## "Transactivation" control signals in the promoter of the herpesvirus thymidine kinase gene

(HeLa cell transient transcription system/linker scanning mutants)

Adele ElKareh<sup>\*</sup>, Andrew J. M. Murphy<sup>†</sup>, Timothy Fichter<sup>\*</sup>, Argiris Efstratiadis<sup>†</sup>, and Saul Silverstein<sup>\*</sup>

Departments of \*Microbiology and †Human Genetics and Development, College of Physicians and Surgeons, Columbia University, 710 West 168th Street, New York, NY 10032

Communicated by I. S. Edelman, October 9, 1984

ABSTRACT The herpes simplex virus thymidine kinase (TK) gene is transcriptionally activated in *trans* ("transactivated") by virus-encoded proteins during the infectious cycle. We show that TK plasmids introduced into a HeLa cell transient transcription assay system are also transactivated after infection with a TK<sup>-</sup> virus. Several aspects of this response are similar to regulation during the normal infectious cycle. Assay of TK promoter deletion and 5- to 10-base-pair substitution mutants in this system reveals that the transactivation response depends on the intactness of 109 base pairs of 5' gene flanking sequence. Differences between these results and analogous assays in the *Xenopus* oocyte system are discussed. A model for the putative binding of transactivator(s) to the promoter region is presented.

Expression of herpes simplex virus (HSV) genes in infected cells is temporally regulated in a cascade fashion (1). Viral mRNAs transcribed immediately after infection from parental templates are translated into gene products ( $\alpha$  polypeptides) that can positively regulate in *trans* ("transactivation") the transcription of a second group of genes ( $\beta$  gene group) (2, 3), which includes the thymidine kinase (TK) gene (4, 5). Subsequently, transcriptional events that requires the presence of both  $\alpha$  and  $\beta$  polypeptides is initiated from a third gene group ( $\gamma$ ). This complex interactive regulatory circuitry is presumably controlled by recognition of multiple DNA elements by specific virus regulatory proteins.

The 5' flanking region of TK has been used as a model system for the analysis of transcriptional signals (6). Assay of deletion mutants has shown that these *cis*-acting signals reside in the DNA sequence extending from the capping site of the gene (position  $\pm$ 1) to position  $\pm$ 109 (TK promoter) (6, 7). The sequences of this region are required for proficient conversion of TK<sup>-</sup> cells to the TK<sup>+</sup> phenotype (7, 8) and for "unregulated" transcription in *Xenopus* oocytes (6, 9).

Injection of a series of *in vitro* generated 5- to 10-base-pair (bp) substitution mutants [linker-scanning (LS) mutants] of the TK promoter into *Xenopus* oocytes (9) revealed that, by the criterion of this assay, this promoter is composed of three discrete domains: the proximal signal (nucleotides -16to -32), which includes the "ATA" box, the first distal signal (-47 to -61), and the second distal signal (-80 to -105). All three signals are required for quantitative expression. In addition, the proximal signal is involved in specifying transcriptional start sites (9).

Previous results (10, 11) suggested that sequences residing in the same 109 bp of the TK promoter, when present in the chromatin of stably transformed cells, can be positively regulated in *trans* by infection with HSV.

In order to determine more precisely the set of the promoter sequences that respond to virus transactivation factors and to compare them with the signals that are recognized in the amphibian oocytes, we examined the transcriptional performance of LS mutant DNAs in a HeLa cell transient transcription assay system (12).

## **MATERIALS AND METHODS**

**Cells and Viruses.** The viruses used in this study were  $D_2$ , a TK<sup>-</sup> deletion mutant of HSV type 1 (HSV-1) (13), and F (14). They were grown and titrated in Vero cells (15). HeLa cells were grown in Dulbecco's modification of Eagle's medium (DME medium) supplemented with 10% calf serum and were seeded at a density of  $1-3 \times 10^6$  cells per 10-cm dish the day before transfection.

**Recombinant DNAs.** The deletion mutant (6, 7) and LS plasmids (9) were prepared as described (16).

**Transfection.** HeLa cells were transfected as described (17) with a calcium phosphate precipitate containing 5–15  $\mu$ g of plasmid DNA and 15  $\mu$ g of carrier salmon sperm DNA. At 24 hr after transfection, the cells were infected with 3–10 plaque-forming units of D<sub>2</sub> for 1 hr. The virus was removed, and the cells were incubated for 4–12 hr in DME medium containing 1% calf serum.

**RNA Analysis.** Total cell RNA was purified by the guanidine thiocyanate/CsCl procedure (18) and was analyzed by S1 nuclease protection (19) using homogenously labeled single-stranded probes synthesized from recombinant M13 templates by a modification of published procedures (20).

## RESULTS

Validity of the Transient Transcription Assay. To examine whether the HeLa cell transient transcription system could provide an appropriate assay for regulated expression of the viral TK gene, we performed the following control experiments.

We first examined whether TK that is no longer associated with the viral chromosome was regulated in the same way as it is during the course of the virus infectious cycle. HeLa cells transfected with plasmid DNA carrying TK were cultured for 36 hr, and then they were either mock-infected or infected with virus  $D_2$ . Cells were harvested 6 hr later, and the amount of TK transcripts from each culture was assayed by S1 nuclease protection (Fig. 1*a*). From these results, we

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TK, thymidine kinase; HSV, herpes simplex virus; HSV-1 and -2, HSV types 1 and 2; bp, base pair(s); LS, linker scanning.

estimate that the level of expression of this gene in the infected cultures is 3-4 orders of magnitude higher than the basal level.

A comparison of the level of TK transcripts in cells infected with wild-type virus and cells transiently expressing a TK gene with a deletion in the 5' noncoding region of the mRNA (pseudo-wild-type gene; ref. 9) after infection with D<sub>2</sub> is shown in Fig. 1b. The amount of TK transcripts in the transfected cells, which were transactivated by D<sub>2</sub>, was comparable to that in the cells infected with wild-type virus ( $\approx$ 50% in this example), if we assume that 1% to 10% of the cells received DNA. This experiment also demonstrates that, in the transfected cells, the RNA is transcribed from the plasmid DNA and not from contaminating TK<sup>+</sup> virus, as evidenced by the short length of the protected probe.

We infer from the results of Fig. 1 a and b that the basal level of the TK transcripts in the transient transcription system is much lower than that in stably transformed L cells



FIG. 1. Level of accumulated TK RNA in transiently expressing HeLa cells. (a) The concentration of TK transcripts in HeLa cells transfected with 10  $\mu$ g of plasmid DNA containing wild-type sequences was measured by S1 nuclease protection after hybridization of total cellular RNA with an excess of a 258-nucleotide uniformly <sup>32</sup>P-labeled single-stranded probe (lane P) synthesized on a M13mp8 template extending from a synthetic BamHI site at position -109 (6) to the Hpa II site at position +122 (Lower). Hybridizations were performed in 10-15 µl as described (21) at 52°C for 16 hr and then were diluted in 300  $\mu$ l of S1 digestion buffer (300 mM NaCl/50 mM NaOAc, pH 4.5/1 mM ZnSO<sub>4</sub>) containing 70 units (22) of S1 nuclease and digested at 37°C for 1 hr. Protected fragments were electrophoresed on a denaturing polyacrylamide gel (23). Lanes: A (overexposed), probe protected after hybridization with 40  $\mu$ g of RNA from mock-infected cells; B-E, probe protected by 100, 200, 400, and 800 ng of RNA from D<sub>2</sub>-infected cells, respectively; M, DNA size markers (Hpa II fragments of M13mp7 replicative form DNA). (b) S1 protection of the same probe as in a by RNA isolated from HeLa cells that were either transfected with 10  $\mu$ g of pseudo-wildtype DNA and subsequently infected at a multiplicity of infection of 3 with  $D_2$  (10 and 1  $\mu g$  in lanes A and B, respectively) or mocktransfected and then infected with wild-type HSV (0.1 and 1  $\mu$ g in lanes C and D, respectively). (Lower) Transcripts from wild-type DNA can protect 126 nucleotides of the probe, while transcripts from the pseudo-wild-type DNA can protect 90 nucleotides. Restriction sites: B, BamHI; E, EcoRI; Bg, Bgl II; H, Hpa II.

because, in the selected L cell convertants, the TK RNA (5– 10 copies per cell) was only 100-fold less than in virus-infected cells (7). Our results are consistent with the observation that another  $\beta$  group gene is barely expressed in the absence of transactivation in a similar transient transcription system (24).

To demonstrate that the transactivation of TK requires expression of  $\alpha$  gene products as it does during the normal infectious cycle (2, 4, 5), we carried out the following experiment. HeLa cells transfected with TK DNA were cultured for 36 hr and then were infected with  $D_2$  in the presence or absence of cycloheximide. At 6 hr after infection, a sample from each culture was removed for RNA isolation. At the same time the treated cultures were washed free of drug, and both sets of cultures were incubated for an additional 6 hr prior to harvesting and RNA isolation. Fig. 2a shows the results of the S1 nuclease analyses of the RNA from these samples. Cells maintained in the presence of cycloheximide (lane A) accumulated very little TK RNA when compared with cells incubated without the drug (lane C). Reversal of the action of the protein synthesis inhibitor by washing (lane B) permitted the translation of  $\alpha$  mRNAs that accumulated during the inhibition. In turn, the translation products transactivate the TK promoter. The amount of TK transcripts that accumulate during the 6 hr of incubation after drug removal (lane B) is similar to that accumulating at 12 hr after infection in the untreated culture (lane D). Thus, TK transcription in the HeLa cell system is dependent on the prior synthesis of  $\alpha$  polypeptides.



(a) Transient transcription of TK in the presence of a FIG. 2. protein synthesis inhibitor. Lanes: A-D, fragments of the probe (shown in lane P) protected from S1 nuclease by 25  $\mu$ g of RNA isolated from D<sub>2</sub>-infected cells maintained in the presence of cycloheximide at 50  $\mu$ g/ml for 6 hr (lane A), infected cells maintained in the presence of cycloheximide for 6 hr and then washed free of drug and incubated for an additional 6 hr (lane B), and untreated infected cells after 6 hr (lane C) or 12 hr (lane D) of incubation; M, DNA size markers as in Fig. 1. (b) Transactivation of TK promoter deletion mutants. HeLa cells were cotransfected with deletion mutant DNAs and pseudo-wild-type TK DNA and then infected with D<sub>2</sub>. Lanes: A-F, probe fragments protected by hybridization to RNA prepared from cells transfected with deletion mutants -74 (lane A), -80 (lane B), -81 (lane C), -85 (lane D), -95 (lane E), and -109 (lane F, which contains 1/5 equivalents of the digestion products in lanes A-E); G, as lane F but with four times the amount of the S1 nuclease digestion products; P, undigested probe; M, DNA size markers as in Fig. 1. Fragment g is the one protected by the transcript of the mutant gene and fragment  $\psi$  is the one protected by the transcript of the control pseudo-gene.

**Transactivation of TK Promoter Deletion Mutants.** To define the 5' limit of the sequences that are necessary for the transactivation response, we first used deletion mutants. In the absence of transactivation, we were unable to detect a difference in the level of transcripts from wild-type or deletion mutant templates. We note, however, that the constitutive level of transcription is close to the lower limits of sensitivity of our assay.

Replicate cultures of HeLa cells were transfected with a mixture of 5' deletion mutant and pseudo-wild-type DNA (internal control); at 20 hr after transfection, the cultures were infected with  $D_2$ , and RNA was isolated 6 hr later. Fig. 2b shows that deletion mutants that terminate at -85 and -95 (lanes D and E, respectively) could still be transactivated. However, the ratio of the level of transcripts from these mutants to the control (5% and 13%, respectively) was lower than that of the -109 template (lanes F and G). Templates containing <85 bp of 5' flanking sequence were not efficiently transcribed after infection (lanes A, B, and C). Because mutant -109 was transactivated close to the pseudo-wild-type level, we conclude that the 5' limit of the transactivation recognition sequence is around this position.

Transactivation of TK Promoter LS Mutants. We assayed the LS substitution mutants in the HeLa cell transient expression system to examine the role of short regions in the DNA of the TK promoter in the transactivation response. Each LS mutant was cotransfected along with the pseudowild-type control DNA into HeLa cells, and 20 hr later the cells were infected with D<sub>2</sub>. RNA was prepared at 6 hr after infection and was analyzed by S1 nuclease protection. Fig. 3 shows the autoradiogram of a gel from a typical assay. For quantitation, protected probe bands corresponding to transcripts from each pair of control and LS mutant DNAs were excised from the gel and assayed by liquid scintillation. Ratios (mutant/control) were determined and normalized to the ratio of LS16. Fig. 4 presents these data from four independent experiments in the form of a histogram. Analogous data (9) from the frog oocyte system are also shown for comparison. From this analysis we conclude that the promoter sequence contains two types of signals. One of them includes



FIG. 3. Transactivation of TK promoter LS mutants. HeLa cells were cotransfected with control pseudo-wild-type DNA and each of the LS mutants (10), LS1 (-6/-16), LS2 (-12/-21), LS3 (-18/-29), LS4 (-32/-42), LS5 (-37/-47), LS6 (-46/-56), LS7 (-49/-59), LS8 (-61/-70), LS9 (-69/-70), LS10 (-70/-80), LS11 (-74/-84), LS12 (-85/-95), LS13 (-95/-105), LS14 (-101/-111), LS15 (-105/-115), and LS16 (-109/-119). After infection with D<sub>2</sub>, RNA was isolated and analyzed by S1 nuclease protection of the same probe as in Fig. 1 (lane P). Fragments g and  $\psi$  are as in Fig. 2b. The numbers above the lanes refer to the LS mutants, as indicated above. Control (lane C) consisted of digestion products of a mock reaction carried out with RNA from untransfected HeLa cells.

the ATA box and seems to be recognized in an analogous manner in both assay systems. The 3' limit of the other signal by the criterion of the HeLa cell assay corresponds approximately to position -39. Differences in the behavior of the LS mutants spanning the region between positions -39 and -109 in the two assay systems are discussed below.

## DISCUSSION

We examined a series of mutants in the promoter of the HSV TK gene in a HeLa cell transient expression assay system to determine which sequences are required for regulated expression. In the absence of virus infection, neither the wild-type nor the mutant genes (not shown) accumulate transcripts at significant levels. However, infection of transfected cultures results in accumulation of high levels of TK RNA. Thus, this transactivation mimics the response of the gene during the course of the wild-type virus infectious cycle and is dependent on the integrity of the sequences that compose the TK promoter. Mutations of distinct subsets of the promoter sequence alter differentially the level of transactivation.

The LS mutants that affect the ATA box region behave in a similar manner in both the oocyte and the HeLa cell assay system and produce defective phenotypes, though the reduction in transcription by the LS2 mutation is less severe in the latter case. Despite the lack of formal proof in either case, these results can be explained by postulating that the mutations in this region affect the interaction with an ATA-binding factor (25, 26).

Mutants LS6, -7 and -11 through -14 are transcribed poorly in the oocyte system and define the first and second distal transcriptional signals (9). However, DNAs with mutations that map between these two regions (LS8 to LS10) are phenotypically very close to wild type. In contrast, in the HeLa cell system, all of the DNAs with mutations between positions -39 and -109 (LS5 to LS15) are not efficiently transcribed. Levels of transcripts are between 40% and 50% of that of the control, with the exception of two mutants LS7 and LS11, for which accumulation is even lower (24% and 15%, respectively). Though the latter two mutations map in regions that are parts of the two distal signals, the overall picture of the sequences that are required for transcription is clearly different between the two assay systems. This difference cannot be attributed to competition between the pseudo-wild-type and LS mutant templates for transcriptional factors in the oocyte but not in the HeLa system because in both cases the total template concentration is at subsaturating levels (ref. 9 and our unpublished observations).

The most straightforward way to interpret our results is to assume that the transactivation factor(s) bind to the TK promoter region approximately between positions -37 and -109. The uniformity in transcriptional reduction (with the exceptions of LS7 and LS11) makes it likely that bases distributed throughout the entire region are involved.

To examine whether a plausible (and testable) model can be formulated for the putative interaction between the TK transactivator and the DNA recognition sequence, we searched the promoter region affected by the LS mutations for dyad symmetries because in other cases (27–31) it has been shown that a pair of antiparallel  $\alpha$  helices of a regulatory protein dimer can recognize a palindrome in the DNA sequence. Fig. 5 shows that the center of this region is a hyphenated palindrome. There are also weak symmetries that flank this region. If one assumes that these symmetries are important for recognition by the transactivator(s), we can now explain the severely defective phenotype of the LS11 mutant. The mutations in this DNA affect the central dyad symmetry more than the mutations which have a less defective phenotype (LS8, -9, and -10). A difficulty with this



FIG. 4. Comparison of the phenotypes of LS mutant DNAs in HeLa cells and amphibian oocytes. (*Upper*) The ratio of the amount of transcripts accumulated in HeLa cells by LS mutant and pseudo-gene (control) DNAs was determined and normalized to that corresponding to LS16. The mean values from four independent experiments (like the one presented in Fig. 3) are plotted in the form of a histogram for visual display. The extreme values from all of the experiments corresponding to the middle of each bar and representing the range of experimental variability have been connected (stippled area). The dots indicate the level of transcripts accumulated by each of the LS mutants in *Xenopus* oocytes (data from ref. 9). (*Lower*) Nucleotide substitutions of the TK wild-type promoter sequence in each of the LS mutants (LS1 to LS16) are indicated. The proximal and the two distal transcriptional signals are underlined. Arrows indicate a hyphenated central palindrome in the sequence (see Fig. 5).

TK-2	ŶŎĬĊŎŶŔĸĬĹŔĹĿŢŎŎŎĸĹŎŔŦ <mark>ŎĿŢŎ</mark> ŶŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎŎ
TK-1	CTATGATGACAAAACCCCCCCCCCCCCCCCCCCCCCCCC
GD	CCGTGTGACACTATCGTCCATACGEACCACACGAGAAAATCCCCCCA AGGGGGACGAGGGCCATIIIIIIACGAGGAGGGGTATAACAAAGGTCTGTCTTTAAAAAGCA GGCACACTGTGATAGCAGGTATGGCTGGIIGIIIGGCTGCTTAGGGGGT TCCCCCCICCCGGGTAAAATGCCCCCCAAAATTGCTCCCCGATATTGTTLCAGACAGAAATTTTTCGT
DUTPASE	ACAGGECGEGCTACATCAGECCTACCAGECCAGCAGCAGECCATCACTACTACCACAGECCATCACCACAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECCATCAGECATCAGECATCAGECAGECATCAGECAGECATCAGECAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECATCAGECAGAGECAGECAGECAGECAGECAGECAGECAGECA
ALKEXO	CGAGACCAACACCACGEC TGGGETATEACGTCC CAGAGEGCATCCENTCCCACACTCCCAACATTCCCAACATTCCCCACGTCGCGCGCCCCAATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCACATCCCCAACATCCCCAACATCCCCAACATCCCCAACATCCCCACATCCCCACATCCCCACACATCCCCACATCCCCACACATCCCCACATCCCCACACATCCCCACATCCCCACACATCCCCACATCCCCACACATCCCCACACATCCCCACACATCCCCACACATCCCCACACACCCCCC
42	TICGC6664AACCC6AGACC6TITG666G9ETGICITGITITTTGTGCCCCCGIEG6G4TIGGCIACC6ATG66GTACCCATG66GTATATAGCACTICCCATATAAGACTCCCCCCA AAGC66CCTTTG6CTCT66CAAACCCC6GCACG6GACGAAGAAACCC666GACCCCTAAAGACTACCCAACG6GTACCCAACG6GTATATACG5GAAGGACAATCCC6AG
140	ACACACCCCCGTGACTCAGGACATCGGCGTGTGTTTTCGTTTTCACTGAAAQTGGTCCGCGCGCGCGCGCTGCGCGCGCGCGCGCGCGCGC
REDUCTASE	ATCCATGACCCTGTATGTCACGGAGAAAGCGGAGAGGGACCCCCCAGCCCCCCGGTCCTGTCCGCGCGCG

FIG. 5. Dyad symmetries in the promoter regions of HSV  $\beta$  group genes. The TK promoters from HSV-2 (32) and HSV-1 (TK-2 and TK-1, respectively) are compared. Asterisks indicate nucleotide substitutions. Dyad symmetries are boxed. The sequence for TK-1 is from strain MP (33). A cross indicates a G-C bp, which is C-G in HSV-1, strain C1 101 (34). The borders of the two distal transcriptional signals in the TK-1 promoter are indicated, and the hexanucleotide inverted repeat (9) they contain is overlined. All other sequences are from HSV-1 and correspond to the genes encoding glycoprotein D (gD) (24), dUTPase (35), alkaline exonuclease (alkexo) (36), and ribonucleotide reductase (37). The promoters of the genes encoding two proteins of unknown function of 42 and 140 kilodaltons (35, 38) are also shown (42 and 140). Known or putative ATA boxes are indicated by dashed double lines. Three of the genes have two ATA boxes, either one of which or both are functional. The mRNAs of all of the genes shown have multiple capping sites.

hypothesis is that, to explain the phenotype of LS7, one would have to argue that the single symmetric guanosine at position -54 is more important than the G·G doublet at -50/-51, which is affected by the substitution in LS5. Unfortunately, a mutation in the corresponding symmetric cytidine at position -95 is not available for analysis. Nevertheless, the general scheme of a dimeric transactivator recognizing the dyad symmetries of the TK promoter seems consistent with the universal appearance of such symmetries in the 5' flanking regions of all of the sequenced  $\beta$  group genes of HSV, which are also transactivated (Fig. 5). In most of these cases, the center of the dyad symmetry is almost equidistant from the ATA box.

The model is consistent with the following evolutionary argument. The TK gene we are studying is that of HSV-1, strain MP (33); TK of HSV type 2 (HSV-2) strain 333 has been sequenced also (32). The sequence divergence in the TK 5' flanking regions between HSV-1 and HSV-2 is only 10% (20 substitutions in 200 sites), and it is the same (9%) in the promoter regions (10 substitutions in 109 sites), while it is 19.5% (220 substitutions in 1131 sites) in the region encoding the TK polypeptide. Though 45% of the coding region substitutions lead to amino acid replacements, the two TK polypeptides exhibit equivalent enzymatic function (39). Furthermore, the HSV-1 TK gene can be transactivated by HSV-2 virus and vice versa (40). Because both the promoter and the coding region must be under evolutionary constraint, the significant difference in the degree of divergence between these two areas (9% vs. 19.5%) indicates that much stronger negative selection operates on the promoter. Accordingly, the observed substitutions between the TK promoters of the two HSV types can be viewed as fixed neutral mutations. It is striking that none of these substitutions affects the dyad symmetries in the sequence and that the entire region containing the central palindrome is completely conserved (Fig. 5). We note that the available experimental evidence (41) excludes the possibility that the putative binding of the oocyte factor that substitutes for the transactivator can be explained as the binding of a dimer to the hexanucleotide inverted repeats present in the two distal transcriptional signals (9, 41). We also note that one of the nucleotide substitutions between the TK promoters of HSV-1 and HSV-2 is present in the first distal signal, while the second distal signal contains four substitutions, one of them affecting the hexanucleotide inverted repeat.

The same 109 bp of 5' flanking sequence is required for the expression of the TK gene in transformed cells (7, 10, 11), in microinjected Xenopus oocytes (7), and in the transient expression/viral infection system described here. In each case a gene that contains 109 bp of upstream sequence transcribes or transforms at levels comparable to those of the wild type, while a gene with only 95 bp of 5' flanking DNA has an activity diminished to 10-60% of that of the control. Our results (Figs. 3 and 4) indicate that the infected HeLa cell utilizes the region between -109 and the ATA box quite differently than the injected oocyte. Thus, the TK promoter can be viewed as consisting of two overlapping elements, each acting in concert with a common ATA box. One of these elements consists of the first and second distal transcriptional signals described by McKnight (7, 9). It is active in amphibian oocytes and perhaps in transformed mammalian cells and promotes transcription in the absence of specific regulation. The other element is comprised of sequences between -39 and -109 (which contain a central palindrome of possible significance) and is specifically activated by the viral early genes.

We gratefully acknowledge the generous gift of TK plasmids from S. McKnight. We also thank E. K. Wagner for helpful discussions. This work was supported by grants from the National Institutes of Health to S.S. (CA17477 and AI 19724) and A.E. (AM32085).

- 1. Honess, R. W. & Roizman, B. (1974) J. Virol. 14, 8-19.
- 2. Honess, R. W. & Roizman, B. (1975) Proc. Natl. Acad. Sci.
- USA 72, 1276–1280.
- 3. Preston, C. M. (1979) J. Virol. 29, 275-284.
- 4. Garfinkle, B. & McAuslan, B. R. (1974) Biochem. Biophys. Res. Commun. 58, 822-829.
- Leiden, J. M., Buttyan, R. & Spear, P. G. (1976) J. Virol. 20, 413-424.
- McKnight, S. L. & Gavis, E. R. (1980) Nucleic Acids Res. 8, 5931-5948.
- McKnight, S. L., Gavis, E. R., Kingsbury, R. & Axel, R. (1981) Cell 25, 385-398.
- Zipser, D., Lipsich, L. & Kwoh, J. (1981) Proc. Natl. Acad. Sci. USA 78, 6276–6280.
- 9. McKnight, S. L. & Kingsbury, R. (1982) Science 217, 316-324.
- ElKareh, A., Silverstein, S. & Smiley, J. (1984) J. Gen. Virol. 65, 19-36.
- 11. Smiley, J. R., Swan, H., Pater, M. M., Pater, A. & Halpern, M. E. (1983) J. Virol. 47, 301-310.
- 12. Banerji, J., Rusconi, S. & Schaffner, W. (1981) Cell 27, 299-308.
- 13. Smiley, J. R. (1980) Nature (London) 285, 333-335.
- 14. Ejercito, P. M., Kieff, E. D. & Roizman, B. (1968) J. Gen. Virol. 2, 357-364.
- 15. Nishioka, Y. & Silverstein, S. (1977) Proc. Natl. Acad. Sci. USA 74, 2370-2374.
- Kahn, M., Kolter, R., Thomas, C., Figurski, D., Meyer, R., Renaut, E. & Helinski, D. R. (1979) *Methods Enzymol.* 68, 268-280.
- Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G. & Chasin, L. (1979) Proc. Natl. Acad. Sci. USA 76, 1373– 1376.
- Chirgwin, J. M., Przybla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* 18, 5294–5299.
- Favaloro, J., Treisman, R. & Kamen, R. (1980) Methods Enzymol. 68, 718-749.
- 20. Hu, N. & Messing, J. (1982) Gene 17, 271-277.
- Casey, J. & Davidson, N. (1977) Nucleic Acids Res. 4, 1539– 1552.
- 22. Vogt, V. M. (1973) Eur. J. Biochem. 33, 192-200.
- 23. Sanger, F. & Coulson, R. C. (1978) FEBS Lett. 87, 107-110.
- 24. Everett, R. D. (1983) Nucleic Acids Res. 11, 6647-6666.
- 25. Davison, B. L., Egly, J.-M., Mulvihill, E. R. & Chambon, P. (1983) Nature (London) 301, 680-686.
- 26. Parker, C. S. & Topol, J. (1984) Cell 36, 357-369.
- Lewis, M., Jeffrey, A., Wang, J., Ladner, R., Ptashne, M. & Pabo, C. O. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 435-440.
- Matthews, B. W., Ohlendorf, D. H., Anderson, W. F., Fisher, R. G. & Takeda, Y. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 427-434.
- 29. Steitz, T. A., Weber, I. T. & Matthew, J. B. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 419-426.
- 30. Tjian, R. (1981) Cell 26, 1-2.
- 31. Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980) Cell 20, 269-281.
- Swain, M. A. & Galloway, D. A. (1983) J. Virol. 46, 1045– 1050.
- 33. McKnight, S. L. (1980) Nucleic Acids Res. 8, 5949-5964.
- 34. Wagner, M. J., Sharp, J. A. & Summers, W. C. (1981) Proc.
- Natl. Acad. Sci. USA 78, 1441–1445. 35. Hall, L. M., Draper, L. K., Frink, R. J., Costa, R. H. & Wag-
- ner, E. K. (1982) J. Virol. 43, 594–607. 36. Costa, R. H. Draper, K. G. Banks, L. Powell, K. L. Co-
- Costa, R. H., Draper, K. G., Banks, L., Powell, K. L., Cohen, G., Eisenberg, R. & Wagner, E. K. (1983) J. Virol. 48, 591-603.
- Draper, K. G., Frink, R. J. & Wagner, E. K. (1982) J. Virol. 43, 1123–1128.
- Frink, R. J., Draper, K. G. & Wagner, E. K. (1981) Proc. Natl. Acad. Sci. USA 78, 6139-6143.
- Maitland, N. J. & McDougall, J. K. (1977) Cell 11, 233–241.
- 40. Minson, A., Bell, S. E. & Bastow, K. (1982) J. Gen. Virol. 58, 127-138.
- 41. McKnight, S. L., Kingsbury, R. C., Spence, A. & Smith, M. (1984) Cell 37, 253-262.