promoter activity requires sequences extending no further than 29 bases 5′ of the transcript cap site. Total RNA was isolated 8 h p.i. from rabbit skin fibroblasts and hybridized to the *Hha*I U_L38–β-galactosidase probe described in Materials and Methods and the legend to Fig. 3. After treatment with RNase, the protected fragments were fractionated on an 8% denaturing acrylamide gel with *Hae*III-digested pBR322 DNA fragments as size standards (M). (A) The left panel shows results of RNase protection with 10 μg of RNA from cells infected with wt, −48/+99, and −29/+99 gC recombinant viruses. This panel also includes a lane in which 600,000 cpm (Cerenkov) of undigested probe was loaded (lane P). The full-length probe was 197 bp because it included pGEM polylinker sequences. The panel on the right shows the results of using 10 μg of total RNA from cells infected with wt, −48/+99, and −29/+99 R_L recombinant viruses. (B) RNase protection with 10 μg of total RNA isolated from two gC recombinant viruses, −29/+99 and −23/+99.

DISCUSSION

One of the reasons for initiating the experiments described in this communication was to develop an internally controlled system in which the expression of HSV promoters bearing defined modifications could be measured and compared with that of the unmodified parental promoter from the same viral genome. It is important to emphasize that differences in expression do not reflect some sort of nonspecific interference due to transcription through the inserted reporter gene-promoter construct since all RNase protection and Northern blot analyses confirm the lack of detectable levels of such readthrough RNA. Such a system is clearly feasible on the basis of the data presented here, which also demonstrate that the activity of a model promoter controlling the same transcript can vary by more than a factor of 5 depending on where it is expressed on the viral genome. Such a result has practical significance in the design of HSV vectors for moving and expressing specific genes into specific infected tissues or cells. It is also important in interpreting the results of deletional and substitutional mutagenesis studies on HSV promoter function and in the identification of promoter sequences controlling the kinetics of expression of particular HSV genes.

Our picture of the essential sequence elements in γ promoters is generally based on the modifications of the gC promoter in situ by Homa et al. (11–13). These researchers have demonstrated that a very limited amount of sequence 5′ of the cap site, as well as some elements 3′, of it are all that are required for full activity. Further, these workers have suggested that the sequence of the TATA box element itself (−28 to −23 of the cap) is a determining factor in its activity (11–13). Similar in situ modification analyses on other model γ promoters such as that controlling the expression of mRNAs encoding the gH (U_L22), γ42 (U_L49), and U_s11 proteins generally confirm the limited extent of the 5′ sequence required to control strict late transcription, although it is clear that no one specific TATA box sequence is conserved (16, 17, 22, 24). The present data on the U_L38 promoter, which removes the influence of any possible sequence elements extending farther upstream than the regions modified in such in situ promoter modification studies, fully support this general picture of HSV γ promoters and demonstrate that very limited amounts of sequence elements upstream of the transcription cap, along with other potential elements 3′ of this site, provide all the information necessary for the expression of transcripts following viral DNA replication (6).

The data on the extent of the U_L38 promoter upstream of the cap site presented here are in general agreement with our transient-expression data presented earlier (6), with the notable exception that in the viral genome the region extending between −29 and −23 corresponding to the sequence TTTAAA is absolutely essential for promoter function; thus, the canonical sequence TATA occurring at −17 of the cap site cannot alone mediate any appreciable transcriptional activity. This result is in contrast with the situation found in transient-expression assays, in which constructs extending to −27 and −23 of the cap site (the latter eliminating the whole upstream T+A-rich element) had activities on the order of 30% of full promoter expression. Such data provide a good demonstration of the limitations of the use of transient-expression assays as a sole measure of promoter activity in mutational modification analyses.

Although all available evidence demonstrates that the 5′ extent of HSV γ promoters is quite limited and is transportable as cassettes to different locations of the viral genome, it is clear that the environment of this core promoter can have a significant role in the levels of its activity. In the U_L38 promoter, the addition of a relatively small amount of extra sequence corresponding to the U_L37 promoter and upstream sequences in their normal position relative to U_L38 make the γ promoter significantly more active in the gC region used for generating recombinants.

Two observations suggest that the low level of activity of the minimally active U_L38 promoter in gC is not due to the general inability of a promoter to be efficiently expressed in this region. First, inclusion of extra sequence significantly increases activity. Second, experiments currently in progress in which 219 bases of the dUTPase promoter is being used to drive the expression of similar chimeric β-galactosidase-containing constructs demonstrate essentially equivalent levels of mRNA expression compared with the wt

mRNA. It is not clear, however, whether the low activity of the minimal U_L38 promoter is due to its kinetic class or whether minimal promoter elements of other classes of transcripts would also be expressed at lower than wt levels in this region.

Although the core U_L38 promoter is sufficient to define the kinetic class of this gene, its significantly lower activity in the gC location than in the R_L implies that this minimal core promoter does not contain all the information requirements for normal levels of mRNA expression. Although DNA sequence elements inserted upstream of the core U_L38 promoter in the gC region are important in influencing its level of expression, they do not appear to be highly specific. This conclusion is based on the fact that the same core promoter placed in the R_L region in an orientation in which the promoter replaces 220 bases of the LAT promoter is active at levels approaching those seen for the promoter in its normal location. Further, the same level of activity is seen when the whole cassette is inverted in the R_L region so that the sequences immediately 5′ of the U_L38 core promoter are completely different. Thus, three different sets of upstream sequences, only one of which encodes another promoter-regulatory region, permit essentially full activity of the core promoter.

Other data support our conclusion that the local environment of the promoter, as well as the specific sequences within it, is important in influencing levels of expression of HSV transcripts. Roemer et al. (20) found that the location of an inserted SV40 promoter-enhancer-controlled reporter gene in the viral genome affected the levels of expression. Further, removal of recognizable α4-binding sites within the promoter of the β TK (U_L23) transcript has a profound activity upon its expression in transient-expression assays, but substitution of such modified promoters for the wt one in the viral genome does not lead to loss of TK expression (15). A similar observation has been reported for gD (23, 25). This result has been interpreted as indicating that the viral genome has a high degree of redundancy in sequence elements which bind either cellular or viral transcription factors so that loss of one or a few in any given region can be readily compensated for. Our data would suggest that the DNA sequence upstream of the core U_L38 promoter which includes the U_L37 promoter acting in the opposite direction and the DNA sequence in the R_L regions in which this promoter was recombined contain enough such sites that U_L38 promoter activity is high. However, our data also show that such sites are not uniformly distributed throughout the viral genome since the region within the gC locus used for recombination is a significantly poorer environment for transcription than is the R_L region.

ACKNOWLEDGMENTS

This work was supported by U.S. Public Health Service grants CA11861 and AI06246. Further support was provided by the UC Irvine Research Unit in Animal Virology. J.F.G. is a predoctoral trainee and the recipient of training grant T32GM07311 ("Structure and synthesis of biological macromolecules").

REFERENCES

1. **Anderson, K. P., R. J. Frink, G. B. Devi-Rao, B. H. Gaylord, R. H. Costa, and E. K. Wagner.** 1981. Detailed characterization of the mRNA mapping in the *Hin*dIII fragment K region of the herpes simplex virus type 1 genome. J. Virol. **37**:1011–1027.
2. **Anderson, K. P., L. E. Holland, and E. K. Wagner.** 1980. Characterization of herpes simplex virus type 1 RNA present in the absence of de novo protein synthesis. J. Virol. **34**:9–27.
3. **Blair, E. D., C. C. Blair, and E. K. Wagner.** 1987. Herpes simplex virus virion stimulatory protein mRNA leader contains sequence elements which increase both virus-induced transcription and mRNA stability. J. Virol. **61**:2499–2508.
4. **Dobson, A. T., F. Sedarati, G. B. Devi-Rao, W. M. Flanagan, M. J. Farrell, J. G. Stevens, E. K. Wagner, and L. T. Feldman.** 1989. Identification of the latency-associated transcript promoter by expression of rabbit beta-globin mRNA in mouse sensory nerve ganglia latently infected with a recombinant herpes simplex virus. J. Virol. **63**:3844–3851.
5. **Draper, K. G., G. B. Devi-Rao, R. H. Costa, E. D. Blair, R. L. Thompson, and E. K. Wagner.** 1986. Characterization of the genes encoding herpes simplex virus type 1 and type 2 alkaline exonucleases and overlapping proteins. J. Virol. **57**:1023–1036.
6. **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 U_L38, a true late gene involved in capsid assembly. J. Virol. **65**:769–786.
7. **Flanagan, W. M., and E. K. Wagner.** 1987. A bi-functional reporter plasmid for the simultaneous transient expression assay of two herpes simplex virus promoters. Virus Genes **1**:61–71.
8. **Frink, R. J., R. Eisenberg, G. Cohen, and E. K. Wagner.** 1983. Detailed analysis of the portion of the herpes simplex virus type 1 genome encoding glycoprotein C. J. Virol. **45**:634–647.
9. **Gilman, M.** 1989. Preparation and analysis of RNA, p. 4.7.1–4.7.8. *In* F. Ausubel, R. Brent, R. Kingston, J. Moore, J. Seidman, J. Smith, and K. Struchl (ed.), Current protocols in molecular biology. Wiley-Interscience, New York.
10. **Holland, L. E., K. P. Anderson, C. Shipman, 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.
11. **Homa, F. L., J. C. Glorioso, and M. Levine.** 1988. A specific 15-bp TATA box promoter element is required for expression of a herpes simplex virus type 1 late gene. Genes Dev. **2**:40–53.
12. **Homa, F. L., A. Krikos, J. C. Glorioso, and M. Levine.** 1991. Functional analysis of regulatory regions controlling strict late HSV gene expression, p. 207–222. *In* E. K. Wagner (ed.), Herpesvirus transcription and its regulation. CRC Press, Inc., Boca Raton, Fla.
13. **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.
14. **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.
15. **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.
16. **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.
17. **Johnson, P. A., C. MacLean, H. S. Marsden, R. G. Dalziel, and R. D. Everett.** 1986. The product of gene U_s11 of herpes simplex virus type 1 is expressed as a true late gene. J. Gen. Virol. **67**:871–883.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **McGeoch, D. J.** 1989. The genomes of the human herpesviruses: contents, relationships, and evolution. Annu. Rev. Microbiol. **43**:235–265.
20. **Roemer, K., P. A. Johnson, and T. Friedmann.** 1991. Activity of the simian virus 40 early promoter-enhancer in herpes simplex virus type 1 vectors is dependent on its position, the infected cell type, and the presence of Vmw175. J. Virol. **65**:6900–6912.
21. **Roizman, B., and A. E. Sears.** 1990. Herpes simplex viruses and their replication, p. 1795–1842. *In* B. N. Fields, D. M. Knipe,

R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Virology, 2nd ed. Raven Press, New York.

22. **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.
23. **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 D (gD) promoter are not essential for efficient gD transcription during virus infection. J. Virol. **66**:623–631.
24. **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.
25. **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.
26. **Wagner, E. K.** 1985. Individual HSV transcripts: characterization of specific genes, p. 45–104. *In* B. Roizman (ed.), The herpesviruses, vol. III. Plenum Press, New York.
27. **Wagner, E. K., G. B. Devi-Rao, L. T. Feldman, A. T. Dobson, Y. F. Zhang, W. M. Flanagan, and J. G. Stevens.** 1988. Physical characterization of the herpes simplex virus latency-associated transcript in neurons. J. Virol. **62**:1194–1202.
28. **Wagner, E. K., W. M. Flanagan, G. B. Devi-Rao, Y. F. Zhang, J. M. Hill, K. P. Anderson, and J. G. Stevens.** 1988. The herpes simplex virus latency-associated transcript is spliced during the latent phase of infection. J. Virol. **62**:4577–4585.
29. **Weinheimer, S. P., and S. L. McKnight.** 1987. Transcriptional and post-transcriptional controls establish the cascade of herpes simplex virus protein synthesis. J. Mol. Biol. **195**:819–833.
30. **Zhang, Y. F., G. B. Devi-Rao, M. Rice, R. M. Sandri-Goldin, and E. K. Wagner.** 1987. The effect of elevated levels of herpes simplex virus alpha-gene products on the expression of model early and late genes in vivo. Virology **157**:99–106.
31. **Zhang, Y. F., and E. K. Wagner.** 1987. The kinetics of expression of individual herpes simplex virus type 1 transcripts. Virus Genes **1**:49–60.