Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 5215–5219, September 1982
Biochemistry

In vitro transcription of the thymidine kinase gene of herpes simplex virus

(uninfected HeLa cell extract/early and late genes/transcriptional control elements/RNA polymerase II/promoter sequences)

G. SULLIVAN READ AND WILLIAM C. SUMMERS

Departments of Therapeutic Radiology, Molecular Biophysics and Biochemistry, and Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510

Communicated by Dorothy M. Horstmann, June 3, 1982

ABSTRACT **We transcribed *in vitro* a cloned 3.5-kilobase fragment of herpes simplex virus type 1 DNA that contains the gene for the viral thymidine kinase. Extracts from uninfected HeLa cells produced five *in vitro* transcripts, one of which initiated at the *in vivo* start site for the thymidine kinase mRNA (an early viral message). A second *in vitro* transcript initiated at or near the start site for a major late *in vivo* viral mRNA. The remaining three *in vitro* transcripts may correspond to minor *in vivo* mRNA species. Sequences similar to the "T-A-T-A" and "C-A-A-T" boxes, which may be involved in the control of transcription of a variety of viral and cellular genes, were found to precede the initiation site of each of the five *in vitro* transcripts. Considerable overlap of transcription units was observed.**

The genome of herpes simplex virus (HSV) is known to contain a gene for thymidine kinase (TK; ATP:thymidine 5′-phosphotransferase, EC.2.7.1.21) (1–3). The HSV TK gene is expressed during lytic infections and in cells that have been "biochemically transformed" from the TK^- to the TK^+ phenotype by exposure to either UV-inactivated virus (4, 5) or to purified DNA fragments containing the HSV TK gene (6). During lytic infections, expression of the HSV TK gene is apparently regulated transcriptionally, in part by other virus-specified proteins (7–10). The enzyme is expressed as an "early" virus function, which means that the onset of synthesis of TK mRNA is dependent upon the prior synthesis of one or more "immediate–early" viral polypeptides (7–10). Moreover, an immediate–early protein, ICP4, is continuously required for maintained synthesis of high levels of early and late mRNAs (10).

In contrast to the situation during lytic infections, the HSV TK gene in transformed cells can be expressed in the apparent absence of other viral gene products. The TK gene is transcribed from the same promoter in lytic infections and in at least four stable TK-transformed cell lines as shown by the finding that the 5′ termini of the TK mRNAs are identical (ref. 11; unpublished data). Even though the resident HSV TK gene in many transformed clones no longer needs ICP4 for its expression, it remains responsive to viral control molecules because superinfection of the cells with TK^- virus leads to enhanced expression of the resident HSV TK gene (12, 13).

A number of factors make the HSV TK gene well suited for studies of eukaryotic gene expression. The gene has been cloned (14–16), the sequence has been determined (17, 18), and the structure of the TK mRNA has been characterized (17, 18), including a mapping of its 5′ terminus (17, 18). Furthermore, the availability of convenient methods for selecting both TK^+ and TK^- virus and cells (3–6) makes possible the isolation of mutants with altered expression of the TK gene. In an effort to better understand the regulation of TK gene expression, we have studied the *in vitro* transcription of the HSV TK gene utilizing the *in vitro* transcription system developed by Manley *et al.* (19). In this paper we report the characterization of the *in vitro* transcripts arising from a 3.5-kilobase (kb) *Bam*HI restriction fragment of HSV type 1 (HSV-1) DNA that contains the HSV TK gene.

MATERIALS AND METHODS

Preparation of DNA. The plasmid pX1 was constructed by insertion of the *Bam*HI Q fragment of HSV-1 CL101 DNA into the unique *Bam*HI site of pBR322 (14). The plasmids pX1 and pBR322 were maintained in the *Escherichia coli* strain HB101 (20). Supercoiled plasmid DNA was prepared from clarified lysates of chloramphenicol-treated cells by banding on two successive CsCl gradients containing ethidium bromide (21, 22). DNA was digested with restriction endonucleases under conditions suggested by the supplier (New England BioLabs). After digestion, the DNA was extracted twice with phenol/chloroform, 1:1 (vol/vol), and twice with chloroform, precipitated twice with ethanol, and resuspended in 10 mM Tris·HCl, pH 7.9/1 mM EDTA.

Purification of DNA Restriction Fragments. DNA fragments were separated by electrophoresis on 1% agarose gels and then electroeluted from slices of the ethidium bromide-stained gel (23). After two extractions with 2-butanol (23), the DNA in 2 mM Tris·HCl, pH 8/5 mM sodium acetate/1 mM EDTA was loaded onto a column of DE-52 (Whatman) equilibrated with the same buffer. The DNA fragments were eluted with 1 M NaCl (23), precipitated twice with ethanol, and resuspended in 10 mM Tris·HCl, pH 7.9/1 mM EDTA.

Preparation of HeLa Cell Extracts. HeLa cells were grown in RPMI 1640 medium (GIBCO) containing 10% calf serum. *In vitro* transcription extracts were prepared from 2–3 liters of uninfected HeLa cells grown to a density of $4–5 \times 10^5$ cells per ml by the procedure described by Manley *et al.* (19).

***In Vitro* Transcription and Purification of the RNA Products.** The 25-μl transcription reactions contained 15 μl of uninfected HeLa cell extract, 12 mM Hepes (pH 7.9), 60 mM KCl, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 1.3 mM dithiothreitol, 10% glycerol, 20 μM cold UTP, 10 μCi of [α-^{32}P]UTP (410 Ci/mmol, Amersham-Searle; 1 Ci = 3.7×10^{10} becquerels), 10 mM creatine phosphate, 50 μM each of cold ATP, CTP, and GTP, and, except where indicated, 50 μg of DNA per ml. Reactions were incubated at 30°C for 1 hr (19), at which time they were terminated by the addition of 250 μl of 8 M urea/1% $NaDodSO_4$/10 mM Tris chloride, pH 7.9/10 mM EDTA. Reaction mixtures were then extracted twice with phenol/chlo-

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, HSV type 1; kb, kilobase.

roform, 1:1 (vol/vol), and twice with chloroform, precipitated three times with ethanol, and resuspended in 30 mM NaH_2PO_4–Na_2HPO_4, pH 6.8/0.2% sarkosyl/1 mM EDTA (19). The *in vitro* transcripts were then denatured by incubation at 50°C for 1 hr in 10 mM NaH_2PO_4–Na_2HPO_4, pH 6.8/50% dimethyl sulfoxide/1 M glyoxal/0.067% sarkosyl/0.33 mM EDTA after which they were electrophoresed on 1.4% agarose gels cast in 10 mM NaH_2PO_4–NA_2HPO_4, pH 6.8 (24). Molecular weight markers were restriction endonuclease fragments of pX1 and pBR322 that had been denatured with glyoxal as described above. After electrophoresis, the gels were stained with ethidium bromide to visualize the size markers, photographed, and dried. Dried gels were exposed to Kodak XAR-5 x-ray film in the presence of intensifying screens.

S1 Nuclease Mapping. Unlabeled *in vitro* transcripts were purified as described above from standard reactions containing *Bam*HI-cleaved pX1 as the template. They were digested at 37°C for 10 min with RNase-free DNase (50 μg/ml), extracted once with phenol/chloroform and once with chloroform, and precipitated with ethanol. Single-stranded probe DNA consisted of a 131-nucleotide fragment of the TK coding strand that was 5′-end-labeled at the *Bgl* II site at nucleotide 780 and extended to the *Eco*RI site at nucleotide 650. Probe DNA was prepared by labeling the 5′ termini of the *Bgl* II–*Eco*RI subfragments of pX1 with ^{32}P (25). Single strands were separated by electrophoresis of denatured DNA on a 5% polyacrylamide gel (25), located by autoradiography, electroeluted from gel slices, and precipitated with ethanol (17). Each hybridization reaction contained 3 ng of probe DNA, *in vitro* transcripts from a 125-μl transcription reaction, and rabbit liver ribosomal RNA (850 μg/ml) in 20 μl of 80% formamide/0.4 M NaCl/40 mM Pipes, pH 6.5/1 mM EDTA. Samples initially were denatured by heating at 80°C for 10 min and then hybridized at 42°C for 18 hr. The samples were then diluted with 200 μl of 30 mM Na acetate, pH 4.5/0.25 M NaCl/3 mM $ZnSO_4$/5% glycerol and digested with S1 nuclease (1,000 units/ml) for 30 min at 25°C, followed by 15 min at 4°C. S1 nuclease-resistant hybrids were then extracted with phenol/chloroform, precipitated with ethanol, resuspended in 80% formamide/10 mM NaOH/1 mM EDTA, and electrophoresed on an 8% polyacrylamide gel (25). Appropriate guidelines specified by the National Institutes of Health were followed for all manipulation of recombinant DNA.

RESULTS AND DISCUSSION

Analysis of Transcripts. Fig. 1*a* is a map of restriction enzyme sites located on the 3.5-kb *Bam*HI Q fragment of HSV-1 DNA (14), which maps between coordinates 0.29 and 0.31 on the P orientation of the HSV-1 genome (26) and is known to contain the HSV-1 TK gene (6). The base sequence is known for a region of DNA extending leftward approximately 2,500 nucleotides from the right *Bam*HI terminus of the fragment (17, 18; unpublished data). Two major and several minor species of polyadenylylated RNA that hybridize to the *Bam*HI Q fragment have been isolated from the cytoplasm of HSV-1 infected cells and partially characterized (11, 17, 18; unpublished data). The 5′ terminal sequences of both major *in vivo* mRNA species (Fig. 1*b*) have been determined. The TK message is unspliced and is transcribed to the left from approximately nucleotide 727 to somewhere between nucleotides 2,030 and 2,034 (17, 18). Transcription of the major 4-kb late mRNA, designated 4kbL, begins approximately at nucleotide 2,055 and proceeds leftward to a point beyond the end of the *Bam*HI Q fragment. At present it is unclear whether this mRNA is spliced, and the polypeptide which it encodes is unknown. Nevertheless, accumulation in the cytoplasm of the 4kbL mRNA is delayed relative to that of

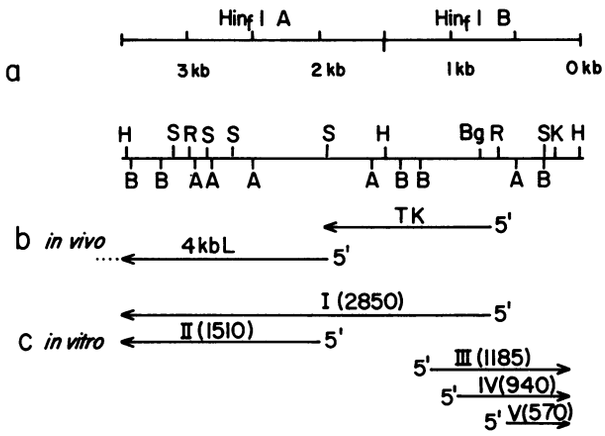

FIG. 1. Map of transcripts from *Bam*HI Q fragment of HSV-1 DNA. (*a*) Restriction enzyme sites. Symbols for restriction enzymes are: H, *Hin*fI; B, *Bgl* I; K, *Kpn* I; R, *Eco*RI; A, *Alu* I; Bg, *Bgl* II; S, *Sma* I. Nucleotide 1 of the sequence reported by Wagner *et al.* (17) is located at nucleotide 318 in this figure. (*b*) Major *in vivo* mRNAs. Arrows indicate the direction (5′ to 3′) and extent of transcribed sequences. (*c*) *In vitro* transcripts. Lengths of transcripts (in nucleotides) are shown in parentheses.

TK mRNA by 2–3 hr, which suggests that it is a member of the late class of viral messages (unpublished data). The several minor species of *in vivo* mRNA appear to be transcribed to the right, beginning at points within the right half of the *Bam*HI Q fragment; however, more precise data mapping the 5′ termini of these mRNAs are not yet available.

We transcribed the *Bam*HI Q fragment of HSV-1 DNA *in vitro*, utilizing extracts of uninfected HeLa cells prepared as described by Manley *et al.* (19). A summary of the *in vitro* transcripts that were observed is shown in Fig. 1*c*. In experiments with a variety of viral and mammalian cell genes as templates in this transcription system, *in vitro* transcripts were produced that had capped 5′ termini identical to those of *in vivo* mRNAs; thus, transcription seems to initiate accurately *in vitro* (19). However, efficient splicing and polyadenylylation of these primary transcripts is generally not observed. This results in the generation of runoff transcripts when linear DNA fragments are used as templates. Our strategy for mapping *in vitro* transcripts was to transcribe different restriction enzyme-generated subfragments of *Bam*HI Q and to measure the sizes of the resulting runoff transcripts by electrophoresis on agarose gels.

When linear *Bam*HI-cleaved pX1 DNA was used as a template, six prominent *in vitro* transcripts were observed (Fig. 2). One of these, with a length of 2,100 ± 150 nucleotides, also was observed when *Bam*HI-cleaved pBR322 was used as a template and was interpreted as arising from transcription of pBR322 sequences. The other five transcripts were specific to the reaction containing *Bam*HI-cut pX1, strongly indicating that they arose from transcription of HSV sequences. Labeled I through V in Fig. 2, these *in vitro* transcripts had lengths of 2,850 ± 200, 1,510 ± 100, 1,185 ± 75, 940 ± 50, and 570 ± 50 nucleotides, respectively. Production of the *in vitro* transcripts was dependent upon the presence of added DNA (Fig. 2, lane 1) and was strongly inhibited by the presence of α-amanitin (1 μg/ml) (Fig. 2, lane 3), indicating that the transcription was due to RNA polymerase II. Qualitatively similar patterns of *in vitro* transcription were seen over a range of DNA concentrations from 50 to 200 μg/ml (unpublished data).

In order to map more finely the *in vitro* transcripts, we next transcribed the isolated *Hin*fI-generated subfragments of *Bam*HI fragment Q. The subfragment *Hin*fI A extends from nucleotide 1,535 to within ≈20 bases of the left end of *Bam*HI

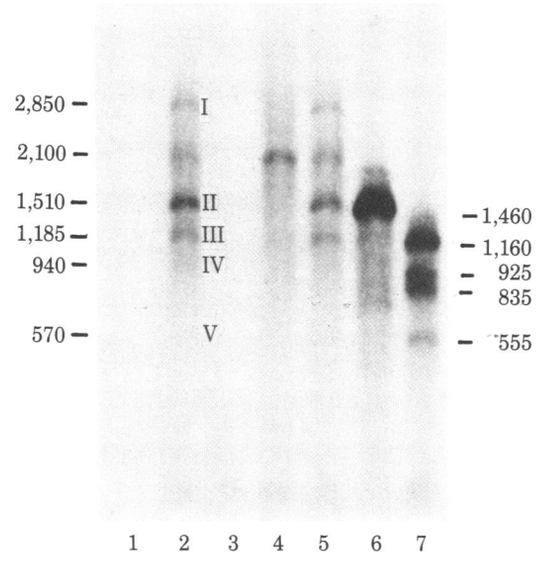

FIG. 2. Autoradiogram of gel of *in vitro* transcripts. RNA was synthesized in standard reactions that contained the following DNAs as templates. Lanes: 1, no added DNA; 2 and 5, *Bam*HI-cleaved pX1 (100 μg/ml); 3, *Bam*HI-cleaved pX1 (100 μg/ml) with α-amanitin (1 μg/ml); 4, *Bam*HI-cleaved pBR322 (100 μg/ml); 6, *Hin*fI subfragment A (100 μg/ml); 7, *Hin*fI subfragment B (100 μg/ml). The RNA products were extracted, denatured with glyoxal, and electrophoresed on a 1.4% agarose gel. The sizes of the RNA species are shown in nucleotides.

Q, whereas the subfragment *Hin*fI B extends from nucleotide 11 to nucleotide 1,535. As can be seen in Figs. 2 and 3*a*, when *Hin*fI A was used as a template, one major *in vitro* transcript was observed with a length of 1,460 ± 100 nucleotides. To establish the polarity of this transcript, we used as templates *Hin*fI A DNA that had been further cleaved with *Bgl* I, *Eco*RI, or *Alu* I, which shorten the fragment from the left end by approximately 300, 500, and 1,000 nucleotides, respectively. The resulting *in vitro* transcripts were shortened from 1,460 nucleotides to 1,150 ± 75, 970 ± 50, and 490 ± 50 nucleotides, respectively (Fig. 3*a*). This established the existence of a leftward *in vitro* transcript initiating 490 ± 50 nucleotides to the right of the *Alu* I site at nucleotide 2,567. Thus, *in vitro* transcription initiated at or near the start site of the *in vivo* 4kbL mRNA.

When *Hin*fI subfragment B was used as a template, four prominent *in vitro* transcripts were observed with lengths of 1,160 ± 75, 925 ± 50, 835 ± 50, and 555 ± 50 nucleotides (Figs. 2 and 3*a*). The transcripts of 1,160, 925, and 555 nucleotides corresponded closely in size to the *in vitro* transcripts III, IV, and V that were observed when *Bam*HI-cleaved pX1 was transcribed. This indicated that RNA species III, IV, and V arose from transcription of sequences that were contained almost entirely within *Hin*fI B, probably from runoff transcription from the right end of *Bam*HI Q. The 835-nucleotide transcript was a new species that was not observed when *Bam*HI-cleaved pX1 was used as the template. This band was consequently interpreted as a runoff transcript from the left end of *Hin*fI B.

This conclusion was corroborated by the observed production of a 1,200 ± 50 nucleotide transcript when *Bam*HI/*Sma* I-cleaved pX1 was used as a template (Fig. 3*b*, lane 1). *Sma* I cleaves the *Bam*HI Q fragment at nucleotides 285 and 1,941 and at several other locations (Fig. 1*a*) but does not cleave pBR322. Initiation of leftward transcription at the TK mRNA start site should have produced runoff transcripts of approximately 808, 1,214, and 2,791 nucleotides when *Hin*fI subfragment B, *Bam*HI/*Sma* I-cleaved pX1, and *Bam*HI-cleaved pX1 were used as templates, respectively. These predictions agreed with the observed sizes of the *in vitro* transcripts to within the limits of resolution of the gels. Transcription of *Bam*HI/*Sma* I-cleaved pX1 also resulted in the production of transcripts of 910 ± 50 and 650 ± 50 nucleotides. The 910-nucleotide transcript was consistent with initiation of a rightward runoff transcript at or near nucleotide 1,185 (1,185–285), whereas the 650 nucleotide band was interpreted as arising from the comigration of a rightward runoff transcript initiating at or near nucleotide 940 and

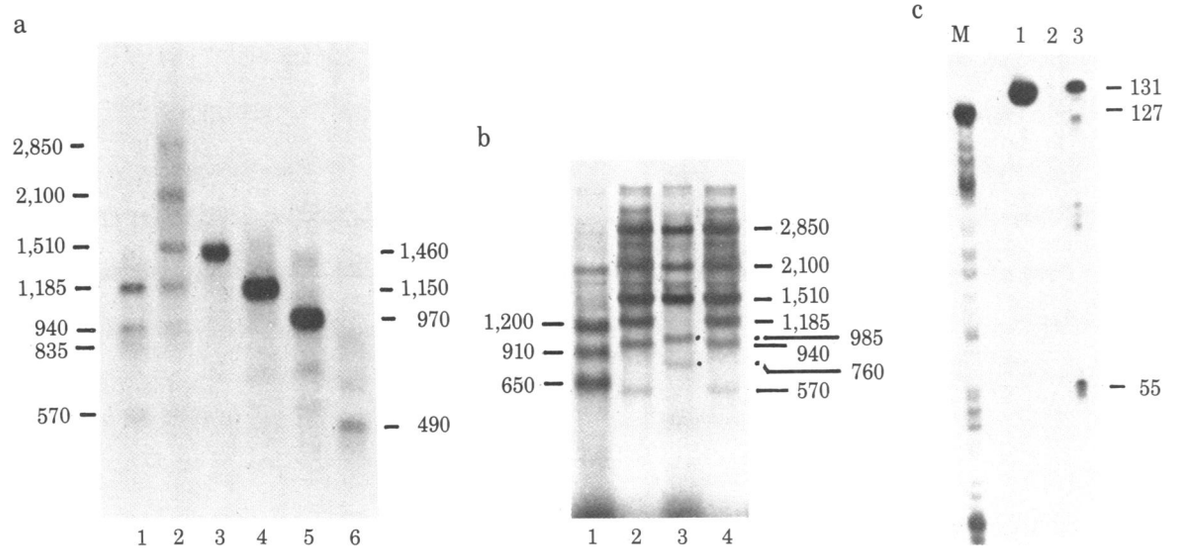

FIG. 3. Analysis of *in vitro* transcripts. (*a* and *b*) Labeled *in vitro* transcripts synthesized in standard reactions were extracted, denatured, and electrophoresed on a 1.4% agarose gel. (*a*) Reaction templates. Lanes: 1, *Hin*fI subfragment B (100 μg/ml); 2, *Bam*HI-cleaved pX1 (100 μg/ml); 3, *Hin*fI subfragment A (100 μg/ml); 4, *Bgl* I-cleaved *Hin*fI A (100 μg/ml); 5, *Eco*RI-cleaved *Hin*fI A (100 μg/ml); 6, *Alu* I-cleaved *Hin*fI A (100 μg/ml). (*b*) Reaction templates. Lanes: 1, *Bam*HI/*Sma* I-cleaved pX1 (50 μg/ml); 2 and 4, *Bam*HI-cleaved pX1 (50 μg/ml); 3, *Bam*HI/*Kpn* I-cleaved pX1 (50 μg/ml). (*c*) S1 nuclease protection experiment. 5′ End-labeled probe DNA was hybridized in the following reactions and then either digested with S1 nuclease or mock digested. Lanes: 1, probe with rabbit rRNA and *in vitro* transcripts, mock digested; 2, probe with rabbit rRNA, digested with S1 nuclease; 3, probe with rabbit rRNA and *in vitro* transcripts, digested with S1 nuclease. Products were electrophoresed on a sequence determination gel with a track of molecular size markers (M), with sizes shown in nucleotides.

of a truncated leftward transcript initiating at or near the start site for the *in vivo* 4kbL mRNA.

In order to determine the polarities of the *in vitro* transcripts III, IV, and V, we next transcribed *Bam*HI-cut pX1 that had been further cleaved with *Kpn* I. *Kpn* I cleaves once within the *Bam*HI Q fragment, shortening it from the right end by approximately 210 nucleotides. *Kpn* I fails to cut pBR322. *Kpn* I cleavage of the *Bam*HI-cut pX1 template resulted in the disappearance of the 1,185-nucleotide transcript III and the appearance of a transcript of 985 ± 50 nucleotides that migrated fractionally more slowly than the 940-nucleotide transcript IV in the adjoining lane. These data indicated that transcript III was a runoff transcript from the right end of *Bam*HI Q. *Kpn* I cleavage of the template resulted in the disappearance of the 940-nucleotide band and the appearance of a new transcript of length 760 ± 50 nucleotides, indicating that a transcript of about 970 (760 + 210) ± 50 nucleotides was transcribed to the right. Within experimental limits, this transcript coincides with transcript IV and implies the existence of an initiation site at nucleotide 940 ± 50. *Kpn* I cleavage also resulted in the disappearance of the 570-nucleotide transcript V, indicating that it was a runoff transcript from the right end of *Bam*HI Q. As would be expected, *Kpn* I cleavage of the *Bam*HI-cut pX1 template did not affect production of the leftward 2,850- and 1,510-nucleotide transcripts or of the 2,100-nucleotide transcript of pBR322 sequences.

In order to finely map the 5′ terminus of *in vitro* transcript I, we performed an S1 nuclease protection experiment utilizing cold *in vitro* transcripts and a single-stranded DNA probe consisting of a 131-nucleotide fragment of the TK coding strand, 5′ end-labeled at the *Bgl* II site located 54–56 nucleotides downstream from the *in vivo* TK mRNA start site and extending to the *Eco*RI site 77 nucleotides upstream from the start site (Fig. 3*c*). Hybridization of the probe with the *in vitro* transcripts followed by digestion of the resulting hybrids with S1 nuclease resulted in the production of a major labeled S1 nuclease-resistant DNA fragment 54–56 nucleotides in length. An identical result has been obtained for *in vivo* TK mRNA (11, 18). Thus, a significant amount of *in vitro* transcription initiated at, or within several nucleotides of, the *in vivo* start site for TK mRNA. An appreciable amount of full-length probe also was protected from S1 nuclease digestion, indicating the existence of either end-to-end transcripts of the entire *Bam*HI Q fragment or of some transcripts initiating within the fragment but to the right of the *Eco*RI site at nucleotide 650. The existence of S1 nuclease-resistant bands of length 90, 100, and 120 nucleotides indicated the existence of *in vitro* transcripts initiating at or near nucleotides 690, 680, and 660.

Promoter Sites. Examination of the sequences flanking the 5′ ends of a large number of genes that are transcribed by RNA polymerase II has revealed that approximately 30 nucleotides upstream from the transcriptional start site of many genes is located a derivative of the Goldberg–Hogness or "T-A-T-A" sequence (27, 28), whereas 70–80 bases upstream from the mRNA start site is often found a derivative of the "C-A-A-T" sequence (28, 29). In addition, the first transcribed base of many mRNAs is an A surrounded by pyrimidines (28, 30). Examination of the sequences of the HSV *Bam*HI Q fragment that flank the start sites of the five *in vitro* transcripts revealed in each case the presence of a derivative of the canonical T-A-T-A box preceded approximately 45 bases upstream by a derivative of the C-A-A-T box. These sequences are listed in Fig. 4 along with the sizes of the transcripts that would be predicted if transcription initiated at an A residue located approximately 31 nucleotides downstream from the first base of the T-A-T-A box and then continued to the end of the *Bam*HI Q fragment. The experimentally determined size of each *in vitro* transcript corresponded closely to the size of a transcript predicted from the DNA sequence.

Regulation of Transcription Units. Our data suggest that there may be considerable overlap of transcription units within the *Bam*HI Q fragment of HSV-1 DNA. The *in vivo* TK mRNA contains two copies of the sequence A-A-U-A-A-A located shortly before the 3′ end of the transcribed portion of the message (17). This hexanucleotide may be part of the polyadenylylation signal for many mRNAs (31, 32). Therefore, it is interesting that the second copy of this hexanucleotide is part of the derivative of the T-A-T-A box preceding the *in vivo* 4kbL mRNA. Thus, any 5′ control sequences flanking the transcribed portion of this gene are necessarily within the transcribed portion of the TK gene. In addition, the 5′ termini of *in vitro* transcripts III and IV overlap the 5′ end of the TK transcript,

Observed in vitro transcripts	Predicted in vitro transcripts	CAAT box		TATA box		Predicted RNA start site
Consensus		GG(C/T)CAATCT	40 bp	TATA(A/T)A(A/T)	14-22 bp	Y---YAYYYYY
2850 (L)	2791 (L)	GGCGAATTC (6/9)	47 bp	CATATTA (5/7)	19 bp	CGAACACCGAG (727)
1510 (L)	1463 (L)	AGACAATAC (5/9)	39 bp	AATAAAA (6/7)	21 bp	TGTTCATAAAC (2055)
1185 (R)	1195 (R)	TGTGAA CT (6/9)	48 bp	CATAAGG (4/7)	24 bp	TTGTCATTACC (1195)
940 (R)	970 (R)	GGCCACCAG (5/9)	43 bp	TATATAA (7/7)	14 bp	TGGGCATTTTC (970)
570 (R)	560 (R)	CGCCAATGA (6/9)	48 bp	AATATAT (6/7)	18 bp	CCAGCAAACGC (560)

FIG. 4. Promoter-like sequences located within the *Bam*HI Q fragment near *in vitro* initiation sites. Derivatives of the consensus C-A-A-T and T-A-T-A boxes are shown along with the sequences surrounding the predicted mRNA start sites. Y designates pyrimidine. The location of the first base of each C-A-A-T and T-A-T-A box is shown in parentheses beneath the sequence along with the number of bases that match the consensus sequence. Consensus sequences were obtained from ref. 28. The predicted and observed directions of transcription are indicated in parentheses by an R for rightward and an L for leftward. Each predicted start site is at an A located approximately 31 nucleotides downstream from the first base of the T-A-T-A box. The location of each predicted start site is indicated in parentheses below the sequence. Predicted start sites at nucleotides 727 and 2,055 are those observed for the TK and 4kbL *in vivo* mRNAs, respectively. Predicted and observed sizes of the transcripts are for reactions containing *Bam*HI-cleaved pX1 as the template.

whereas the C-A-A-T boxes for transcript V and the TK mRNA share nucleotide pairs. One can imagine ways in which the overlapping of transcription units might be important in coordinating the expression of different genes within this region of the genome.

The HSV TK gene apparently can be recognized and accurately transcribed by the unmodified cellular transcriptional apparatus in biochemically transformed cells and after microinjection into frog oocytes. In addition, we observed that extracts from uninfected cells initiated *in vitro* transcription at or near the start sites for both early (TK) and late (4kbL) viral genes. Why, then, during lytic infections is production of any detectable TK mRNA dependent upon the prior synthesis of immediate–early viral proteins (9)? Although it is well established that the production of early and late viral mRNAs is subject to positive regulation by an immediate–early protein, the available data are consistent with the possibility that transcription of early and late genes is also subject to some form of negative control when these genes are introduced into the cell as part of a virion. After infection, complete uncoating of early and late genes might be dependent upon the prior synthesis of immediate–early proteins. Alternatively, immediate–early proteins might be required in order to alter the conformation of or to chemically modify the virion DNA or to alleviate inhibition of early and late gene transcription by virion proteins. In any case, early and late genes that are not introduced into a cell as part of a virus particle may be transcribable, at least at a low level, in the absence of immediate–early proteins.

We thank R. Millette, M. J. Wagner, J. A. Sharp, W. P. Summers, U. DasGupta, P. Liebowitz, and D. Pintel for helpful discussions. We thank J. Manley for help at the start of this work. This work was supported by grants from the U.S. Public Health Service (CA 13515 and CA 16038). G.S.R. was supported by a training grant from the U.S. Department of Health and Human Services (5 T32 CA 09159).

1. Dubbs, D. R. & Kit, S. (1964) *Virology* **22**, 493–505.
2. Honess, R. W. & Watson, D. H. (1974) *J. Gen. Virol.* **22**, 171–185.
3. Summers, W. P., Wagner, M. & Summers, W. C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4081–4084.
4. Munyon, W., Kraiselburd, E., Davis, D. & Mann, J. (1971) *J. Virol.* **7**, 813–820.
5. Munyon, W., Buchsbaum, R., Paoletti, E., Mann, J., Kraiselburd, E. & Davis, D. (1972) *Virology* **49**, 683–689.
6. Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C. & Axel, R. (1977) *Cell* **11**, 223–232.
7. Honess, R. W. & Roizman, B. (1974) *J. Virol.* **14**, 8–19.
8. Preston, C. M. (1979) *J. Virol.* **29**, 275–284.
9. Leung, W.-C., Dimock, K., Smiley, J. R. & Bacchetti, S. (1980) *J. Virol.* **36**, 361–365.
10. Watson, R. J. & Clements, J. B. (1980) *Nature (London)* **285**, 329–330.
11. McKnight, S. L., Gavis, E. R., Kingsbury, R. & Axel, R. (1981) *Cell* **25**, 385–398.
12. Lin, S. & Munyon, W. (1974) *J. Virol.* **14**, 1199–1208.
13. Leiden, J. M., Buttyan, R. & Spear, P. G. (1976) *J. Virol.* **20**, 413–424.
14. Enquist, L. W., Vande Woude, G. F., Wagner, M., Smiley, J. R. & Summers, W. C. (1979) *Gene* **7**, 335–342.
15. Colbere-Garapin, F., Chousterman, S., Horodniceanu, F., Kourilsky, P. & Garapin, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3755–3759.
16. Wilkie, N. M., Clements, J. B., Boll, W., Mantei, N., Lonsdale, D. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 859–877.
17. Wagner, M. J., Sharp, J. A. & Summers, W. C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1441–1445.
18. McKnight, S. L. (1980) *Nucleic Acids Res.* **8**, 5949–5964.
19. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
20. Boyer, H. & Rowland-Dussiox, D. (1969) *J. Mol. Biol.* **41**, 459–472.
21. Clewell, D. B. & Helinski, D. R. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 1159–1166.
22. Clewell, D. B. (1972) *J. Bacteriol.* **110**, 667–676.
23. Smith, H. O. (1980) *Methods Enzymol.* **65**, 371–380.
24. Carmichael, G. G. & McMaster, G. K. (1980) *Methods Enzymol.* **65**, 380–391.
25. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
26. Locker, H. & Frenkel, N. (1979) *J. Virol.* **32**, 429–441.
27. Gannon, F., O'Hare, K., Perrin, F., LePennec, J. P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B. & Chambon, P. (1979) *Nature (London)* **278**, 428–434.
28. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
29. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Cell* **21**, 653–668.
30. Busslinger, M., Portmann, R., Irminger, J. C. & Birnstiel, M. L. (1980) *Nucleic Acids Res.* **8**, 957–977.
31. Proudfoot, N. J. & Brownlee, G. G. (1974) *Nature (London)* **252**, 359–362.
32. Fitzgerald, M. & Shenk, T. (1981) *Cell* **24**, 251–260.