JOURNAL OF VIROLOGY, Aug. 1983, p. 301–310
0022-538X/83/080301-10$02.00/0

Vol. 47, No. 2

Positive Control of the Herpes Simplex Virus Thymidine Kinase Gene Requires Upstream DNA Sequences

JAMES R. SMILEY,* HELEN SWAN, MARY M. PATER,† ALAN PATER,† AND MARNIE E. HALPERN

Pathology Department, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

Received 14 February 1983/Accepted 9 May 1983

We examined the ability of deleted versions of the herpes simplex virus type 1 thymidine kinase gene, present in biochemically transformed mouse cells, to be induced in *trans* to a higher level of expression by superinfecting herpes simplex virus immediate early gene products. The results demonstrate that sequences mapping between −200 and −80 and between −70 and −12 are required for induction. As these regions are largely coincident with the previously identified thymidine kinase gene promoter, the results suggest that herpes simplex virus immediate early gene products or their metabolic product activate thymidine kinase expression by acting at the promoter region to increase the rate of transcription.

The genes of herpes simplex virus (HSV) fall into at least three differentially regulated classes, whose expression is controlled at the level of transcription by viral proteins (9, 11, 20, 27). Expression during lytic infection of the viral gene for thymidine kinase (TK), a delayed early gene, requires the prior synthesis and continued activity of at least one immediate early (IE) viral protein, ICP4 (11, 20, 27). This requirement for IE function is partially removed when the purified TK gene is transfected into uninfected cultured cells so that stable transformants can be isolated which express low levels of TK in the absence of IE proteins (28). In addition, the purified TK gene is expressed after microinjection into *Xenopus* oocytes (16). In both of these cases, transcription initiates at the same site as in lytic viral infections (16, 25; M. J. Wagner, W. C. Summers, J. A. Sharp, J. R. Smiley, and W. P. Summers, submitted for publication). The DNA sequences required for basal level expression in uninfected cells have been mapped by McKnight and coworkers, who demonstrated that at least three separate upstream regions are needed (15–17). Although the mechanisms which allow low-level expression in uninfected cells in the absence of IE proteins are obscure, it is clear that the transfected gene remains susceptible to viral positive control. Superinfection of transformed cells with TK-deficient HSV results in a substantial increase in both TK enzymatic activity and TK mRNA specified by the resident TK gene (10, 12). This enhancement requires the same IE functions that activate TK transcription from the intact viral genome during lytic infection. These observations suggest that the resident TK DNA gene present in transformed cells retains control sequences which respond to IE transcriptional regulators supplied by the superinfecting virus. We sought to identify these presumed target sequences by testing the ability of deleted versions of the TK gene to respond to the IE proteins supplied in *trans* by superinfecting virus. Our experimental strategy was first to transform TK-deficient LtA cells to the TK^+ phenotype with TK genes bearing various deletions and then to assess the effects of each of these deletions on the response to IE proteins after infection of the resulting transformants with TK-deficient HSV type 1 (HSV-1). The results demonstrate that positive control requires the integrity of DNA sequences located upstream from the transcription initiation site.

† Present address: Department of Microbiology, College of Medicine and Dentistry, New Jersey Medical School, Newark, NJ 07103.

MATERIALS AND METHODS

Virus and cells. TK-deficient HSV-1 B2006 (3) was grown and titrated on Vero cells. Vero and LtA cells were grown in α-minimal essential medium containing 5 and 10% fetal bovine serum, respectively, and maintained at 37°C in a humidified atmosphere containing 5% CO_2.

Plasmids. pX1 contains the 3.5-kilobase *Bam*HI Q fragment of HSV DNA cloned into the *Bam*HI site on pBR322 (4). pJS1 contains the larger *Bgl*II subfragment of *Bam*HI-Q inserted at the *Bam*HI site on pBR322. p1EN1, obtained from W. C. Leung, carries the *Eco*RI N fragment of HSV-1 DNA at the *Eco*RI site of pBR325. pTK173, obtained from R. McKinnon and F. L. Graham, contains the largest *Pvu*II subfragment of *Bam*HI-Q in the *Pvu*II site of pBR322. Plas-

mids were maintained in *Escherichia coli* LE392 or HB101.

pΔ1 and pΔ35 were derived from pX1 by limited *Bal* 31 digestion at the single *Bgl*II site, followed by recircularization with T4 DNA ligase. Twenty nanograms of *Bgl*II-cleaved DNA were incubated in 100 μl of 12 mM $CaCl_2$–12 mM $MgCl_2$–600 mM NaCl–40 mM Tris-hydrochloride–0.2 mM EDTA (pH 8.1) containing 4 units of *Bal* 31 at 30°C. Samples (20 μl) were withdrawn after 0.5, 1, 1.5, 3, and 6 min of incubation into 0.2 ml of ice-cold 200 mM EDTA (pH 8.0). The DNA was extracted twice with phenol and once with ether then ethanol precipitated and suspended in 100 μl of 66 mM Tris-hydrochloride (pH 7.6)–6.6 mM $MgCl_2$–10 mM dithiothreitol–0.4 mM ATP. After the addition of 1 unit of T4 DNA ligase, the reaction mixture was incubated for 4 h at 20°C and then used to transform *E. coli* LE392 to ampicillin resistance. Plasmids were screened by restriction endonuclease cleavage site mapping, and in the case of pΔ1 and pΔ35, the deletion endpoints were determined by DNA sequence analysis (13).

pΔCAT1 and pΔCAT9 were isolated as follows: pTK173 (20 μg) was partially cleaved with *Eco*RI and then digested with 500 units of S1 nuclease in 0.5 ml of 200 mM NaCl–50 mM sodium acetate–1 mM $ZnSO_4$ (pH 4.5) at 20°C for 30 min. The solution was made to 100 mM Tris-hydrochloride (pH 8.0)–10 mM EDTA–0.1% sodium dodecyl sulfate, and the DNA was extracted with phenol and ether and then ethanol precipitated. Singly cleaved, unit-length DNA was purified by electrophoresis through a 1% agarose gel, recovered by electroelution, and circularized with ligase as described above. After transformation of HB101 to ampicillin resistance, three plasmids containing only a single *Eco*RI site were characterized. pΔCAT1 and pΔCAT9 both lacked the viral *Eco*RI site, whereas pAMP-2 lacked the pBR322 *Eco*RI site.

pΔ7 and pΔ17 were derived from pAMP-2 by cleaving 20 μg with *Eco*RI and *Bgl*II, followed by digestion with S1 nuclease and T4 DNA ligase as described above.

Transformation of mammalian cells. Cultures of LtA cells growing in 60-mm tissue culture dishes (5×10^5 cells per dish) were exposed to a calcium phosphate coprecipitate formed with 50 or 100 ng of circular plasmid and 5 μg of LtA carrier DNA exactly as described by Graham et al. (6). The precipitate (0.5 ml) was added directly to the growth medium (5 ml of α-minimal essential medium containing 10% fetal bovine serum) and left in contact with the cells at 37°C for 16 h. The medium was then replaced with fresh α-minimal essential medium. The medium was switched to α-HAT (α-minimal essential medium containing hypoxanthine, aminopterin, and thymidine) (28) 48 h after the initial exposure to the precipitate and replaced every 4 to 5 days thereafter. TK^+ colonies were scored after 2 weeks, and well-isolated clones were selected with a cloning cylinder. Cell lines were maintained in α-HAT.

TABLE 1. Deletion endpoints[a]

Mutant	Endpoint	
	5′	3′
p1EN1	−80	—
pΔ1	−11	+189
pΔ7[b]	−87	+99
pΔ17	−81	+59
pΔ35	−85	+85
pΔCAT1	−77	−72
pΔCAT9	−80	−70
pΔJS1	+56	—

[a] The numbering system is that of McKnight and Kingsbury (17).

[b] The precise endpoints of pΔ7 are ambiguous (see text for details).

Induction of TK activity in transformed cell lines. Transformed cell lines growing in 75-cm^2 tissue culture flasks were infected at the indicated multiplicities of infection with TK-deficient HSV-1 B2006 (3) and harvested 12 h later. Cell extracts were prepared and assayed for TK activity with [^{14}C]thymidine as described by Summers et al. (24), except that the DE81 papers were washed individually in scintillation vials with 30 mM ammonium formate three times then once with water and once with ethanol. Each wash was for 30 min. The TK-specific activities obtained after infection were normalized to that obtained for mock-infected cells, so that the results are expressed as the number of fold induction above the basal level.

Gel electrophoresis and Southern transfers. The indicated amounts of DNA were cleaved with a 20-fold excess of restriction enzyme, phenol and ether extracted, ethanol precipitated, and suspended in 10 mM Tris-hydrochloride–1 mM EDTA (pH 7.6). Electrophoresis was through a 1.4% agarose gel containing 40 mM Tris-acetate–2 mM EDTA (pH 8.3). Southern transfers to nitrocellulose (23) were carried out by the method of Wahl et al. (26).

RESULTS

Transformation with mutant TK genes. Several plasmid clones containing the HSV-1 TK gene were used as substrates for the isolation of deletion mutants as described above; their structures and those of the resulting deletion mutants are indicated in Fig. 1. The deletion endpoints, as determined by DNA sequencing (13), are summarized in Table 1. Note that the precise endpoints of Δ7 were ambiguous because the deletion arose by fusion of two nearly perfect 10-base direct repeats. Thus, although the 5′ endpoint of Δ7 is located between −97 and −94, the net effect is to largely preserve the wild-type sequence up to −88, with a single T→ C transition at −93.

To establish cell lines containing and expressing each mutant gene, supercoiled DNA from each mutant was used to transform TK-deficient LtA cells to TK^+ by using a modification (6) of the calcium phosphate coprecipitation technique of Graham and Van der Eb (7) (Table 2). We chose not to excise the viral inserts from the plasmid vector, hoping to minimize loss of viral sequences during integration into the host cell chromosomes. As has been reported previously

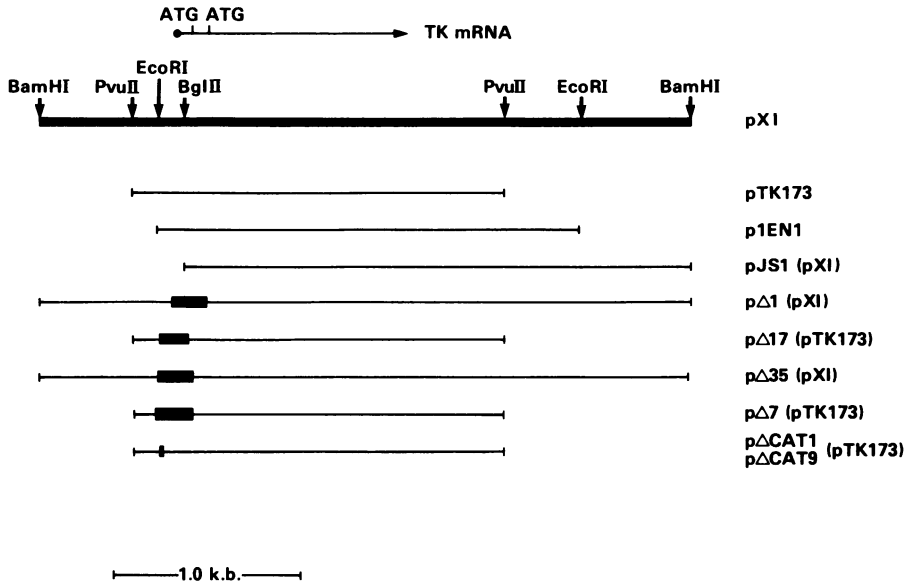

FIG. 1. TK sequences carried by plasmids used. The extents of the internal deletions carried by several plasmids are indicated by the raised filled regions. The location of TK mRNA coding sequences is from McKnight (14) and Wagner et al. (25). k.b., kilobases.

(2), the wild-type *Bam*HI and *Pvu*II fragments gave rise to transformants at roughly comparable frequencies (pX1 and pTK173, with 726 and 200 bases of 5'-flanking DNA, respectively). The various mutant genes fell into at least three classes with respect to their ability to transfer TK. Several mutants gave rise to TK^+ transformants at roughly the same frequency as the wild-type *Pvu*II fragment (Δ1, Δ7, ΔCAT1, and ΔCAT9). Several mutants transformed at a reduced efficiency (p1EN1, Δ17, and Δ35), whereas one mutant failed to produce any transformed colonies (pJS1). The results of such transformation assays, taken alone, are difficult to interpret in a straightforward way, because deletions can be expected to impair gene expression by a variety of mechanisms, including inactivation of promoters and deleterious effects on the stability or translational efficiency of mRNA. Furthermore, we observed variations in the relative transforming activities of plasmids between experiments. For these reasons we feel that trans-

TABLE 2. Transforming activity of plasmid DNA relative to pTK173[a]

Plasmid	Experiment no. with following amounts (ng) of circular plasmid:							
	1		2		3		4	
	50	100	50	100	50	100	50	100
pX1	0.47	1.07	0.37	0.76	—[b]	—	—	—
p1EN1	0.02	0.04	0.09	0.05	—	—	—	—
pJS1	0	0	—	—	—	—	—	—
pΔ1	—	—	0.54	0.64	—	—	0.64	0.40
pΔ7	—	—	0.88	1.36	—	—	0.45	0.66
pΔ17	—	—	0.40	0.40	—	—	0.09	0.07
pΔ35	—	—	0.13	0.23	—	—	0.002	0.01
pΔCAT1	—	—	—	—	0.30	0.85	—	—
pΔCAT9	—	—	—	—	1.11	1.44	—	—
pTK173	1581/5	1211/5	404/4	406/4	434/4	383/4	565/4	1147/4

[a] Quadruplicate cultures of 5×10^5 LtA cells were each exposed to a calcium-phosphate precipitate formed with 50 or 100 ng of circular plasmid and 5 μg of LtA carrier DNA exactly as described by Graham et al. (6). Medium was changed to HAT (28) 48 h later, and TK^+ colonies were scored after 2 weeks. pTK173 was included in each experiment, and the results are expressed relative to the number of colonies obtained with pTK173 (i.e., pTK173 = 1.0). In the case of pTK173, the data shown represent the total number of colonies observed over the indicated number of plates. No colonies were observed in the absence of plasmid DNA.

[b] —, Not done.

formation is a poor assay for promoter function; our primary aim was to establish cell lines containing each mutant gene. Nevertheless, several interesting points have emerged from the results which allow us to draw some tentative conclusions regarding the sequence requirements for basal level expression of TK in uninfected cells. First, as has also been reported previously (16), sequences upstream from the *Eco*RI site at −80 are required for efficient transformation (p1EN1 is impaired). McKnight et al. (16) reported that the 5′ boundary of sequences affecting transformation frequency is located between −109 and −95. Second, efficient transformation does not require the wild-type TK mRNA starting site nor does it require the presumed translation initiation site at +110. The mutant Δ1, which is deleted for both of these sequences, is not dramatically impaired compared with the wild-type in this assay. Therefore, the promoter sequences required for basal level expression in LtA cells appear to be located upstream from −11, again in keeping with the results obtained in *Xenopus* oocytes. As the cell lines transformed by Δ1 express wild-type levels of enzymatically active TK, although the mutant lacks the first potential ATG initiation codon for the TK polypeptide (+110), it would appear that either the amino-terminal portion of the protein is dispensable for catalytic activity or the physiological initiator is located further downstream (for example, at +245). Preston and McGeoch (21) have presented evidence that two translational initiators are used during lytic viral infections, and Roberts and Axel (22) have suggested that transcripts initiated at approximately +200 encode catalytically active TK.

The third conclusion that can be drawn from the transformation data is that the sequences immediately surrounding the *Eco*RI site at −80 are dispensable for efficient transformation. ΔCAT1 lacks 5 bases and ΔCAT9 lacks 10 bases centered at this site, and both transfer TK with reasonable efficiency. The deletion in ΔCAT9 removes all but two bases of a so-called "CAAT" consensus sequence (1) found in this region (GGCGAATTC). This result implies that this CAAT sequence does not form an indispensable part of the TK gene promoter, which is in agreement with the data of McKnight and Kingsbury (17), who have found that this CAAT sequence is also dispensable for expression in *Xenopus* oocytes.

We examined three deletion mutants which remove all of the sequences between the *Eco*RI site at −80 and the *Bgl*II site located in the region encoding the nontranslated leader of TK mRNA. The interpretation of the transformation data obtained with these mutants is especially difficult. Two of the mutants, Δ17 and Δ35, transformed at a variable efficiency, whereas the third, Δ7, appeared to transform almost as efficiently as the wild type. We therefore hesitate to draw conclusions about the sequence requirements for transformation based on these mutants. It is worth noting, however, that all three of these mutants lack two of the three upstream regions required for efficient expression in oocytes (15–17), yet at least one of them transforms cells with an efficiency comparable to that of the wild-type genes in some experiments.

DNA sequences present in transformants. Because the major aim of this study was to examine the effects of mutations on viral positive control of TK expression, it was necessary to demonstrate that each of the transformed cell lines contained at least one complete, unrearranged copy of the appropriate mutant gene. Accordingly, the structures of the TK sequences present in cell lines transformed by each deletion mutant were analyzed by the blotting technique of Southern (23). In each case, high molecular weight cellular DNA was cleaved with a combination of restriction endonucleases designed to generate diagnostic deletion-spanning fragments. An example of this analysis for the cell line Δ35/1 is shown in Fig. 2A. This line contains multiple copies of the appropriate deletion-bearing fragment extending from the upstream *Pvu*II site at −200 to the *Sma*I site at the 3′ end of the TK gene (Fig. 2A, lane 3). The viral sequences present in this line extend downstream beyond the 3′ *Eco*RI site, which is well outside the TK gene (Fig. 2A, lanes 6 and 8). Based on the signal intensity of the deletion-bearing *Pvu*II fragment with respect to the marker plasmid DNA, the line contains a minimum of 100 copies of unrearranged mutant TK DNA per diploid equivalent of cellular DNA. Δ35/1 cellular DNA also contains fragments corresponding to the plasmid vector in addition to viral TK sequences. *Pvu*II/*Sma*I and *Pvu*II/*Eco*RI double digests gave rise to fragments which comigrate with all of the fragments of the transforming plasmid (Fig. 2A, lanes 3 and 6). This could arise in one of two ways. Either the plasmid is integrated at many sites and the breakpoints on the plasmid differ at each insertion site, or the plasmid is integrated at fewer sites as large concatameric arrays. As neither *Hin*dIII nor *Sal*I digests gave rise to a strong band comigrating with linear unit-length plasmid (Fig. 2A, lanes 10 and 11, more clearly visible in shorter exposures), the former explanation appears to be correct.

Fourteen additional cell lines (Table 3) were analyzed in analogous experiments with comparable results; that is, each line contained at least one complete copy of the appropriate mutant gene (data not shown). During this analysis, we

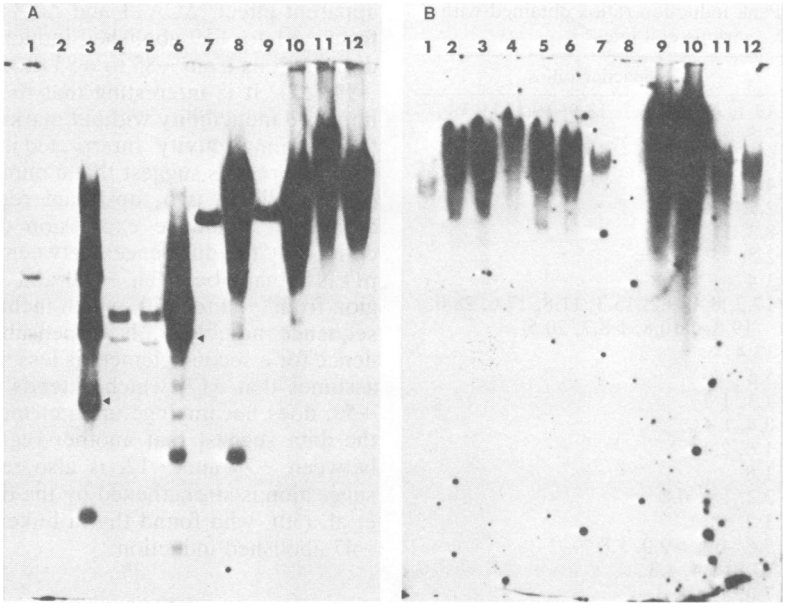

FIG. 2. Plasmid DNA sequences present in transformants. (A) Analysis of Δ35/1 cellular DNA. Five micrograms of Δ35/1 cellular DNA and 200 pg of marker plasmid DNA were cleaved to completion with the indicated restriction endonucleases, and the resulting fragments were separated by electrophoresis through a 1.4% agarose gel. After transfer to nitrocellulose (23), the fragments were hybridized to 2×10^7 dpm of nick-translated pX1 probe DNA (specific activity, 10^8 dpm/μg according to Wahl et al. (26). Lane 1, pX1 marker (*Pvu*II/*Sma*I); lane 2, pΔ35 marker (*Pvu*II/*Sma*I); lane 3, cellular DNA (*Pvu*II/*Sma*I); lane 4, pX1 marker (*Pvu*II/*Eco*RI); lane 5, pΔ35 marker (*Pvu*II/*Eco*RI); lane 6, cellular DNA (*Pvu*II/*Eco*RI); lane 7, pΔ35 marker (*Eco*RI); lane 8, cellular DNA (*Eco*RI); lane 9, pΔ35 marker (*Hin*dIII); lane 10, cellular DNA (*Hin*dIII); lane 11, cellular DNA (*Sal*I); lane 12, cellular DNA (*Xba*I). An 18-h exposure of the autoradiogram is shown. Deletion-bearing fragments are indicated by arrowheads in some lanes. (B) Comparison of copy numbers between cell lines. Ten micrograms of the indicated cellular DNAs were cleaved to completion with *Xba*I and analyzed as in (A). Lane 1, Δ17/1; lane 2, Δ7/3; lane 3, Δ7/1; lane 4, Δ1/1; lane 5, p1EN1/B; lane 6, p1EN1/A; lane 7, pX1/B; lane 8, pX1/A; lane 9, Δ35/1; lane 10, Δ35/2; lane 11, ΔCAT9/1; lane 12, ΔCAT1/1. The autoradiogram was exposed for 1 week.

noticed substantial differences in the amount of transforming DNA present in the various cell lines. To roughly compare the relative copy numbers of the transforming DNA in some of these cell lines, we compared the hybridization signal intensities in *Xba*I digests of cellular DNA (Fig. 2B). The results indicated that the copy number ranged from about 100 copies for Δ35/1 and Δ35/2 to about one for pX1/A. In contrast to the results of Zipser et al. (30), we did not observe an obvious negative correlation between the transforming efficiencies of plasmids and the copy number in the resulting transformants. For example, Δ35 and Δ17 transformed LtA cells at roughly the same efficiency but gave rise to cell lines with very different copy numbers. We have not yet examined enough cell lines to determine whether each plasmid consistently generates cell lines with a characteristic copy number.

Control of mutant gene expression. To assess the effects of the mutations on the response of the TK gene to viral positive regulators, cells transformed by the various plasmids were superinfected with TK-deficient HSV-1 B2006 (3), and the levels of TK enzymatic activity were determined 12 h later. Control experiments demonstrated that 12 h postinfection corresponds to the peak of the induction response for the cell line pX1/A (data not shown). Figure 3 shows the induction of TK activity as a function of the multiplicity of superinfecting virus for a number of cell lines. The peak induction values obtained in comparable experiments with other cell lines are listed in Table 3. Superinfection resulted in a marked increase in TK activity in cell lines transformed by several of the plasmids. As others have reported (18), we noticed some variation in the peak induction values obtained in separate experiments with the same cell lines, and we are at present unable to account for these differences. Nevertheless, the various plasmids clearly fell into two classes, those which gave rise to transformants which induced as well as cells containing the wild-type gene (induction ratios of greater than five), and those which

TABLE 3. Peak induction ratios obtained with various cell lines[a]

Cell line	Induction ratios
pX1/A	17.1, 43.4, >6.7, 13.9, 19.1, 11.3
pX1/B	7.5
pTK173/1[b]	8.2, 7.5
pTK173/6[b]	4.0, 2.2
pTK173/7[b]	2.7, 3.3
pTK173/8[b]	8.3, 11.8
p1EN1/A[b]	2.5, 1.7
p1EN1/B[b]	1.4
Δ1/1[b]	17.3, 8.4, 9.2, 15.3, 11.8, 17.6, 28.0, 19.8, >10.8, >8.7, 20.5
Δ1/2	13.4
Δ1/3	5.6
Δ7/1	1.1, 1.5
Δ7/3[b]	0.8, 1.4
Δ17/1[b]	1.6, 2.4
Δ17/2[b]	1.8
Δ35/1[b]	2.2, 1.4, 0.8
Δ35/2[b]	1.2
ΔCAT1/1[b]	5.6, 4.9, >9.0, 5.8
ΔCAT1/2	13.8, 6.4, 4.3
ΔCAT1/3	7.0, 3.0
ΔCAT1/4	7.7, 4.4
ΔCAT9/1	6.4, >10.7, 7.2
ΔCAT9/2	>12.0, 11.3
ΔCAT9/3	10.5, 9.8

[a] Cell lines were superinfected with HSV-1 B2006 at 0.5, 1.0, 2.0, and 5.0 PFU per cell, and TK activity was measured 12 h postinfection. The induction ratio is the TK specific activity in infected cells/TK specific activity in uninfected cells. Each value gives the peak ratio obtained in a separate experiment. The basal level of TK activity in uninfected cells was similar between cell lines, varying no more than twofold.

[b] Cell lines were examined by Southern blotting (see text for details).

produced only transformants which induced poorly or not at all (induction ratios of two or less). By correlating the extents of the deletions in the transforming genes with the inducibility of the resulting cell lines, it should be possible to assemble a map of the sequences required for induction. First, cells transformed by both the *Bam*HI (pX1) and *Pvu*II (pTK173) TK fragments induced well after superinfection, suggesting that, at most, 200 bases of 5′-flanking viral DNA are required, in accord with the results of Wilkie et al. (29). Furthermore, because Δ1 transformants were also inducible, the sequences from −11 to +189 are dispensable. Therefore, any upstream induction sequences map entirely between −200 and −12 and apparently are capable of regulating expression from novel transcription initiation sites (Δ1 must necessarily initiate at a novel site). Deletion of sequences 5′ to −80 severely impaired the inducibility of TK expression (p1EN1), suggesting that sequences located somewhere between −200 and −80 are required, whereas deletions from −80 to −70 had no apparent effect (ΔCAT1 and ΔCAT9). Deletion from −81 to +59 abolished induction (Δ17), as did deletions from −85 to +85 (Δ35) and −87 to +99 (Δ7). It is interesting that the Δ7 deletion impaired inducibility without markedly affecting transforming activity. Interpreted in the simplest way, the results suggest that a minimum of one, or more likely two, upstream regions are required for inducible expression of TK. One, defined by the differences between pTK173 and p1EN1, maps between −200 and −80. The region from −80 to −70, which includes a CAAT sequence homology, is dispensable. The evidence for a second element is less secure. If one assumes that Δ17, which extends from −81 to +59, does not impinge upon element one, then the data suggest that another region, mapping between −70 and −12, is also required. This suggestion is strengthened by the data of Zipser et al. (30), who found that a linker insertion at −47 abolished induction.

DISCUSSION

A central question bearing on the interpretation of the results presented here is the mecha-

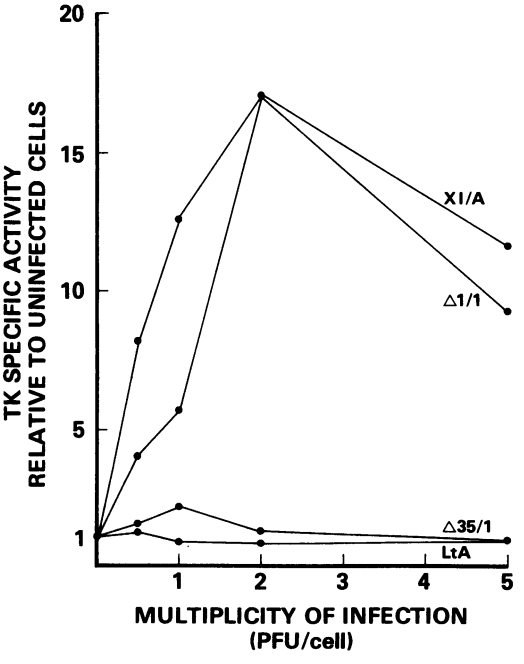

FIG. 3. *trans*-Induction of TK activity by Superinfecting HSV-1. Approximately 2×10^7 cells were infected at the indicated multiplicities with TK-deficient HSV-1 B2006. Twelve hours later the cells were harvested, and the extracts were assayed for TK enzymatic activity (see text). The results are expressed as the number of fold enhancement above the basal level in uninfected cells.

nism by which superinfecting HSV activates the expression of the TK gene present in transformed cells. We believe that this activation results from an increased frequency of transcription initiation events effected by IE proteins. The evidence which supports this belief is, first, that the increase in TK enzymatic activity after superinfection with HSV is paralleled by a comparable increase in the concentration of TK mRNA transcribed from the transforming gene (A. El Kareh and S. J. Silverstein, personal communication; Wagner, Summers, Sharp, Smiley, and Summers, submitted for publication) and, second, that the 5′ boundary of the sequences necessary for this enhancement map upstream from the transcription initiation site (as observed in this study). We therefore believe that post-transcriptional events such as RNA processing and transport are unlikely to play a major role in induction.

The TK genes that we analyzed fall into two classes with respect to induction by superinfecting virus, those which induce normally and those which are impaired. The interpretation of mutants which are inducible is straightforward: they must retain all of the necessary sequences. Thus, we conclude that any upstream sequences involved in induction map between −200 and −12 (pTK173 versus pΔ1). Furthermore, the region from −80 to −70 is dispensable (pΔCAT1 and pΔCAT9). The interpretation of the phenotypes of the remaining noninducible mutants is more complicated, because deletions could inactivate induction either by removing an essential sequence element or by altering critical spacing arrangements between elements which themselves are not directly affected by the deletion. The latter possibility is more of a concern when considering internal deletions (pΔ1, pΔ7, pΔ17, and pΔ35) than when considering external deletions (e.g., p1EN1). As p1EN1, which lacks only sequences upstream from −80, is impaired, we conclude that induction requires sequences located between −200 and −80. We believe that the phenotypes of the remaining mutants cannot be accounted for solely by alterations in the spacing of otherwise intact elements for the following reason. The wild-type induction phenotype of Δ1 demonstrates that the relative spacing of elements located upstream and downstream from the Δ1 endpoints is not critical for induction. As the Δ7, Δ17, and Δ35 deletions all have their 3′ endpoints within the Δ1 deletion, we believe that these mutants are impaired because their 5′ endpoints extend further upstream than that of Δ1, which removes one or more essential sequence elements. The alternative, which is that they leave all essential elements intact but alter their spacing, is unlikely because the spacing of these (putative) upstream and downstream elements is also altered by Δ1 without effect.

If one accepts that Δ7, Δ17, and Δ35 inactivate induction by eliminating essential sequences, then the data suggest that these sequences map between −81 and −12. Since residues −80 to −70 are dispensable, the relevant sequences are probably located between −70 and −12. This conclusion agrees with the finding of Zipser et al. (30) that a linker insertion at −47 abolishes induction. In summary, the data presented here, taken with that of Zipser et al. (30) suggest that induction requires the integrity of at least two separate upstream DNA sequences, one mapping between −200 and −80 and the other mapping between −70 and −12. Recently, the 5′ boundary of the most distal of these two regions necessary for induction has been mapped more precisely to between −109 and −95 by using the deletion mutants constructed by McKnight et al. (16; El Kareh and Silverstein, personal communication).

We believe that it is likely that each of these regions may contain several functionally distinct sequences whose role in induction remains to be clarified. In particular, the region between −70 and −12 encompasses two sequences found to be important for constitutive expression in *Xenopus* oocytes. This includes a guanosine-rich region extending from about −60 to −47 and the TATA homology at −28 (16, 17). Figure 4 shows the nucleotide sequence of the upstream region which appears to be necessary for inducible expression of TK and also indicates the regions identified by McKnight and Kingsbury to be necessary for basal level expression in oocytes (17). Although the induction sequences have not yet been mapped as precisely as those required for basal level expression, it is interesting that these two sets of sequences overlap at least partially. In particular, the region from −80 to −70 is dispensable for both phenomena, whereas sequences upstream and downstream from −80 are required. However, the results of Zipser et al. (30) demonstrate that inducible and basal level expression are dissociable by mutation, implying the existence of sequences specific for each process.

A surprising feature of our results and those of McKnight and Kingsbury (17) is the dispensability of the CAAT sequence homology extending from −82 to −75. McKnight and Kingsbury demonstrated that this sequence is not required for expression in oocytes (17), and we find that it is not required for either basal level or inducible expression in mouse cells. On the surface, therefore, it would appear that this CAAT sequence does not play a major role in either constitutive expression of TK or in the positive control of its expression. We wish to point out, however, that

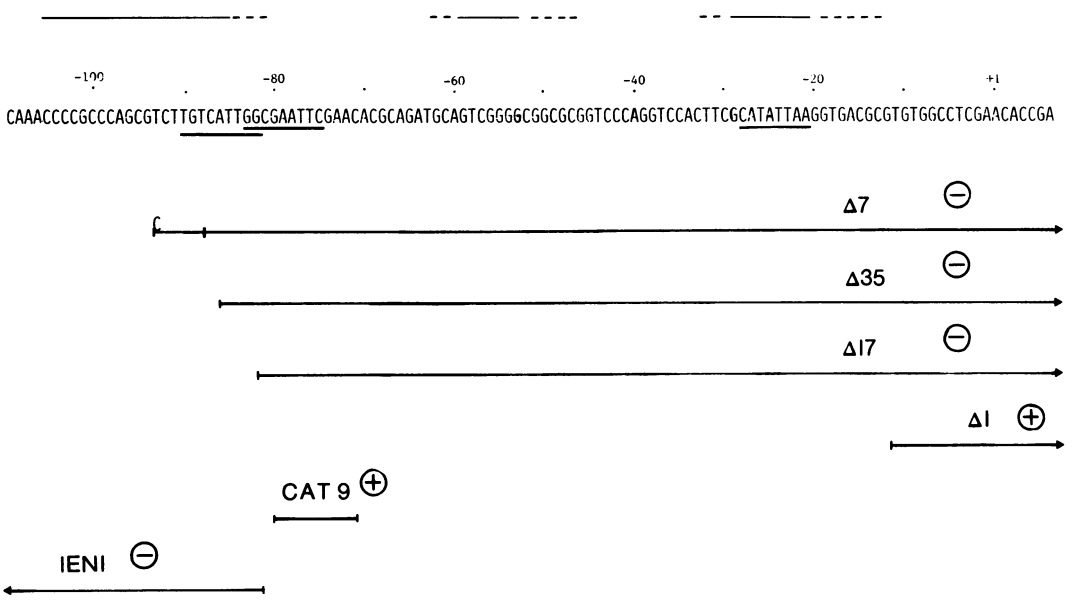

FIG. 4. Nucleotide sequences required for positive control. The nucleotide sequence for the TK gene extending from −109 to +7 is shown (14, 16). The locations of the CAAT and TATA homologies are underscored, and regions found by McKnight and Kingsbury (17) to be required for basal level expression in oocytes are indicated by the lines above the sequence. Uncertainties about the precise boundaries of these regions are indicated by the broken portions of the lines. The sequences absent from the deletion mutants described in the text are shown. Symbols: +, those mutants which are inducible by superinfecting virus; −, those mutants which are not inducible by superinfecting virus. The ambiguity in assigning the precise endpoints for Δ7 can be seen.

a second partially overlapping sequence extending from −92 to −82 (TGTCATTGG) matches the literature consensus sequence (GG$^{C}_{T}$CA$^{A}_{T}$CT) as well as the −82 to −75 sequence (Fig. 4). This second sequence remains intact in our ΔCAT1 and ΔCAT9 mutants. Perhaps only one functional CAAT sequence is required. In this regard, it is interesting that mutations that extend from −84 to −74, destroying both CAAT sequences, reduce expression in oocytes (17). We believe, therefore, that the CAAT homology may prove to be crucial for the control of gene expression and that its reiteration in the TK control region reflects its importance. This suggestion is in agreement with the recent finding that the CAAT sequence is essential for the expression of the rabbit β-globin gene in vivo (8).

As noted above, the boundaries of the region that we have identified as being essential for induction of TK expression by HSV IE proteins overlap with those of the constitutive TK promoter identified by McKnight and coworkers (14–17) in *Xenopus* oocytes. Although our results leave open the possibility that sequences upstream from the most distal constitutive promoter element are also involved in induction, the data of El Kareh and Silverstein (personal communication), which demonstrate that sequences upstream from −109 are not required, make this unlikely. The most straightforward interpretation of the available data is that viral IE proteins, or their metabolic product, interact with DNA sequences in the promoter region to activate the TK transcription. McKnight and Kingsbury (17) have hypothesized that constitutive expression in oocytes depends on base pairing between a CCGCCC sequence (−103 to −98) with GGGCGG (−55 to −49), which exposes an RNA polymerase binding site within the intervening melted region. Polymerase, then, is directed to an initiation site by the TATA sequence. This is an appealing idea, as the exposed regions would include the CAAT sequences which we suggest may facilitate the initial interaction of polymerase with the DNA. If this general scheme is correct, then perhaps ICP4 induces expression by an analogous mechanism. By binding to sequences surrounding the CAAT region, directly or complexed with a cellular protein (5), ICP4 may facilitate loading of polymerase, either by stabilizing a similar cruciform structure or by allowing polymerase to bind to fully duplexed DNA. By either model, ICP4 is postulated to play a direct, structural role in transcription. This is in keeping with the finding that it is required continuously throughout lytic infections to maintain transcription of delayed early and late HSV genes (27).

Alternatively, it is conceivable that the sequences we have identified are required for induction, not because they contain recognition sites for viral positive regulators, but only because they constitute the promoter site. By this model, the recognition site for the regulatory signal is located downstream, within the TK coding sequence, rather than around the promoter site, but a functional promoter is required for the response to the regulatory signal. Although we have no direct evidence to exclude this possibility, we consider it to be very unlikely for several reasons. First, Zipser et al. (30) have demonstrated that some mutations in the upstream region destroy the constitutive promoter function without affecting induction and, conversely, that other mutations destroy inducibility without affecting constitutive expression. Second, Post et al. (19) have demonstrated that TK is converted to an immediate early gene after fusion of the TK coding sequence to an IE promoter, demonstrating that for IE HSV genes, temporal control signals reside in the upstream region. Third, the putative downstream control element would have to be largely indifferent to its spacing relative to the promoter site, as Δ1 remains inducible. Although we do not yet know where transcription initiates in cells transformed by noninducible mutants, it seems likely that secondary internal promoter sites are used, as demonstrated by Roberts and Axel (22). One would then expect that the postulated intragenic control element would act to induce expression from these sites, a result that is not obtained. For these reasons we favor the first explanation, which is that the inducing signal acts upon sequences located in the promoter region. A definitive demonstration that this is the case will require showing that the region that we have identified is not only necessary but also sufficient for induction by superinfecting virus.

ACKNOWLEDGMENTS

We thank W. C. Leung, R. McKinnon, and F. L. Graham for gifts of plasmids and R. J. Redfield for assistance in the initial phases of the study.

This work was supported by the Medical Research Council and the National Cancer Institute of Canada. J.R.S. is a research scholar of the National Cancer Institute of Canada.

LITERATURE CITED

1. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
2. **Colbere-Garapin, F., S. Chousterman, F. Horodniceau, P. Kourilsky, and A. Garapin.** 1979. Cloning of the active thymidine kinase gene of herpes simplex virus type 1 in *E. coli* K12. Proc. Natl. Acad. Sci. U.S.A. **76**:3755–3759.
3. **Dubbs, D. R., and S. Kit.** 1964. Mutant strains of herpes simplex deficient in thymidine kinase-inducing activity. Virology **22**:493–502.
4. **Enquist, L. W., G. F. Van de Woude, M. J. Wagner, J. R. Smiley, and W. C. Summers.** 1979. Construction and characterization of a recombinant plasmid encoding the gene for the thymidine kinase of herpes simplex type 1 virus. Gene **7**:335–342.
5. **Freeman, M. J., and K. L. Powell.** 1982. DNA-binding properties of a herpes simplex virus immediate early protein. J. Virol. **44**:1084–1087.
6. **Graham, F. L., S. Bacchetti, and R. McKinnon.** 1980. Transformation of mammalian cells with DNA using the calcium technique, p. 3–32. *In* Introduction of macromolecules into viable mammalian cell. Alan R. Liss, Inc., New York.
7. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of the infectivity of adenovirus 5 DNA. Virology **52**:456–467.
8. **Grosveld, G. C., A. Rosenthal, and R. A. Flavell.** 1982. Sequence requirements for the transcription of the rabbit β-globin gene *in vivo*: the −80 region. Nucleic Acids Res. **10**:4951–4971.
9. **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.
10. **Leiden, J. M., R. Buttyan, and P. G. Spear.** 1976. Herpes simplex virus gene expression in transformed cells. I. Regulation of the viral thymidine kinase gene in transformed L cells by products of superinfecting virus. J. Virol. **20**:413–424.
11. **Leung, W.-C., K. Dimock, J. R. Smiley, and S. Bacchetti.** 1980. Herpes simplex virus thymidine kinase transcripts are absent from both nucleus and cytoplasm during infection in the presence of cycloheximide. J. Virol. **36**:361–365.
12. **Lin, S.-S., and W. Munyon.** 1974. Expression of the viral thymidine kinase gene in herpes simplex virus-transformed cells. J. Virol. **14**:1199–1208.
13. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end labelled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
14. **McKnight, S. L.** 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acids Res. **8**:5949–5964.
15. **McKnight, S. L.** 1982. Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. Cell **31**:355–365.
16. **McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel.** 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell **25**:385–398.
17. **McKnight, S. L., and R. Kingsbury.** 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**:316–324.
18. **Minson, A., S. E. Bell, and K. Bostow.** 1982. Correlation of the virus sequence content and biological properties of cells carrying the herpes simplex virus type 2 thymidine kinase gene. J. Gen. Virol. **58**:127–138.
19. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of α genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. Cell **24**:555–565.
20. **Preston, C. M.** 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant *tsK*. J. Virol. **29**:275–284.
21. **Preston, C. M., and D. J. McGeoch.** 1981. Identification and mapping of two polypeptides encoded within the herpes simplex virus type 1 thymidine kinase gene sequences. J. Virol. **38**:593–605.
22. **Roberts, J. M., and R. Axel.** 1982. Gene amplification and gene correction in somatic cells. Cell **29**:109–119.
23. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
24. **Summers, W. P., M. Wagner, and W. C. Summers.** 1975. Possible peptide chain termination mutants in thymidine kinase gene of a mammalian virus, herpes simplex virus. Proc. Natl. Acad. Sci. U.S.A. **72**:4081–4084.

25. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. U.S.A. **78**:1441–1445.
26. **Wahl, G. M., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. U.S.A. **76**:3683–3687.
27. **Watson, R. J., and J. B. Clements.** 1980. A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. Nature (London) **285**:329–330.
28. **Wigler, M., S. Silverstein, L.-S. Lee, A. Pellicer, Y.-C. Cheng, and R. Axel.** 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell **11**:223–232.
29. **Wilkie, N. B., J. B. Clements, W. Boll, N. Mantei, D. Lonsdale, and C. Waissman.** 1979. Hybrid plasmids containing an active thymidine kinase gene of herpes simplex virus—1. Nucleic Acids Res. **7**:859–877.
30. **Zipser, D., L. Lipsich, and J. Kwoh.** 1981. Mapping functional domains in the promoter region of the herpes thymidine kinase gene. Proc. Natl. Acad. Sci. U.S.A. **78**:6276–6280.