

Analysis of DNA Sequences Which Regulate the Transcription of a Herpes Simplex Virus Immediate Early Gene

CHRIS M. PRESTON,* MIKE G. CORDINGLEY,† AND NIGEL D. STOW

Medical Research Council Virology Unit, Glasgow G11 5JR, Scotland

Received 1 December 1983/Accepted 21 February 1984

The locations and functions of DNA sequences involved in transcription of the gene encoding herpes simplex virus type 1 immediate early (IE) mRNAs 4 and 5 were analyzed by use of a transient-expression assay. The region upstream of the genes encoding IE mRNAs 4 and 5 was fused to the thymidine kinase gene coding sequences, and production of enzyme or RNA was measured after transfection of plasmids into BHK cells. The effect of deletions in the upstream region was determined in the absence or presence of a virus structural component which stimulates herpes simplex virus IE transcription. Two distinct units were identified. One of these was a promoter which required not more than 69 base pairs of DNA specific for the genes encoding IE mRNAs 4 and 5 upstream from the mRNA 5' terminus. The other was a far-upstream region which mediated the response to the virion component and had an upstream boundary between nucleotides -347 and -335. An origin of DNA replication was interposed between these two units. The element TAATGAGATAC, which represents a consensus sequence present in the upstream regions of all herpes simplex virus type 1 IE genes, appeared to be essential for stimulation by the virion component. The activity of this element was modulated by the sequences which flank it, especially by regions having extremely high contents of guanine plus cytosine and which contain a conserved unit CCCGCC or its complement GGGCGG.

Transcription of the herpes simplex virus (HSV) type 1 (HSV-1) genome can be divided into three temporally distinct phases named immediate early (IE), early, and late (5, 11, 33). IE transcription occurs shortly after infection and is rapidly repressed when early RNA synthesis commences. If the transition from IE to early and late phases is prevented, large amounts of IE mRNAs accumulate, showing that IE genes can be transcribed very efficiently under certain circumstances (1, 9, 35).

Figure 1A shows the genome locations for the five major IE mRNAs. The 5' termini of IE mRNAs 4 and 5 map at equivalent loci within the internal and terminal inverted repeats and consequently are adjacent to identical promoter sequences (28, 36). The experiments described here deal with these common promoter sequences, and therefore the conclusions apply equally to both genes. For convenience, they will be referred to as IE gene 4/5. This region of the genome also contains the 5' terminus of IE mRNA 3 and an origin of DNA replication (ORIs; see Fig. 1B) (4, 12, 18, 27, 30, 31), and one aim of the experiments reported here was to define the relationship of promoter and regulatory sequences of IE gene 4/5 to previously identified elements with known functions.

IE gene expression is known to be controlled in at least two ways: it is repressed by an autoregulatory mechanism which involves the IE gene 3 product, Vmw175, and is stimulated by a component of the HSV particle (7, 8, 22, 23, 26, 34). The stimulatory effect has been observed by the use of chimeric plasmids containing IE gene promoter sequences linked to the HSV-1 thymidine kinase (TK) structural sequences. When such plasmids are introduced into tissue culture cells, production of TK is strongly stimulated upon

infection with HSV, even in situations where no viral polypeptides are synthesized (2, 7). An increase in mRNA which is specific to the IE-TK chimeric gene can also be detected, indicating that the effect represents stimulation of transcription (7).

Attempts to locate the IE-specific sequences which respond to the virion component have shown that far-upstream regions (at least 110 base pairs [bp] and possibly up to 730 bp from the mRNA 5' terminus) are involved (7, 13, 14). The most extensively studied example is IE gene 3. In this case, a region between nucleotide positions -174 and -331 (relative to the mRNA 5' terminus) contains the target for stimulation by the virion component (7, 13). This far-upstream region is distinct from an essential promoter element located between -50 and -108. It shows some functional similarity to enhancer sequences, since it stimulates expression of chimeric IE-TK genes even in the absence of the virion component (7).

Since IE genes 1, 2, 3, and, as shown here, 4 and 5 all respond to stimulation by the virion component (14), it may be expected that common sequence elements exist in their far-upstream regions. Comparison of such regions revealed a number of consensus sequences, the most obvious being 5'-TAATGARATTC (R = purine) and 5'-GGGCGGG (or its complement 5'-CCCCGCC). These are usually represented in multiple copies with various extents of homology with the consensus (14, 19, 37, 39). Other, less stringent, guanine-plus-cytosine (G+C)-rich, conserved sequences can also be detected (C. Preston, unpublished data). To determine the functional importance, if any, of these consensus sequences, we have constructed plasmid deletion mutants which remove specific components of the far-upstream region of IE gene 4/5. The altered upstream regions have been linked to the HSV TK structural gene, so that IE gene 4/5-specific transcription and its regulation could be measured in a transient expression assay. The results show that transcription of IE gene 4/5 is controlled in the same

* Corresponding author.

† Present address: Laboratory of Tumor Virus Genetics, National Institutes of Health, Bethesda, MD 20205.

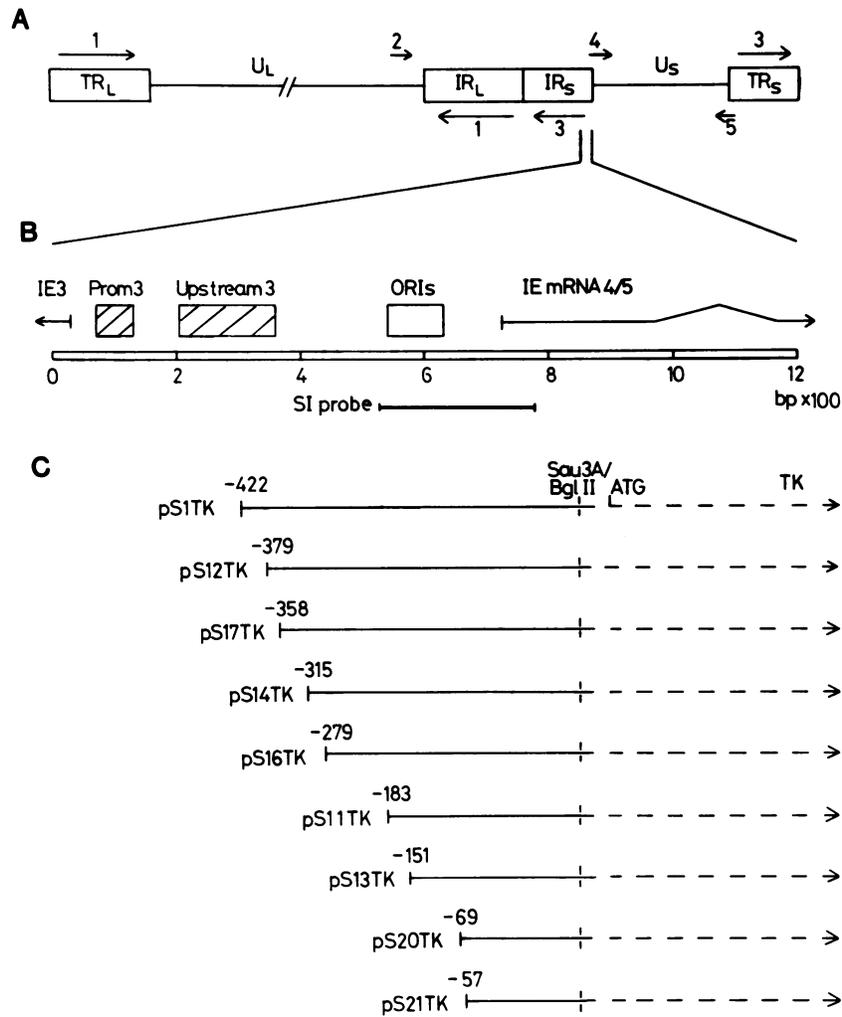


FIG. 1. (A) Genome location of five IE mRNAs. (B) Expanded scale of the 1,200 bp which encompass the initiation sites for IE mRNAs 3 and 4/5. The location of the *Ava*I fragment used for nuclease S1 mapping is also shown (S1 probe). Previously identified promoter (Prom), far-upstream regions of IE gene 3, and ORIs are shown above the line. (C) Structures of pSTK plasmid deletions; detailed derivations are described in the text. The solid lines represent IE gene 4/5 upstream sequences; the broken lines represent TK gene sequences. The *Sau*3A/*Bgl*II hybrid site (*Sau*3A/*Bgl*II) and the ATG initiation codon for TK are also shown.

way as the other IE genes and indicate an essential regulatory role for a sequence which closely resembles the TAATGARATTC consensus. A modulatory effect of G+C-rich flanking sequences was also demonstrated.

MATERIALS AND METHODS

Cells and virus. BHK-21 clone 13 cells (15) were used throughout. The HSV-1 mutant *tsK* (16, 25) was inactivated by irradiation with UV light as described previously (21) to reduce the titer by a factor of 5×10^5 .

Transfection of cells. Plasmid DNAs were transfected into cells as calcium phosphate precipitates. Precipitates were formed as described by Shen et al. (29), by using 2 μ g of plasmid DNA and 4 μ g of calf thymus DNA per petri dish (diameter, 35 mm) containing 10^6 cells. Cell monolayers were overlaid with the calcium phosphate precipitate and incubated at 37°C for 40 min. Plating medium (2 ml), prepared as described by Shen et al. (29) but with Eagle medium instead of Dulbecco medium, was added to cells, and incubation was continued at 37°C. After 3 h, cell

monolayers were treated with 25% dimethyl sulfoxide, as described by Stow and Wilkie (32), and incubated a further 1 h at 38.5°C. Monolayers were then either mock infected or treated with an amount of *tsK* mutant which was equivalent to 3 PFU of unirradiated virus per cell. Cells were incubated at 38.5°C for a further 16 h, after which time extracts were prepared either for TK assays or for isolation of cytoplasmic RNA. Increasing the plasmid DNA to 4 μ g and decreasing the calf thymus DNA to 2 μ g gave a twofold increase in the production of TK, showing that a linear relationship between plasmid DNA and enzyme production existed.

TK assays. Cytoplasmic extracts were prepared, and TK assays were performed as described previously (7). Both the quantity of extract and time of incubation were altered to ensure a linear relationship between phosphorylation of tritiated thymidine and these two variables. Values which fell outside the linear range were not used. To standardize the results, plasmid pS1TK was included in all experiments as a positive control. The average TK expression by this plasmid from 15 independent determinations was calculated,

and the actual value for pS1TK in a given experiment was corrected to the average. The activities of the other plasmids tested in a given experiment were adjusted by the same factor.

Nuclease S1 mapping. Cytoplasmic RNA was extracted as described previously (24). A 5- μ g amount of the RNA was hybridized with a 5' terminally labeled *Ava*II fragment derived from plasmid pGX35, exactly as described by Rixon and Clements (28). The genome location of this fragment is indicated in Fig. 1B.

Plasmid deletions. The two sets of plasmid deletions (the pSTK and pTKN series) used in the experiments reported here were prepared by standard cloning procedures, and their structures are shown in Fig. 1C and 5. All deletion endpoints were determined by DNA sequencing.

The construction of the plasmids is shown in Fig. 2. The pSTK series was derived from the pS plasmids as described by Stow and McMonagle (31). The starting point was pS1, a *Sau*3AI fragment representing sequences from -422 to +114 with respect to the 5' terminus of IE gene 4/5, cloned as a *Bam*HI fragment in pAT153. BAL 31 deletions were made from the *Sal*I site of pAT153, and a *Sal*I linker was inserted at deletion endpoints. The small *Bam*HI-*Sal*I fragment was recloned into pAT153. The appropriate *Bgl*II-*Eco*RI fragment from pTK1 containing TK coding sequences (40) was cloned between the unique *Bam*HI and *Eco*RI sites of the pS plasmids to yield pS12TK, pS17TK, pS14TK, pS16TK, pS11TK, and pS13TK. Plasmid pS1TK contained the entire *Sau*3AI fragment, whereas pS20TK and pS21TK were made by inserting a *Sal*I linker at *Sma*I sites -69 and -57, respectively, from the mRNA 5' terminus in pS12TK, followed by *Sal*I cleavage and religation.

The starting point for the pTKN series was pN1, which is HSV-1 *Bam*HI cloned into pAT153 from which the *Eco*RI site had been deleted. Exonuclease III was used to generate deletions from the single *Eco*RI site in *Bam*HI, to give the pN plasmids as described previously (7). A *Sal*I linker was inserted into the *Sma*I sites at -362 or -299 (see Fig. 4), and a TK-containing fragment was ligated to this site via a *Sal*I linker at the -299 *Sma*I site in pS12TK or the *Sal*I linker at the deletion endpoint of pS14TK, pS16TK, or pS13TK. Plasmids pTKN7 and pTKN8 therefore contained a *Sal*I linker at the -299 *Sma*I site. Plasmid pTKN9 was made by subcloning a *Sau*3AI-*Eco*RI fragment of pN9 (endpoint at -335) into pAT153 and subsequent insertion of the large *Bgl*II-*Bam*HI fragment of pTK1 into the *Bam*HI site generated by the *Sau*3AI-*Bam*HI fusion. Therefore, pTKN9 did not have a *Sal*I linker at -299. Plasmid pTKN12 represented a fusion of the -362 *Sma*I site to the deletion endpoint of pS14TK.

RESULTS

Upstream boundaries for promoter and regulatory sequences of IE gene 4/5. To identify DNA sequences important for transcription of IE gene 4/5, chimeric plasmids were constructed in which the IE mRNA 5' terminus and upstream regions were fused to TK structural sequences. The *Bgl*II site located 54 bp upstream from the first ATG of TK was ligated to a *Sau*3AI site in the 5' noncoding region of IE mRNA 4/5. Production of TK was therefore controlled by IE gene 4/5 promoter and regulatory sequences, but the transcript did not contain the intron which is normally present in the 5' noncoding region of IE mRNA 4/5 (28, 36).

Expression of the chimeric plasmids was measured by

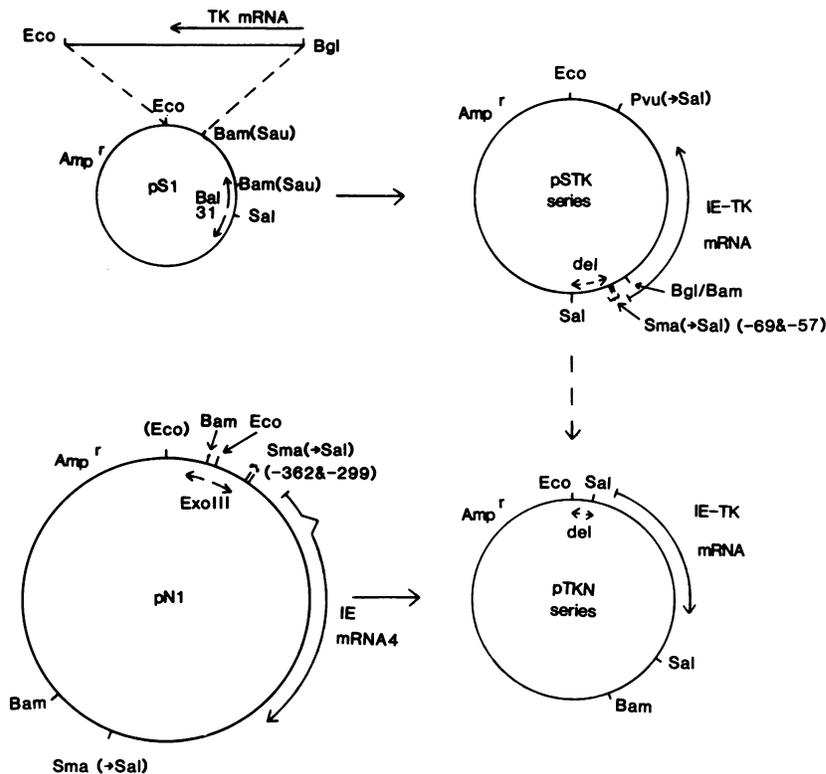


FIG. 2. Derivation of pSTK and pTKN plasmids, as described in the text. The gross positions of the deletions are shown (del).

transient expression of TK after transfection of duplicate sets of BHK cell monolayers, by the calcium phosphate precipitation technique of Shen et al. (29). One set of monolayers was mock infected, and another was treated with UV-inactivated *tsK* mutant. After incubation at 38.5°C for 16 h, TK activity in cell lysates was determined. Mock-infected cells gave an estimate of unstimulated expression of the chimeric plasmids, whereas the ratio of mutant *tsK*-treated to mock-infected cell levels defined the degree of response to the virion component.

Figures 1C and 2 depict a group of plasmids which were based on deletions previously used to locate ORIs (31). These plasmids sequentially remove DNA sequences from -422 to -57 relative to the mRNA 5' terminus. When unstimulated expression was assayed, all plasmids directed similar levels of TK, except pS21TK, which was approximately threefold less active (Table 1). This experiment shows that sequences located not more than 69 bp upstream from the mRNA 5' terminus are sufficient for transcription and suggests an important role for the region between -69 and -57, which contains the motif GGCGGG. The element located within 69 bp of the mRNA 5' terminus is referred to as the promoter.

When cells transfected with pS1TK were treated with inactivated *tsK* mutant, TK production was stimulated approximately eightfold (Table 1); values from individual experiments varied between sixfold and twelvefold. This result suggests that production of IE mRNAs 4 and 5, like that of other IE mRNAs, is increased by a virion component. Expression of plasmid pS12TK was stimulated to the same extent as pS1TK, but the degree of response with pS17TK was reduced to an average value of five- to sixfold. Although this reduction is not significantly greater than the variation in degree of stimulation between experiments, it was consistently observed within individual experiments. When the deletion extended to -315, as in pS14TK, or further, no effect of inactivated *tsK* was detected, thereby localizing an upstream boundary between -358 and -315 for sequences which are essential for the response to the virion component.

Synthesis of IE gene 4/5-specific RNA in transfected cells. To investigate whether stimulation of TK levels represented an increase in correctly initiated IE gene 4/5-specific transcripts, cytoplasmic RNA from transfected cells was analyzed by S1 nuclease mapping (Fig. 3). In mock-infected cells transfected with pS12TK or pS14TK, small amounts of 48- and 46-nucleotide bands were observed (lanes 5 and 6). These bands are equivalent to those representing the 5' termini of IE mRNA 4/5, shown in lane 1 (28). In mutant *tsK*-treated cells, even cultures transfected with calf thymus

TABLE 1. Transient expression of TK by pSTK plasmids

Plasmid	TK activity (cpm/min of assay per μ g of protein per μ g of plasmid in:		Degree of stimulation (-fold)
	<i>tsK</i> -treated cells	Mock-infected cells	
pS1TK	911	116	7.9
pS12TK	920	115	8.0
pS17TK	667	107	5.6
pS14TK	112	107	0.9
pS16TK	105	117	1.0
pS11TK	126	111	1.1
pS13TK	110	113	1.1
pS20TK	101	98	1.0
pS21TK	38	44	0.9
pAT153	<16	<16	

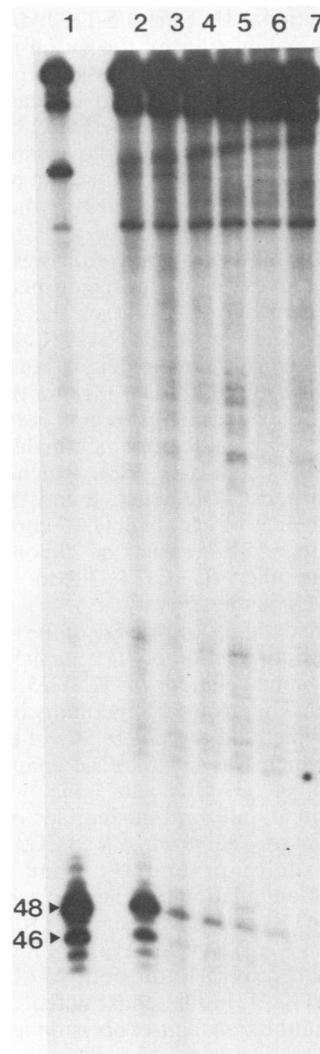


FIG. 3. Nuclease S1 mapping of cytoplasmic RNA from transfected cells. Lane 1, hybridization of 0.2 μ g of IE RNA to the 5'-end-labeled *AvaII* A fragment of pGX35 (28; Fig. 1B). The 48- and 46-nucleotide bands, which represent hybrids formed by the 5' portion of IE mRNAs 4 and 5, are indicated. Lanes 2 to 7, hybridization of 5 μ g of cytoplasmic RNA from transfected cells to the probe. Lanes 2 and 5, transfection with pS12TK; lanes 3 and 6, transfection with pS14TK; lanes 4 and 7, transfection with no plasmid DNA. RNA samples were from cells treated with inactivated *tsK* mutant (lanes 2 to 4) or mock-infected cells (lanes 5 to 7). The intense bands at the top of all lanes are due to reassociation of both the probe and a pAT153 band which copurified with the *AvaII* A fragment.

DNA alone contained some IE gene 4/5-specific RNA (lane 4), presumably because this small target escaped total inactivation by UV irradiation. Cells transfected with pS12TK, however, showed an increase in correctly initiated transcripts (lane 2), whereas cells transfected with pS14TK did not (lane 3). This experiment demonstrates that the levels of TK measured in the transient-expression assay reflect the production of plasmid-directed cytoplasmic transcripts with IE mRNA 4/5-specific 5' termini.

Since plasmids pS12TK and pS14TK contained IE gene 4/5 upstream sequences fused to the TK structural gene, an

increase in functional IE gene 4/5-TK hybrid mRNA must account for the stimulation by the virion component. This result is in agreement with a previous report that positive regulation of IE gene 3 is due to increased synthesis of mRNA (7). With the exception of pTKN13, the plasmids to be described in the next section all have small deletions only in the far-upstream region, and changes in production of TK enzyme therefore reflect alterations at the transcriptional level.

Further analysis of far-upstream sequences. Figure 4 shows the nucleotide sequence for the region between -410 and -270, which spans the endpoint of the IE gene 4/5 far-upstream region. Since plasmids pS12TK and pS17TK were stimulated by the virion component, it is clear that neither the TAACGAGGAAC (-391 to -381) nor the GGGCGGGG (-373 to -366) sequences in this region are essential for the response. Two additional features are highlighted in Fig. 4: TAATGAGATAC (-338 to -328), which is very closely related to a consensus sequence found in all HSV-1 IE genes, and two blocks of very high G+C content. One block, (G+C)1 (-375 to -354), consists of 22 consecutive G or C residues, and the other, (G+C)2 (-303 to -276), has one T among 27 G or C residues.

To gain a more detailed knowledge of the sequences which are important for stimulation of IE gene 4/5 transcription, a second set of plasmid deletions (pTKN7-13, see Fig. 2 and 5) was constructed. The upstream endpoints of these plasmids were at -402, which did not delete any of the far-upstream region, -347, which was close to the endpoint of pS17TK but removed all of (G+C)1, and -335, which effectively changed the second and third nucleotides of the consensus sequence at -338 to -328 from AA to CC, since the *EcoRI* linker at the deletion endpoint had the structure GGAATTCC. Manipulation of these plasmids with fragments from pS14TK, pS16TK, and pS13TK enabled internal deletions of TAATGAGATAC (pTKN12), (G+C)1 (pTKN8), (G+C)2 (pTKN10) or both (G+C)1 and (G+C)2 (pTKN11 and pTKN13) to be constructed. The behavior of these plasmids in the transient-expression assay is shown in Table 2.

Expression of pTKN8 was stimulated by inactivated *tsK* to a smaller extent than was pTKN7 expression, confirming that, as in the case of pS12TK and pS17TK, sequences in the (G+C)1 region affect the response to the virion component. The *SalI* linker at position -299 in pTKN7 and pTKN8 did not affect TK expression, since pTKN7 showed comparable activity to pS1TK. The main alteration caused by the linker insertion was to change the first two nucleotides of (G+C)2 from CC to GA. When the deletion extended to -335 (pTKN9), no response to inactivated *tsK* was observed. A boundary for the essential part of the far-upstream region was therefore located between -347 and -335. Plasmid pTKN10 showed similar properties to pTKN8: response to the virion component was reduced, compared with pTKN7, but not abolished. The effects of deleting both (G+C)1 and (G+C)2 appeared to be additive, since pTKN11 was more impaired than either pTKN8 or pTKN10.

These results suggest a modulating rather than absolute role for (G+C)1 and (G+C)2 in the regulation of IE gene 4/5 transcription, a conclusion which is supported by analysis of two further plasmids, pTKN12 and pTKN13. The deletion in pTKN12 removed 47 bp, including TAATGAGATAC, but left (G+C)1 and (G+C)2 essentially intact. No response of this plasmid to the virion component could be detected. The importance of sequences which flank TAATGAGATAC was further emphasized by the extremely reduced response of pTKN13 to stimulation by inactivated *tsK*. This plasmid consisted of 43 bp of the far-upstream region, including TAATGAGATAC but not (G+C)1 or (G+C)2, fused to position -151 from the 5' terminus of IE gene mRNA 4/5. Although the altered distance from promoter sequences may contribute to this finding, a clear conclusion is that TAATGAGATAC per se is not sufficient for stimulation of IE gene 4/5 transcription by the virion component.

DISCUSSION

The experiments described here show that IE gene 4/5 shares an overall similarity with other IE genes in its upstream sequence arrangement. A promoter element is

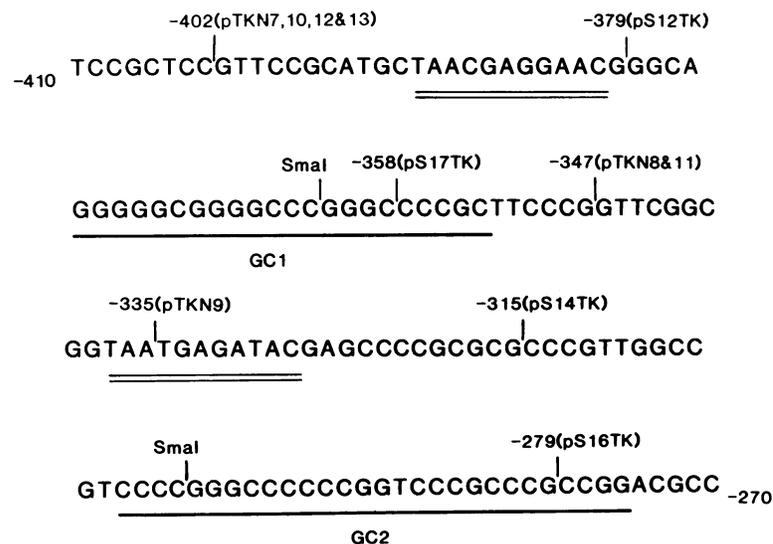


FIG. 4. Nucleotide sequence between positions -410 and -270, relative to the 5' terminus of IE mRNA 4/5. The regions defined as (G+C)1 (GC1) and (G+C)2 (GC2) are underlined, and two 11-bp elements which resemble the TAATGARATTC consensus are double-underlined. Deletion endpoints and *SmaI* sites are also shown.

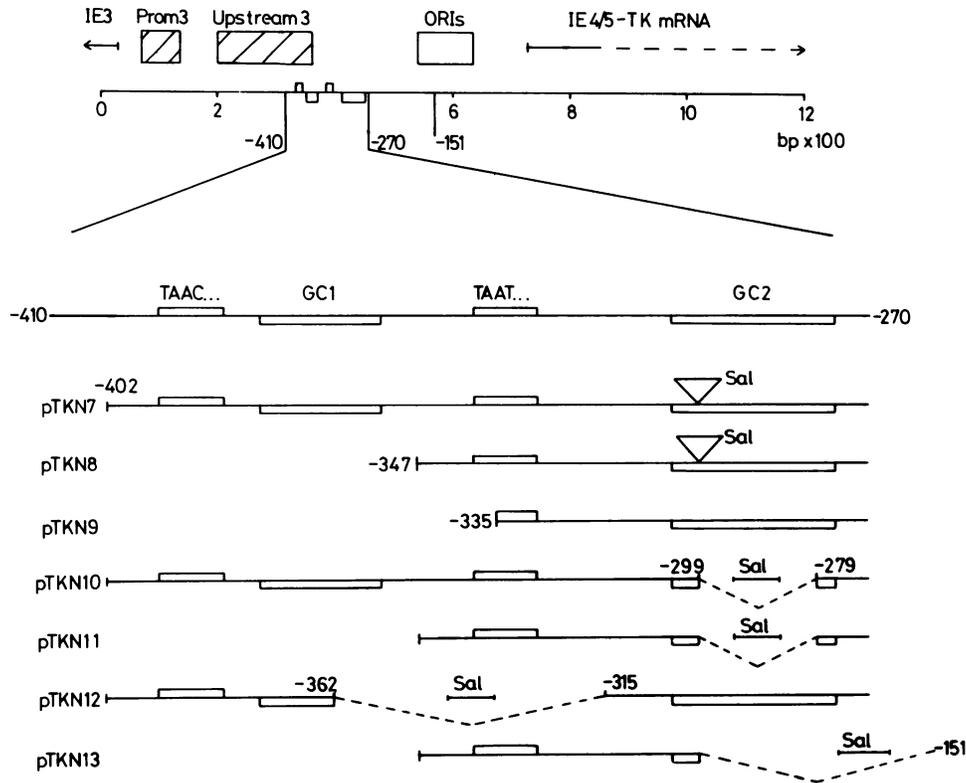


FIG. 5. Structures of far-upstream regions in plasmids pTKN7 to pTKN13. The upper part shows the 1,200-bp region which encompassed the 5' termini of IE mRNAs 3 and 4/5. The (G+C)1 (GC1) and (G+C)2 (GC2) sequences are represented by boxes below the lines, and two sequences which resemble the TAATGARATTC consensus (TAACGAGGAAC [TAAC . . .] and TAATGAGATAC [TAAT . . .]) are represented by boxes above the lines. The lower part shows the region from -270 to -410, the sequence of which is given in Fig. 4, on an expanded scale. The deletion endpoints and the position of an 8-bp *Sal*I linker (GGTTCGACC) are marked. Note that pTKN13 has the sequences from -299 to -347 fused to position -151, the upstream endpoint of pS13TK.

located close to the mRNA 5' terminus, and a far-upstream region contains sequences which act as a target for stimulation by a virion component. The proposed layout of this region is shown in Fig. 6.

Efficient transcription of IE gene 4/5 occurred when only 69 bp of upstream DNA was present, showing that the promoter is located close to the mRNA 5' terminus. The reduced activity of pS21TK, which had only 57 bp of upstream DNA, suggests that, as found in the analysis of IE gene 3 (7), sequences other than the TATA homology are important for promoter function. Further plasmid constructions will be necessary to determine precisely the functionally active DNA sequences in this region. It is clear, however, that ORIs are located between the far-upstream and promoter regions of IE gene 4/5 (31). In the case of simian virus 40 late gene expression, in which a similar arrangement of functional elements exists, it has been shown that DNA synthesis affects transcription (6). It will be interesting to determine whether HSV DNA synthesis influences expression of IE gene 4/5 or, conversely, whether transcription affects the function of ORIs.

The regulatory units for IE genes 3 and 4/5 appear to be separate, since pS17TK responds significantly to the virion component but does not overlap with any of the IE gene 3 far-upstream region (7). It is possible, however, that these adjacent regulatory regions interact when both are present, for example, in viral DNA. In one detailed respect, transcription of IE genes 3 and 4/5 shows differences. The far-upstream region of IE gene 4/5 only stimulated transcription

when the virion component was present and did not detectably augment expression in mock-infected, transfected cells. In contrast, the far-upstream region of IE gene 3 stimulates transcription at homologous and heterologous promoters in the absence of virion components and shows some similarities to enhancer elements (7; C. Preston and D. Tannahill, manuscript in preparation). The molecular basis for this difference remains to be determined.

A detailed analysis of the far-upstream region of IE gene 4/5 localized an essential element in the 49 bp between -347 (defined by pTKN8) and -299 (defined by pTKN10). The 12 bp between -347 and -335 (defined by pTKN9) contained nucleotides which were absolutely required for response to

TABLE 2. Transient expression of TK by pTKN plasmids

Plasmid	TK activity (cpm/min of assay per µg of protein per µg of plasmid in:		Degree of stimulation (-fold)
	<i>tsK</i> -treated cells	Mock-infected cells	
pTKN7	1191	140	8.5
pTKN8	402	122	3.4
pTKN9	123	121	1.0
pTKN10	826	145	5.7
pTKN11	291	137	2.1
pTKN12	113	144	0.8
pTKN13	127	112	1.3
pAT153	<21	<21	

distance of the far-upstream region from the promoter or to changing the relative positions of sequences within the far-upstream region. The net sequence change was only the replacement of 20 bp of (G+C)₂ with an 8-bp *Sal*I linker (GGTCCGACC). Studies with the far-upstream region of IE gene 3 have shown that spacing alterations of this magnitude do not significantly affect its activity (C. Preston and D. Tannahill, unpublished data). Removal of TAATGAGATAC from virtually all its normal flanking sequences, as in pTKN13, severely reduced the response to the virion component, emphasizing that this element alone is not sufficient. In the case of pTKN13, however, the net internal deletion was 139 bp, resulting in a much greater disruption of spacing.

The basis for the effects of flanking regions is difficult to ascertain. The additive nature of (G+C)₁ and (G+C)₂ deletions argues against interactions between them, even though GGGCCCCGGGCC in (G+C)₁ and CCCC GGCCCC in (G+C)₂ contain inverted and direct repeats. It has been suggested that G+C-rich inverted repeats within sequences which flank TAATGARATTC may contribute to the response to the virion component (13, 14). In this respect, (G+C)₁ contains an inverted repeat with no intervening sequences (-371 to -354), whereas a part of (G+C)₂ (-286 to -277) is repeated in an inverted form 19 bp downstream, giving a potential structure with a 10-bp, G+C-rich stem and a 19-bp loop (19, 37). If such elements are important, the loop size must be irrelevant to their mode of action. Figure 7 shows that all IE gene upstream regions have stretches of very high G+C DNA flanking a TAATGARATTC consensus. The sequence CCCGCC or its complement GGGCGGG appeared in all these regions except the region upstream of TAATGATATTC in IE gene 1. These elements are very similar to CCCC GCCC and GGGCGG, which are present in important regions of the early TK promoter (17), and CCGCCC, which is an active component of the simian virus 40 21-bp early promoter (10). Any attempt to assign a function to this sequence in IE gene transcription must take account of the diverse structures and regulatory properties of the genes in which it has been identified. A further possible rationale for the effects of (G+C)₁ and (G+C)₂ is that DNA regions of almost 100% G+C, irrespective of their exact sequence, induce conformational changes at the A+T-rich consensus, thereby potentiating its activity (3). There are great differences in the lengths of the G+C-rich regions and their positions relative to TAATGARATTC (Fig. 7), and such variations might account for subtle differences in the properties of IE gene promoters, as discussed above for IE genes 3 and 4/5.

The concept of flanking sequences modulating the activity of an essential element has been recognized in other eucaryotic systems. A relevant example is the identification of two classes of point mutations within the simian virus 40 72-bp element (38). One class, in the core sequence, destroys transcription enhancing activity, whereas mutations in other regions only reduce the effect. The function of essential elements in the simian virus 40 DNA replication origin and HSV ORIs is also affected by flanking sequences (20, 31).

The results presented here show that the virion component recognizes a relatively large far-upstream domain in IE gene 4/5. It would be expected that a similar situation will apply to other IE genes. These findings provide a basis for detailed examination of the molecular changes caused by the virion component and suggest that the far-upstream region may be a binding site for the component itself or for other factors which interact with it. The outcome of such binding may be conversion of the far-upstream region to a potent

RNA polymerase II entry site or stabilization of structural changes which increase the activity of the normal promoter.

ACKNOWLEDGMENTS

We thank J. H. Subak-Sharpe for critical reading of the manuscript and E. C. McMonagle for skilled technical assistance.

M.G.C. was a recipient of a Medical Research Council Research Studentship.

LITERATURE CITED

- Anderson, K. P., R. H. Costa, 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.
- Batterson, W., and B. Roizman. 1983. Characterization of the herpes simplex virion-associated factor responsible for the induction of α genes. *J. Virol.* **46**:371-377.
- Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss. 1983. Physical characteristics in eucaryotic promoters. *Nucleic Acids Res.* **11**:4521-4540.
- Clements, J. B., J. McLauchlan, and D. J. McGeoch. 1979. Orientation of herpes simplex virus type 1 immediate early mRNAs. *Nucleic Acids Res.* **7**:77-91.
- Clements, J. B., R. J. Watson, and N. M. Wilkie. 1977. Temporal regulation of herpes simplex virus type 1 transcription: location of transcripts on the viral genome. *Cell* **12**:275-285.
- Contreras, R., D. Gheysen, J. Knowland, A. van de Voorde, and W. Fiers. 1982. Evidence for the direct involvement of DNA replication origin in synthesis of late SV40 RNA. *Nature (London)* **300**:500-505.
- Cordingley, M. G., M. E. M. Campbell, and C. M. Preston. 1983. Functional analysis of a herpes simplex virus type 1 promoter: identification of far-upstream regulatory sequences. *Nucleic Acids Res.* **11**:2347-2365.
- Dixon, R. A. F., and P. A. Schaffer. 1980. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. *J. Virol.* **36**:189-203.
- Easton, A. J., and J. B. Clements. 1980. Temporal regulation of herpes simplex virus type 2 transcription and characterization of virus immediate early mRNAs. *Nucleic Acids Res.* **8**:2627-2645.
- Everett, R. D., D. Baty, and P. Chambon. 1983. The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. *Nucleic Acids Res.* **11**:2447-2464.
- Jones, P. C., and B. Roizman. 1979. Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. *J. Virol.* **31**:299-314.
- Mackem, S., and B. Roizman. 1980. Regulation of herpesvirus macromolecular synthesis: transcription-initiation sites and domains of α genes. *Proc. Natl. Acad. Sci. U.S.A.* **77**:7122-7126.
- Mackem, S., and B. Roizman. 1982. Differentiation between the promoter and regulator regions of herpes simplex virus 1: the functional domains and sequence of a movable α regulator. *Proc. Natl. Acad. Sci. U.S.A.* **79**:4917-4921.
- Mackem, S., and B. Roizman. 1982. Structural features of the herpes simplex virus α gene 4, 0, and 27 promoter-regulatory sequences which confer α regulation on chimeric thymidine kinase genes. *J. Virol.* **44**:939-949.
- Macpherson, I., and M. G. P. Stoker. 1962. Polyoma transformation of hamster cell clones—an investigation of factors affecting cell competence. *Virology* **16**:147-151.
- Marsden, H. S., I. K. Crombie, and J. H. Subak-Sharpe. 1976. Control of protein synthesis in herpesvirus-infected cells: analysis of the polypeptides induced by wild-type and sixteen temperature-sensitive mutants of strain 17. *J. Gen. Virol.* **31**:347-372.
- McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control sequences of a eucaryotic protein-coding gene. *Science* **217**:316-324.

18. Mocarski, E. S., and B. Roizman. 1982. Herpesvirus-dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral sequences and linked to an origin of viral DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* **79**:5626-5630.
19. Murchie, M.-J., and D. J. McGeoch. 1982. DNA sequence analysis of an immediate-early gene region of the herpes simplex virus type 1 genome (map co-ordinates 0.950 to 0.978). *J. Gen. Virol.* **62**:1-15.
20. Myers, R. M., and R. Tjian. 1980. Construction and analysis of simian virus 40 origins defective in tumour antigen binding and DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* **77**:6491-6495.
21. Notarianni, E. L., and C. M. Preston. 1982. Activation of cellular stress protein genes by herpes simplex virus temperature-sensitive mutants which overproduce immediate early polypeptides. *Virology* **123**:113-122.
22. Post, L. E., S. Mackem, and B. Roizman. 1981. Regulation of genes by herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. *Cell* **24**:555-565.
23. Post, L. E., B. Norrild, T. Simpson, and B. Roizman. 1982. Chicken ovalbumin gene fused to a herpes simplex virus α promoter and linked to a thymidine kinase gene is regulated like a viral gene. *Mol. Cell. Biol.* **2**:233-240.
24. Preston, C. M. 1977. The cell-free synthesis of herpesvirus-induced polypeptides. *Virology* **78**:349-353.
25. 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**:275-285.
26. Preston, C. M. 1979. Abnormal properties of an immediate early polypeptide in cells infected with the herpes simplex virus type 1 mutant *tsK*. *J. Virol.* **32**:357-369.
27. Rixon, F. J., M. E. Campbell, and J. B. Clements. 1982. The immediate-early mRNA that encodes the regulatory polypeptide Vmw175 of herpes simplex virus is unspliced. *EMBO J.* **1**:1273-1277.
28. Rixon, F. J., and J. B. Clements. 1982. Detailed structural analysis of two spliced HSV-1 immediate-early mRNAs. *Nucleic Acids Res.* **10**:2241-2256.
29. Shen, Y.-M., R. R. Hirschhorn, W. E. Mercer, E. Surmacz, Y. Tsutsui, K. J. Soprano, and R. Baserga. 1982. Gene transfer: DNA microinjection compared with DNA transfection with a very high efficiency. *Mol. Cell. Biol.* **2**:1145-1154.
30. Stow, N. D. 1982. Localization of an origin of DNA replication within the TR_s/IR_s repeated region of the herpes simplex virus type 1 genome. *EMBO J.* **1**:863-867.
31. Stow, N. D., and E. C. McMonagle. 1983. Characterization of the TR_s/IR_s origin of DNA replication of herpes simplex virus type 1. *Virology* **130**:427-438.
32. Stow, N. D., and N. M. Wilkie. 1976. An improved technique for obtaining enhanced infectivity with herpes simplex virus type 1 DNA. *J. Gen. Virol.* **33**:447-458.
33. Swanstrom, R. I., K. Pivo, and E. K. Wagner. 1975. Restricted transcription of the herpes simplex virus genome occurring early after infection and in the presence of metabolic inhibitors. *Virology* **66**:140-150.
34. Watson, R. J., and J. B. Clements. 1980. A herpes simplex virus type 1 function required for early and late virus RNA synthesis. *Nature (London)* **285**:329-330.
35. Watson, R. J., C. M. Preston, and J. B. Clements. 1979. Separation and characterization of herpes simplex virus type 1 immediate-early mRNAs. *J. Virol.* **31**:42-52.
36. Watson, R. J., M. Sullivan, and G. F. Vande Woude. 1981. Structures of two spliced herpes simplex virus type 1 immediate-early mRNAs which map at the junctions of the unique and reiterated regions of the virus DNA S component. *J. Virol.* **37**:431-444.
37. Watson, R. J., and G. F. Vande Woude. 1982. DNA sequence of an immediate-early gene (IE mRNA-5) of herpes simplex virus 1. *Nucleic Acids Res.* **10**:979-991.
38. Weiher, H., M. Konig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-631.
39. Whitton, J. L., F. J. Rixon, A. J. Easton, and J. B. Clements. 1983. Immediate-early mRNA-2 of herpes simplex viruses 1 and 2 is unspliced: conserved sequences around the 5' and 3' termini correspond to transcription regulatory signals. *Nucleic Acids Res.* **11**:6271-6287.
40. Wilkie, N. M., J. B. Clements, W. Boll, N. Mantei, D. Lonsdale, and C. Weissmann. 1979. Hybrid plasmids containing an active thymidine kinase gene of herpes simplex virus 1. *Nucleic Acids Res.* **7**:859-877.