MOLECULAR AND CELLULAR BIOLOGY, Apr. 1983, p. 643–653
0270-7306/83/040643-11$02.00/0

Vol. 3, No. 4

Preparation of a "Functional Library" of African Green Monkey DNA Fragments Which Substitute for the Processing/Polyadenylation Signal in the Herpes Simplex Virus Type 1 Thymidine Kinase Gene

GEORGE M. SANTANGELO AND CHARLES N. COLE*

Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut 06510

Received 27 October 1982/Accepted 14 January 1983

Fragments of African green monkey (*Cercopithecus aethiops*) DNA (3.5 to 18.0 kilobases) were inserted downstream from the thymidine kinase (TK, *tk*) coding region in pTK206/SV010, a gene construct which lacks both copies of the hexanucleotide 5'-AATAAA-3' and contains a simian virus 40 origin of replication, allowing it to replicate in Cos-1 cells. No polyadenylated *tk* mRNA was detected in Cos-1 cells transfected by pTK206/SV010. The ability of simian DNA fragments to restore *tk* gene expression was examined by measuring the incorporation of [^{125}I]iododeoxycytidine into DNA in Cos-1 cells transfected by pTK206/SV010 insertion derivatives. *tk* gene expression was restored by the insertion in 56 of the 67 plasmids analyzed, and the level of expression equaled or exceeded that obtained with the wild-type *tk* gene in 30 of these. In all plasmids examined that showed restoration of *tk* gene expression, polyadenylated *tk* mRNA of discrete size was detected. The sizes of these *tk* mRNAs were consistent with the existence of processing and polyadenylation signals within the inserted DNA fragments. The frequency with which inserted fragments restored *tk* gene expression suggests that the minimal signal for processing and polyadenylation is a hexanucleotide (AAUAAA or a similar sequence). LTK$^-$ cells were biochemically transformed to TK$^+$ with representative insertion constructs. pTK206/SV010 transformed LTK$^-$ cells at a very low frequency; the frequency of transformation with insertion derivatives was 40 to 12,000 times higher.

The hexanucleotide 5'-AAUAAA-3' is an essential element in the signal specifying processing and polyadenylation in most eucaryotic mRNAs (19, 40). This hexanucleotide, or one very closely related to it, is located 10 to 30 base pairs (bp) upstream from the polyadenylate [poly(A)] addition site in most mRNAs from higher eucaryotes (4, 40); deletion of this hexanucleotide from the simian virus 40 (SV40) late (19) or herpes simplex virus type 1 (HSV) thymidine kinase (TK, *tk*; see below) (16) transcription units results in loss of the ability to form functional mRNAs.

Although usually confined to 3' noncoding regions of mRNAs, the AAUAAA hexanucleotide is found within the SV40 early (41) and the adenovirus type 12 E1A (37) transcription units; in neither case are mRNAs detected with 3' ends immediately downstream from the hexanucleotide. This indicates that either additional elements are required to constitute a functional processing and polyadenylation signal or mechanisms exist to prevent the use of these internal AAUAAAs.

We reported previously (16) that processing and polyadenylation signals from SV40 and polyoma virus DNAs could substitute for the HSV *tk* processing and polyadenylation signal. A wide range of levels of *tk* mRNA was seen in Cos-1 cells transfected with a series of HSV *tk* gene constructs containing different processing and polyadenylation signals. This could reflect differences in both the efficiency with which different signals are recognized and the stability of the different mRNA species produced. Polyadenylated [poly(A)$^+$]*tk* mRNA production also occurred when the entire 3' untranslated region of the HSV *tk* gene was replaced by an 88-bp fragment of SV40 DNA derived from the middle of the SV40 early transcription unit containing the hexanucleotide AAUAAA. One possible explanation for this observation is that an AAUAAA alone is sufficient to signal processing and polyadenylation at a nearby downstream site.

In this report, we describe the construction and preliminary characterization of a series of *tk* gene constructs prepared by inserting fragments

of African green monkey (*Cercopithecus aethiops*) DNA at the 3′ end of the HSV *tk* gene, from which both copies of the AAUAAA hexanucleotide and all downstream information had been deleted. The inclusion of an SV40 origin of DNA replication in these constructs permitted their replication in Cos-1 cells and analysis of the transient expression of the HSV *tk* gene (20, 32). This "functional 3′-end regulatory signal library" was prepared to study the following questions. Could *tk* gene expression be restored if fragments of cellular DNA were inserted downstream from *tk* coding sequences in a construct lacking a processing and polyadenylation signal and incapable of *tk* gene expression? If so, with what frequency would functional 3′-end regulatory signals be found? Would signals be isolated that restored *tk* gene expression without signaling polyadenylation? What range in levels of gene expression would be obtained with cellular signals? How far into fragments of cellular DNA would 3′-end regulatory signals be located? Would all signals obtained function in both simian and murine cells?

We screened the functional library for the presence of active *tk* genes by measuring the incorporation of [^{125}I]iododeoxycytidine (IdC) into DNA in Cos-1 cells transfected with DNA from different members of the library. The level of [^{125}I]IdC incorporation into DNA was significantly above the background in 83% of the clones, and in 45% incorporation occurred at a level equal to or greater than that with the wild-type HSV *tk* gene. When *tk* mRNA production was analyzed, a high level of TK was always associated with the synthesis of poly(A)$^+$ mRNA. The sizes of these *tk* mRNAs were consistent with the existence of processing and polyadenylation signals within the inserted monkey DNA fragments.

In addition, we measured the frequency of transformation of LTK$^-$ cells to TK$^+$ with the wild-type HSV *tk* gene and the series of *tk* gene constructs containing the processing and polyadenylation signals from SV40 and polyoma virus in place of the *tk* signal (16). We observed a direct correlation between the level of *tk* mRNA produced in Cos-1 cells ("transient expression") and the frequency with which these constructs transformed LTK$^-$ cells to TK$^+$. Transformation frequencies were also measured for representative members of the functional library. Whereas the TK gene construct lacking a processing and polyadenylation signal (pTK206/SV010; see Fig. 1B) transformed LTK$^-$ cells at an extremely low frequency (20 colonies per pmol of *tk* gene), every member of the functional library examined transformed LTK$^-$ cells at a higher frequency, ranging from 40 to 12,000 times that of pTK206/SV010. Higher frequencies of transformation generally were obtained with constructs that directed the biosynthesis of poly(A)$^+$ *tk* mRNA and the incorporation of more [^{125}I]IdC into DNA in Cos-1 cells.

MATERIALS AND METHODS

Usage of abbreviations for TK. We follow the convention of Ostrander et al. (36), which uses the abbreviation TK for thymidine kinase protein, TK$^+$ and TK$^-$ for the phenotype of cell lines, and *tk* for the genotype or when referring to nucleic acid (DNA, mRNA) coding for thymidine kinase.

Enzymes and chemicals. Enzymes were purchased from New England Biolabs, Beverly, Mass., PL Biochemicals, Milwaukee, Wis., Boehringer Mannheim, Inc., New York, N.Y., and Bethesda Research Laboratories, Bethesda, Md. Some restriction endonucleases were a generous gift from Richard Wang, Yale University. α-^{32}P-labeled deoxynucleotide triphosphates (>800 Ci/nmol or 2,000 to 3,000 Ci/mol) were purchased from Amersham Radiochemical Centre, England. [^{125}I]IdC was purchased from New England Nuclear Corp., Boston, Mass. Oligodeoxythymidylate-cellulose was purchased from Collaborative Research, Waltham, Mass. Tetrahydrouridine was a generous gift from Bill Summers, Yale University.

Bacteria and plasmids. Many of the plasmids used in these studies have been described previously (16) and are shown in Fig. 1. Established methods were used for transfection of bacterial cells (29), screening of minilysates (9), and preparation of plasmid DNA (14).

Plasmid construction. pTK206SA1/SV010 through pTK206SA67/SV010 (Fig. 1I) were constructed by ligating *Bam*HI-digested, high-molecular-weight CV-1 monkey DNA to *Bam*HI-digested, phosphatase-treated pTK206/SV010 (Fig. 1B). Minilysates were prepared, digested with *Bam*HI, and examined by electrophoresis in 1% agarose gels. Of the 100 plasmids examined, 67 contained inserts of cellular DNA fragments ranging from 3.5 to 18 kilobases (kb). Two of these plasmids contained two cellular *Bam*HI fragments.

Cells and DNA. The maintenance of Cos-1 cells (20) has been described previously (16). LTK$^-$ adenine phosphoribosyltransferase-negative (APRT$^-$) cells were maintained in Dulbecco modified Eagle medium with 10% fetal calf serum (growth medium) and subcultured 1:100 once every 7 to 8 days. These cells were subcultured 1:5 and allowed to reach 80 to 90% confluency before being prepared for transformation (see below). High-molecular-weight DNA was isolated from CV-1 monkey kidney cells or LTK$^-$ APRT$^-$ cells by the method of Wigler et al. (51, 52), with the following modifications. Cells were scraped into Tris-buffered saline (TS; 25 mM Tris-chloride [pH 7.5], 137 mM NaCl, 5 mM KCl, 0.6 mM Na_2HPO_4, 0.05 mM $MgCl_2$, 0.7 mM $CaCl_2$) and centrifuged at 1,000 × *g* for 5 min. The pellet was suspended in 1 ml of TNE (10 mM Tris-chloride [pH 8.0], 150 mM NaCl, 10 mM EDTA) per 100-mm dish, and sodium dodecyl sulfate and proteinase K were added to 0.2% and 200 μg/ml, respectively. This lysate was incubated for 12 to 18 h at 37°C and then extracted sequentially once with buffer-saturated phenol and once or twice with phenol-chloroform (1:1). After two volumes of 95% ethanol had been carefully added, DNA was spooled with a

glass rod, transferred to a clean tube, and dissolved in TEN (10 mM Tris-chloride [pH 7.5], 1 mM EDTA, 10 mM NaCl). DNA sequence analysis was performed by the method of Maxam and Gilbert (30).

Transformation of LTK$^-$ APRT$^-$ cells. Biochemical transformation was done by the procedure of Wigler et al. (51) with the following modifications. One day before transformation, cells were seeded at 5×10^5 cells per 100-mm dish. A solution (2× $CaCl_2$-DNA) was prepared that contained high-molecular-weight carrier DNA (10 μg/ml), plasmid DNA (form I), 250 mM $CaCl_2$, and 25 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid). The pH of this solution was tested routinely and found to be 7.0 ± 0.5. The 2× HEPES-buffered saline contained 280 mM NaCl, 25 mM HEPES, and 1.5 mM sodium phosphate. The pH was adjusted to 7.1 ± 0.05. After direct addition of the calcium phosphate-DNA coprecipitate to the growth medium (1 ml per plate), the cells were incubated for 18 to 24 h at 37°C. The medium was then replaced, and the cells were incubated for an additional 18 to 24 h before the medium was replaced with medium containing hypoxanthine (14 μg/ml), aminopterin (0.2 μg/ml), and thymidine (4 μg/ml) (HAT) and 10% fetal calf serum.

Transfection of Cos-1 cells with DEAE-dextran. Confluent dishes of Cos-1 cells were subcultured 1:4 (approximately 2×10^6 cells per 100-mm dish) 24 h before transfection by a modification of the method of Sompayrac and Danna (44). The next day, the growth medium was removed, the cells were washed twice with TS, and the DNA–DEAE-dextran mixture was added. This mixture was prepared by diluting plasmid DNA (unless stated otherwise, 1 μg of DNA purified in CsCl-ethidium bromide gradients) in 0.75 ml of TS per dish, followed by addition of 0.25 ml of a 2-mg/ml solution of DEAE-dextran (molecular weight, 500,000; Pharmacia Fine Chemicals, Piscataway, N.J.) per dish. This mixture (1 ml/100-mm dish) was added to the cells, which were then incubated for 4 h at 33°C. The dish was tilted every 30 to 45 min to ensure constant exposure of the cells to the DNA–DEAE-dextran mixture. After 4 h, the inoculum was removed by aspiration, the cells were washed once with TS, fresh growth medium containing 5% fetal calf serum was added, and the cultures were incubated at 37°C.

TK assay with [^{125}I]IdC. Cos-1 cells in 35-mm dishes were transfected as described above. After 24 h, the growth medium was replaced with growth medium (1 ml per dish) containing 0.4 to 1.0 μCi of [^{125}I]IdC and 150 μg of tetrahydrouridine per ml (46). After 24 to 48 h, the medium was removed, and 1 ml of 5% trichloroacetic acid (TCA)–1% sodium pyrophosphate per dish was added. After 10 min, each dish was washed twice with 5% TCA–1% sodium pyrophosphate, and 0.1 N NaOH (0.4 ml/dish) was added; after 10 min, the sample was transferred to a scintillation vial, acidified with 0.1 ml of 0.5 N HCl, mixed with 4 ml of Aquasol scintillation fluid (New England Nuclear), and assayed in a Beckman scintillation counter.

Analysis of DNA replication in Cos-1 cells. DNA replication in Cos-1 cells was analyzed as previously described (16, 39), except that DNA was digested with *Mbo*I endonuclease before electrophoresis.

Analysis of *tk* mRNA and densitometry. Cytoplasmic RNA was prepared as described by White et al. (50); poly(A)$^+$ RNA was obtained by chromatography on oligodeoxythymidylate-cellulose (2). RNA was analyzed by electrophoresis in 1% agarose–2.2 M formaldehyde gels (21, 27), followed by transfer to nitrocellulose (45, 47) and hybridization with ^{32}P-labeled pTK206/SV010 (16). Quantitative analysis of *tk* mRNA with S1 nuclease was performed by established procedures (7, 18). Densitometric tracings of autoradiograms were obtained with a Beckman spectrophotometer equipped for densitometry; the area under each peak was determined by triangulation.

RESULTS

In a previous report (16) we described the resection of the HSV *tk* gene from its 3' end and the isolation of a series of plasmids lacking different amounts of information downstream from the *tk* coding region. One of these, pTK206, lacked both copies of the hexanucleotide AAUAAA and retained only 27 bp of the 3' untranslated region. This 3' deletion mutant was transferred to pSV010, a vector containing the SV40 origin of DNA replication (35) and, therefore, capable of replication in Cos-1 cells. When Cos-1 cells were transfected with pTK206/SV010 (Fig. 1B), *tk* mRNA was not produced (16). *tk* gene expression was restored when DNA fragments containing the processing and polyadenylation signals from SV40 and polyoma virus were inserted into pTK206/SV010. The experiments described below were conducted to see with what frequency, and by what mechanism, *tk* gene expression could be restored when fragments of African green monkey DNA were inserted downstream from *tk* coding sequences in pTK206/SV010.

Maximal *tk* gene expression in Cos-1 cells after DEAE-dextran transfection of a wild-type *tk* gene. The detection of HSV TK in cells transfected with *tk* gene constructs requires an efficient method of introducing DNA into the cells and a sensitive assay for the presence of the HSV TK. Unlike the cellular TK enzyme, the HSV TK is a deoxypyrimidine kinase and can utilize a wide variety of deoxypyrimidines as substrate (11–13). Only the successful introduction of an active HSV *tk* gene into cells allows the incorporation of [^{125}I]IdC into the DNA (46), providing a rapid assay of *tk* gene transfection efficiency. Recently, Sompayrac and Danna (44) reported conditions that allow the transfection of 10 to 50% of CV-1 cells with SV40 DNA. We adopted this method to introduce *tk* gene constructs into Cos-1 cells. Optimal transfection conditions were determined by measuring the amount of [^{125}I]IdC incorporated into the DNA after each parameter was varied independently. The first parameter tested was the time that cells were exposed to the DNA–DEAE-dextran mixture. Maximal labeled IdC incorporation was obtained when cells were exposed to the mixture for 4 to 6 h at 33°C (Fig. 2A). Substantial

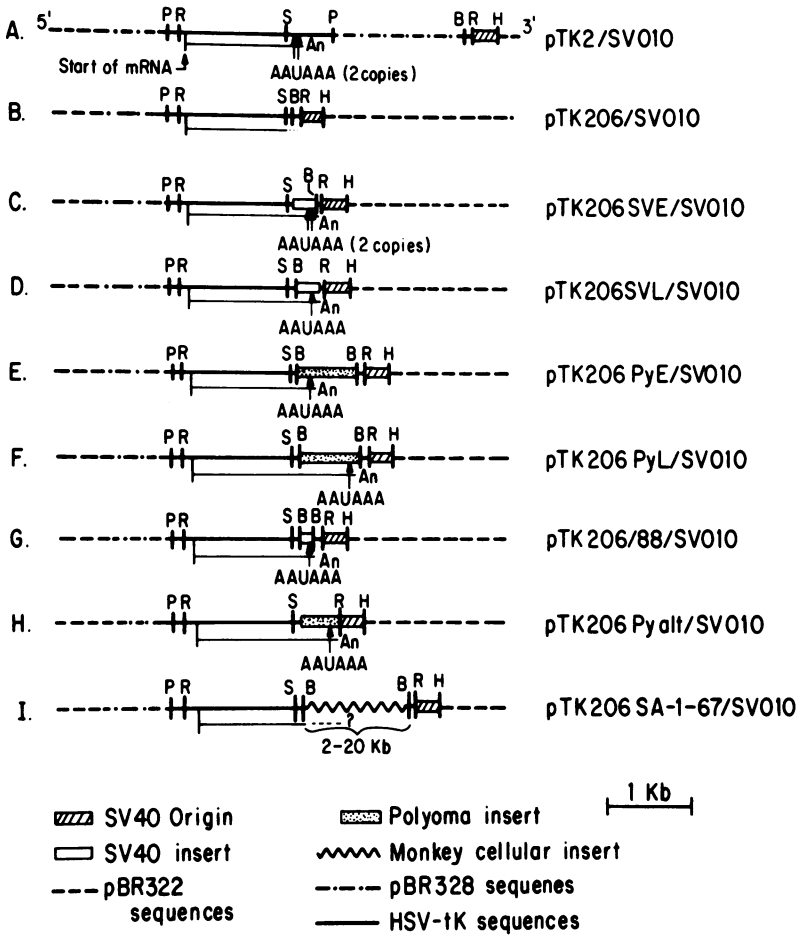

FIG. 1. Diagram of HSV *tk* gene derivatives. (A) pTK2/SV010, containing the HSV *tk* *Pvu*II subfragment, which includes the entire *tk* coding region and all essential 5' and 3' flanking sequences. (B) pTK206/SV010, a 3' deletion construct which lacks all HSV sequences distal to nucleotide 1,676 (16), including both copies of the AATAAA hexanucleotide. (C through G and I) Insertions into the *Bam*HI site of pTK206/SV010. (C) pTK206SVE/SV010, 237-bp *Bam*HI-*Bcl*I fragment of SV40 (nucleotides 2,514 through 2,751 [48], inserted so that the early viral strand is fused to the *tk* transcription unit (5'→3'; 0.189 to 0.144 map units [mu]). (D) pTK206SVL/SV010, same as C but opposite orientation (late viral strand fused to *tk* transcription unit; 5'→3', 0.144 to 0.189 mu). (E) pTK206PyE/SV010, 622-bp *Sau*3A fragment of polyoma virus (nucleotides 2,763 to 3,384 [43]), inserted so that the early viral strand is fused to the *tk* transcription unit (5'→3', 22.7 to 35.4 mu). (F) pTK206PyL/SV010, same as E but opposite orientation (late viral strand fused to *tk* transcription unit; 5'→3', 34.5 to 22.7 mu). (G) pTK206/88/SV010, 88-bp *Bst*NI fragment of SV40 (nucleotides 3,230 to 3,318 [48]), inserted with *Bam*HI linkers so that the early viral strand is fused to the *tk* transcription unit (5'→3', 0.297 to 0.280 mu). (H) pTK206Pyalt/SV010, 369-bp *Sau*3A-*Eco*RI fragment of polyoma virus (nucleotides 1,192 to 1,560 [43]), inserted into *Bam*HI-*Eco*RI-digested pTK206/SVl010 so that the early viral strand is fused to the *tk* transcription unit (5'→3', 93 to 100 mu). (I) pTK206SA1/SV010 through pTK206SA67/SV010, random *Bam*HI fragments of African green monkey DNA. SA1 through SA67 refer to the 67 different members of this library. Abbreviations: P, *Pvu*II; R, *Eco*RI; S, *Sma*I; B, *Bam*HI; H, *Hin*dIII.

detachment of cells from the dish occurred when they were exposed to the inoculum for more than 6 h. The amount of DNA required to give maximal *tk* gene expression was then optimized (Fig. 2B). Between 50 and 100 ng (0.02 to 0.04 pmol) of *tk* gene construct per 35-mm dish resulted in maximum incorporation of [^{125}I]IdC into DNA; the addition of more than 1 μg (0.4 pmol) was slightly inhibitory.

All subsequent experiments used 0.04 to 0.08 pmol of *tk* gene per 0.2 ml per 35-mm dish. These plateau levels of DNA ensured that small differences in the amount of added plasmid DNA would not result in measurable changes in the

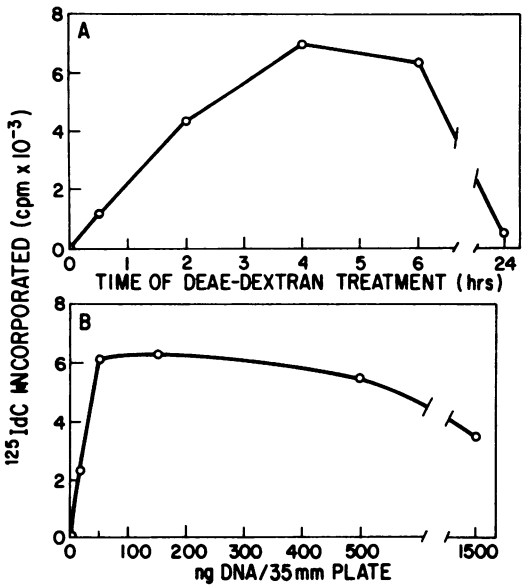

FIG. 2. Optimized transfection of Cos-1 cells with *tk* gene constructs by using DEAE-dextran. (A) Effect on *tk* gene expression of time that Cos-1 cells were exposed to the DNA–DEAE-dextran mixture. Two dishes were assayed at each time point. (B) Effect on *tk* gene expression of varying the amount of DNA in the DNA–DEAE-dextran mixture used to transfect cells. Cos-1 cells were exposed to the mixtures for 4 h.

level of *tk* gene expression. Furthermore, even with the largest plasmid examined in these experiments (22.5 kb), cells were not exposed to amounts of DNA which would have inhibited maximal *tk* gene expression.

Sensitivity of [^{125}I]IdC incorporation assay. To determine the sensitivity with which *tk* gene expression could be detected, we transfected Cos-1 cells with several of the *tk* gene constructs listed in Fig. 1. *tk* gene expression was measured in two different ways: (i) incorporation of [^{125}I]IdC into DNA in transfected cells and (ii) determination of the level of *tk* mRNA present in transfected cells by quantitative nuclease S1 mapping (7, 18) (Table 1). For method ii, total cytoplasmic RNA was extracted from cells transfected with DNA from different *tk* constructs, and S1 nuclease analysis was done. In each case, cells were cotransfected with a plasmid containing the human α-1 globin gene (32) as an internal control. The amount of *tk* mRNA present in cells transfected by each construct was quantitated by densitometry and normalized by using the amount of α-globin mRNA present in the same sample (Table 1).

Comparison of the results obtained with these two assays showed the sensitivity of the [^{125}I]IdC incorporation assay. Although *tk* mRNA could be detected by nuclease S1 analy-

TABLE 1. *tk* gene expression in cells transfected with *tk* gene constructs containing processing and polyadenylation signals from SV40 and polyoma virus

Construct	Relative [^{125}I]IdC incorporation (%) (mean ± SD)[a]	Relative *tk* mRNA levels	Transformation frequency[b]	
			Colonies/pmol	Relative frequency
pTK2/SV010	100	1.0	5,600	1.0
pTK206/SV010[c]	2 ± 0.84	0.0	20	<0.01
pTK206[d]			20	<0.01
pTK206/88/SV010	46 ± 15	1.2	1,980	0.4
pTK206SVE/SV010	173 ± 25	12.3	24,500	4.4
pTK206SVL/SV010	197 ± 20	14.8	26,000	4.6
pTK206Pyalt/SV010	25 ± 7	0.1	740	0.1
pTK206PyL/SV010	49 ± 4	0.4	11,000	2.0
pTK206PyE/SV010	114 ± 5	6.6	97,000	17.3

[a] Each construct was examined in three to eight experiments. All constructs in each experiment were analyzed in duplicate. The values were normalized to incorporation levels in cells transfected by pTK2/SV010. The background was defined as the level of incorporation in mock-transfected cells and was subtracted from every value. When cultures in 35-mm dishes received 1 μCi of [^{125}I]IdC, cultures transfected by pSV010 and pTK2/SV010 routinely incorporated 3,000 and 50,000 cpm, respectively, in a 24-h labeling period.

[b] Values are expressed as colonies per picomole of the HSV *tk* gene to normalize for the size of each construct. Values are based on 5 to 40 dishes of cells transformed with each construct. Most constructs were examined in two to four separate experiments. Relative frequency was determined by normalization to the frequency obtained with pTK2/SV010, which was assigned a value of 1.0.

[c] In every experiment, pTK206/SV010 samples contained slightly more TCA-precipitable radioactivity than pSV010 (see footnote *a*).

[d] pTK206 lacks the SV40 origin of replication and was not examined in Cos-1 cells.

sis in cells transfected with pTK206Pyalt/SV010, an exposure time of 11 days was required. Overnight exposure was sufficient to detect *tk* mRNA in cells transfected with pTK2/SV010, pTK206SVE/SV010, pTK206SVL/SV010, or pTK206PYE/SV010. The detection of TK with the IdC assay, however, was rapid, highly reproducible, and capable of detecting TK levels as low as 3% that of the wild-type *tk* gene. The actual amounts of TK present were very low; the enzyme could be detected in extracts by standard assays only in cells transfected with pTK206PyE/SV010, pTK206SVL/SV010, or pTK206SVE/SV010 (unpublished data). This may reflect inefficient translation of *tk* mRNA in these cells.

Analysis of *tk* gene expression in Cos-1 cells transfected by pTK206/SV010 derivatives containing fragments of monkey DNA inserted at the *Bam*HI site. African green monkey DNA was digested with *Bam*HI endonuclease, and the fragments were inserted into the *Bam*HI site of pTK206/SV010 (Fig. 1). These plasmids were designated pTK206SA1/SV010 through pTK206SA67/SV010. The in vivo IdC assay was used to screen them for the ability to synthesize HSV TK. Cos-1 cells were transfected with each plasmid (as minilysate DNA); pTK206/SV010, pTK206SVL/SV010, pTK2/SV010, and pTK206PYalt/SV010 were used as controls. Based on the levels of [^{125}I]IdC incorporation, the plasmids were divided into groups as shown in the histogram (Fig. 3). Incorporation significantly above background levels was observed for 83% of the isolates, ranging from 2 (group B) to 12 (group E) times background. For comparison, the level of [^{125}I]IdC incorporation with pTK2/SV010 was equal to the lowest levels obtained in group D; the level obtained with pTK206SVL/SV010 was equal to that obtained with the single isolate in group E. All members of group D showed levels of incorporation equal to or greater than that of pTK206PYalt/SV010.

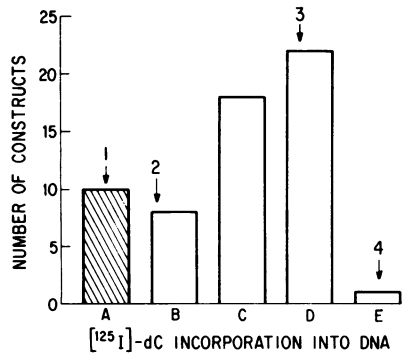

FIG. 3. Histogram of relative levels of *tk* gene expression in Cos-1 cells transfected with *tk* gene constructs containing fragments of monkey DNA. Minilysate DNA was prepared (9) from stationary cultures of *Escherichia coli* HB101 containing plasmids carrying inserts of monkey DNA in pTK206/SV010. Cos-1 cells were transfected, and 24 h later 0.5 μCi of [^{125}I]IdC was added to each 35-mm dish. After another 48 h, the amount of label incorporated into the DNA in each sample was determined. Cultures transfected with pTK2/SV010, pTK206SVL/ SV010, pTK206/SV010, or pTK206PYalt/SV010 served as controls. The plasmids containing inserts of monkey DNA were grouped together based on the amount of [^{125}I]IdC incorporated. Group A contains those plasmids with inserts that did not increase the amount of label incorporated over that for pTK206. Arrows indicate the relative levels of [^{125}I]IdC incorporation for: 1, pTK206/SV010; 2, pTK206PYalt/ SV010; 3, pTK2/SV010; and 4, pTK206SVL/ SV010.

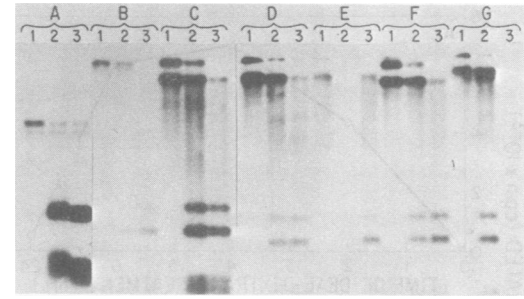

FIG. 4. Replication in Cos-1 cells of *tk* gene constructs carrying inserts of monkey DNA. Cultures were transfected with different constructs and harvested (24, 39) 1, 2, or 3 days later (indicated at top). Each culture also received ^{3}H-labeled M13 bacteriophage RFII DNA (20,000 to 30,000 cpm, about 20,000 cpm/μg; a generous gift from A. Wu and R. Kahn, Yale University). The ^{3}H-labeled phage DNA was included to allow correction for loss of material during the harvesting procedure. After progeny DNA were isolated, a sample was precipitated with TCA to calculate the percent recovery of [^{3}H]DNA. Samples with identical ^{3}H counts were digested with *Mbo*I restriction endonuclease, mixed with sample buffer, and loaded onto a 1% agarose gel in TBE buffer (89 mM Tris-chloride [pH 8.3], 89 mM boric acid, 2.5 mM EDTA). The DNA was transferred to nitrocellulose (45, 47) and hybridized with radiolabeled (42) pSV010 probe. (A) pSV010; (B) pTK206SA1/SV010; (C) pTK206SA2/SV010; (D) pTK206SA4/SV010; (E) pTK206SA9/SV010; (F) pTK206SA13/SV010; (G) pTK206SA16/SV010. The bands in the lower half of the figure represent progeny DNA.

Several plasmids that restored *tk* expression to a high level and a few that did not were chosen for further study. The replication of these representative plasmids was examined after Cos-1 cells were transfected with plasmid DNA. At various times, low-molecular-weight DNA was isolated, digested with *Mbo*I restriction endonuclease to allow the separation of progeny DNA from DNA in the inoculum, subjected to electrophoresis, transferred to nitrocel-

lulose, and hybridized with radiolabeled pTK206/SV010 DNA. The autoradiogram (Fig. 4) indicated that these plasmids replicated in Cos-1 cells but to a much lower extent than did pSV010 (Fig. 4A). In general, plasmids up to 6 kb replicated as well as did pSV010 (16), but larger plasmids replicated to a lower copy number. Plasmid pTK206SA2/SV010 (Fig. 4C) had an insert of 3.5 kb and replicated considerably better than did pTK206SA1/SV010 (Fig. 4B), which had an insert of 10.0 kb. Plasmids pTK206SA4/SV010 (Fig. 4D) and pTK206SA13/SV010 (Fig. 4F) replicated to a level intermediate between those of pTK206SA1/SV010 and pTK206SA2/SV010; they contained inserts of 6.0 and 4.5 kb, respectively. It is not known what level of replication is required for efficient expression of the *tk* gene.

tk gene expression in transfected Cos-1 cells was measured by the [^{125}I]IdC incorporation assay (Table 2). In addition, poly(A)$^+$ mRNA was isolated form Cos-1 cells transfected with several members of the library. After being electrophoresed in denaturing formaldehyde-agarose gels, the RNA was transferred to nitrocellulose and hybridized with radiolabeled pTK206/SV010 DNA. The autoradiographic data were analyzed by densitometry; representative tracings are shown in Fig. 5.

TABLE 2. *tk* gene expression in cells transfected with *tk* gene constructs containing inserts of monkey DNA downstream from *tk* coding sequences[a]

Construct	Approx insert size (kb)[b]	Relative [^{125}I]IdC incorporation (%)	Transformation frequency	
			Colonies/pmol	Relative frequency
pTK2/SV010		100	5,600	1.0
pTK206/SV010		2	20	<0.01
pTK206SA17/SV010	5.5	107	30,000	5
pTK206SA13/SV010	4.5	84	30,000	5
pTK206SA16/SV010	10.0	79	240,000	43
pTK206SA10/SV010	7.0	74	230,000	41
pTK206SA8/SV010	5.5	56	40,000	3
pTK206SA4/SV010	6.0	53	28,000	5
pTK206SA7/SV010	13.0	49	13,000	2
pTK206SA14/SV010	18.0	47	140,000	25
pTK206SA11/SV010	7.0	44	240,000	43
pTK206SA12/SV010	17.0	38	19,000	3
pTK206SA6/SV010	15.0	36	16,000	3
pTK206SA5/SV010	6.0	31	27,000	5
pTK206SA3/SV010	5.5	23	11,000	2
pTK206SA9/SV010	7.0	15	3,500	1
pTK206SA1/SV010	10.0	8	1,300	0.3
pTK206SA2/SV010	3.5	8	800	0.2
pTK206SA15/SV010	8.0	5	1,200	0.3

[a] See Table 1, footnotes *a*, *b*, and *c* for explanations.
[b] The size of each inserted fragment was estimated from electropherograms of *Bam*HI-digested plasmid DNA in 1% agarose gels, using *Ava*I fragments of bacteriophage lambda DNA as size markers.

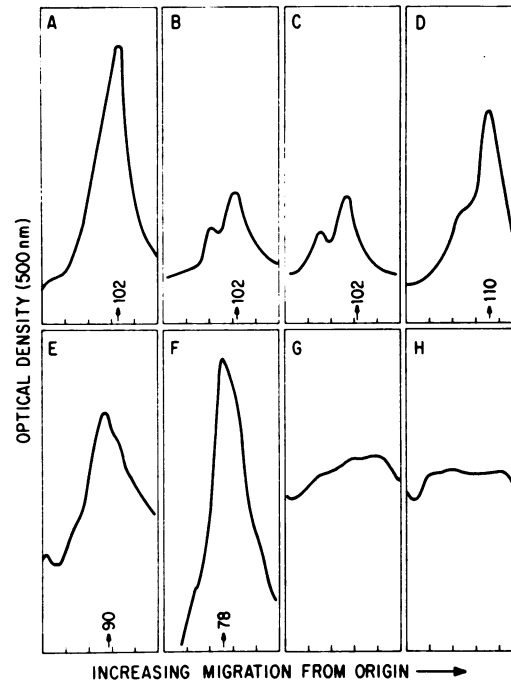

FIG. 5. Northern hybridization analysis of *tk*-specific mRNA in cells transfected by *tk* constructs containing inserts of monkey DNA. The distance from the origin to the position of each peak (in millimeters) is indicated where appropriate. All densitometer tracings were done at the same speed and optical density setting. Only the relevant portion of each tracing is shown. (A) pTK206SVL/SV010; (B) pTK206SA8/SV010; (C) pTK206SA10/ SV010; (D) pTK206SA6/SV010; (E) pTK206SA14/SV010; (F) pTK206SA17/SV010; (G) pTK206SA1/SV010; (H) pTK206SA15/SV010.

No discrete *tk*-specific mRNA was detected in Cos-1 cells transfected with pTK206/SV010 (Table 1) (16), but a discrete 1.5-kb poly(A)$^+$ *tk* mRNA species was detected in Cos-1 cells transfected with pTK206SVL/SV010 (Fig. 5A). The wild-type *tk* mRNA had a size of 1.5 kb, including the poly(A) tail (31, 49). The insertion into pTK206/SV010 of fragments of monkey DNA that contain processing and polyadenylation signals should allow the biosynthesis of *tk*-specific poly(A)$^+$ mRNA of discrete size, whereas genomic fragments lacking a polyadenylation signal should not. pTK206SA6/SV010, pTK206SA8/SV010, pTK206SA10/SV010, pTK206SA14/SV010, and pTK206SA17/SV010 produced major discrete poly(A)$^+$ *tk* mRNA species of approximately 1.4, 1.5, 1.5, 1.9, and 2.3 kb (± 0.25 kb), respectively, including the poly(A) tail (Fig. 5); in three cases (Fig. 5B, C, and D), there appeared to be two discrete

poly(A)$^+$ *tk* mRNA species, perhaps reflecting the presence of multiple processing and polyadenylation signals within the cellular DNA fragment and the failure of the first signal encountered to function with 100% efficiency. All five displayed relatively high levels of [^{125}I]IdC incorporation into DNA in transfected Cos-1 cells (Table 2). There was no correlation between the size of the transcription unit and the level of *tk* gene expression. We concluded that the inserted fragments of monkey DNA in these constructs contained sequences that restored processing and polyadenylation. In contrast, no discrete poly(A)$^+$ *tk*-specific mRNA was detected in Cos-1 cells transfected by pTK206SA1/SV010 or pTK206SA15/SV010 (Fig. 5), and the level of [^{125}I]IdC incorporation was very low (Table 2). We conclude that these inserts either lacked a processing and polyadenylation signal or contained such a weak signal that discrete mRNA species could not be detected by Northern hybridization analysis. In all cases examined, the restoration of *tk* gene expression, as measured by the IdC incorporation assay, was always associated with the production of poly(A)$^+$ *tk* mRNA.

Transformation of LTK$^-$ cells to TK$^+$. Each of the *tk* gene constructs (Fig. 1) was used to transform mouse LTK$^-$ cells by calcium phosphate coprecipitation of construct DNA, followed by HAT selection. Since there was a linear relationship in all cases between the amount of plasmid DNA added and the number of TK$^+$ colonies obtained per dish, it was necessary to use different amounts of plasmid DNA when transforming with different constructs. The total amount of DNA (plasmid plus carrier) added to each 100-mm dish was always 10 μg. The frequencies of transformation are listed in Table 1. The constructs were grouped by the source of the processing and polyadenylation signal (either polyoma virus or SV40). Within the two groups, there was a direct correlation between the level of transient expression in Cos-1 cells ([^{125}I]IdC incorporation into DNA or relative level of *tk* mRNA) and the transformation frequency in LTK$^-$ cells.

LTK$^-$ cells were also transformed with representative gene constructs containing fragments of monkey DNA. The transformation frequencies are listed in Table 2. Transformation frequencies 40 to 12,000 times those obtained with pTK206/SV010 were observed. Constructs that showed background levels of [^{125}I]IdC incorporation into DNA and no discrete poly(A)$^+$ *tk* mRNA were nevertheless able to transform LTK$^-$ cells, but with a frequency only 20 to 30% of that obtained with the wild-type *tk* gene (pTK2/SV010). Most of the other constructs transformed at frequencies between one and five times that observed with pTK2/SV010. Four of the constructs (pTK206SA10, -11-, -14-, and -16/SV010) transformed cells at 25 to 43 times the frequency of pTK2/SV010. There was no correlation between the size of the *tk* mRNA produced in Cos-1 cells (Fig. 5) and the frequency of transformation of LTK$^-$ cells. The results obtained in this experiment are consistent with the existence of processing and polyadenylation signals within the monkey DNA inserts in those constructs that transformed LTK$^-$ cells to TK$^+$ at a frequency equal to or greater than that of pTK2/SV010.

DISCUSSION

In this communication, we described an approach to the isolation of cellular regulatory sequences that are required at the 3′ end of genes for gene expression to occur. We began with plasmid pTK206/SV010, a *tk* gene construct containing the complete *tk* coding sequences and the *tk* promoter but lacking any processing and polyadenylation signal. We previously reported that the insertion into pTK206/SV010 of DNA fragments from SV40 and polyoma virus that carry known processing and polyadenylation signals restored its ability to produce TK and poly(A)$^+$ *tk* mRNA (16). In addition, poly(A)$^+$ *tk* mRNA synthesis was restored to a low level after the insertion of an SV40 DNA fragment containing the hexanucleotide 5′-AAUAAA-3′, which is not normally involved in processing and polyadenylation in SV40.

Processing and polyadenylation signals from fragments of African green monkey DNA. In the experiments reported here, fragments of African green monkey DNA were inserted into pTK206/SV010 downstream from *tk* coding sequences. *tk* gene expression was assayed both in Cos-1 cells (transient expression) and by transformation of LTK$^-$ cells to TK$^+$. In 83% of the clones with inserts of monkey DNA, *tk* gene expression was restored (Fig. 3), suggesting that most fragments of cellular DNA contain information that can direct the formation of a mature 3′ terminus for *tk* mRNA. The insert sizes in these clones ranged from 3.5 to 18.0 kb. The frequency with which *tk* gene expression was restored is consistent with a hexanucleotide (AAUAAA or a similar sequence) being a sufficient signal for the 3′-end maturation of an mRNA.

The level of *tk* gene expression obtained with these constructs showed a wide range. Measurements of the incorporation of [^{125}I]IdC into DNA in Cos-1 cells ranged from background levels (equal to incorporation in cells transfected with pTK206/SV010) to a level equal to that obtained with pTK206SVL/SV010 (Fig. 1), which showed the highest level of *tk* gene

expression in Cos-1 cells of any construct tested. More than half of the *tk* constructs had greater activity than the wild-type *tk* gene (pTK2/SV010; Fig. 1). There are several possible explanations for the different levels of tk gene expression observed.

(i) Cos-1 cells transfected with different *tk* gene constructs contained different numbers of *tk* genes (Fig. 4). All constructs were present at a lower copy number than pSV010, the wild-type *tk* gene pTK2/SV010, or *tk* gene constructs containing processing and polyadenylation signals from SV40 and polyoma virus (16). There was no correlation between the amount of replication (Fig. 4) and the amount of TK, as measured with the [^{125}I]IdC incorporation assay (Table 2); plasmid copy number was inversely related to plasmid size. Thus, the specific activities of the *tk* gene constructs containing fragments of monkey DNA were considerably higher than their measured relative activities (Table 2) when compared with the wild-type gene, pTK2/SV010.

(ii) Different cellular inserts may contain processing and polyadenylation signals that function with different efficiencies.

(iii) The *tk* mRNA species produced from each *tk* construct would be expected to have different stabilities since they differ from one another in the sequence of the 3′ untranslated region, the portion of the mRNA which is most labile (5, 6, 10, 28). The length of the 3′ untranslated region varied from 0.1 to 1.1 kb in the mRNAs examined by Northern hybridization analysis (Fig. 5); however, there was no correlation between the length of the 3′ untranslated region and the level of *tk* gene expression.

(iv) Some of the cellular fragments could contain intervening sequences; the efficiency of excision and transport could then affect the level of *tk* mRNA.

In all cases examined, a relatively high level of [^{125}I]IdC incorporation into DNA was correlated with the biosynthesis of discrete poly(A)$^+$ *tk* mRNA, whereas poly(A)$^+$ *tk* mRNA was not detected in cells transfected with the two constructs that showed near-background levels of [^{125}I]IdC incorporation (Fig. 5, Table 2). In some cases, there appeared to be more than one poly(A)$^+$ *tk* mRNA species.

DNA sequence analysis of the monkey DNA fragments will be required to determine the sequence of the processing and polyadenylation signals they contain; preliminary sequence analysis of the first 200 bp of the monkey DNA fragment in pTK206SA8/SV010 indicated that the hexanucleotide AAUAAA does not occur, but the sequences AUUAAA and AAUUAAA each appear once. These sequences are found in place of AAUAAA in some genes (22, 26). An unspliced *tk* mRNA of 1.5 kb (Fig. 5) would have its 3′ end within the first 200 bp of the monkey sequences. Therefore, it is likely that one or both of these sequences signals processing and polyadenylation for the *tk* gene in pTK206SA8/SV010.

Some histone genes produce nonpolyadenylated mRNAs (1, 23, 25). The production of the discrete 3′ terminus of histone mRNA is thought to be dependent on the presence of a specific sequence distal to sequences appearing in the mRNA and capable of forming a specific, ordered secondary structure (8). From our studies, it appears that none of the constructs examined contained signals for the production of nonpolyadenylated *tk* mRNAs. Alternatively, nonpolyadenylated *tk* mRNAs could be sufficiently unstable that the level of [^{125}I]IdC incorporation into DNA would be extremely low and *tk* mRNA might not be detected. Since most mRNAs are poly(A)$^+$, signals for the production of the 3′ termini of nonpolyadenylated mRNAs may be extremely rare, and the isolation of these could require the preparation of a much larger functional 3′-end signal library. It is also possible that genes that normally give rise to poly(A)$^+$ mRNA cannot function with signals for the production of nonpolyadenylated 3′ termini.

Transformation of LTK$^-$ cells with *tk* gene constructs. Gene expression was also assayed by transformation of LTK$^-$ cells to TK$^+$. All of the constructs examined, including those which did not produce detectable poly(A)$^+$ *tk* mRNA, transformed LTK$^-$ cells at levels substantially above background (Table 2). Whereas pTK206/SV010 had a transformation efficiency below 20 colonies per pmol of *tk* gene, constructs with cellular inserts transformed at efficiencies of 800 to 240,000 colonies per pmol of *tk* gene. Examination of the data (Table 2) allowed the constructs to be placed into three groups. Group 1 (three constructs) transformed LTK$^-$ cells at frequencies below that of pTK2/SV010, the wild-type gene. Cos-1 cells transfected with group 1 constructs incorporated [^{125}I]IdC into DNA at a level only slightly, but reproducibly, above background and produced no poly(A)$^+$ *tk* mRNA. These three either contained a weak signal or lacked a specific signal entirely. Most of the other constructs (group 2, 10 constructs) transformed LTK$^-$ cells at frequencies between one and five times that observed with pTK2/SV010. All group 2 constructs examined produced poly(A)$^+$ *tk* mRNA; therefore, the monkey DNA fragments in these constructs contained processing and polyadenylation signals. Four of the constructs (group 3; pTK206SA10/SV010, pTK206SA11/SV010, pTK206SA14/SV010, and pTK206SA16/SV010) transformed LTK$^-$ cells at frequencies 25 to 43

times that of pTK2/SV010. No constructs were isolated that transformed LTK$^-$ cells at frequencies intermediate between those of group 2 and 3. It is possible that the cellular DNA fragments in group 3 constructs contained enhancer sequences in addition to a processing and polyadenylation signal. The transformation frequencies of group 3 constructs also exceeded those for all constructs containing processing and polyadenylation signals from polyoma virus and SV40 (Table 1), none of which contained enhancer sequences. Additional experiments will be required to determine whether these group 3 constructs contain cellular enhancer sequences.

The process of DNA-mediated gene transfer is thought to involve the formation of a very large structure (38) that contains carrier DNA sequences, selectable gene sequences, and any cotransfected DNA sequences in a large covalent array. Both simian (Table 2) and murine (unpublished data) DNA sequences were capable of restoring processing and polyadenylation to the *tk* transcription unit. The frequency of transformation by pTK206/SV010 was extremely low (Table 2), suggesting that the association of the *tk* sequences in pTK206/SV010 with carrier DNA sequences did not occur soon enough to allow the *tk* gene to function and the cells to survive HAT selection. In these experiments, selection in HAT medium was initiated 48 h after DNA was first added to LTK$^-$ cells.

Correlation between transient expression and biochemical transformation. The HSV *tk* gene can be expressed either episomally or from an integrated copy. The former is defined as transient expression (3, 17, 32–34) and the latter as stable transformation. In this study we reported a direct correlation between the level of *tk* gene expression in Cos-1 cells (transient expression) and the frequency of biochemical transformation of LTK$^-$ cells for two sets of derivatives of pTK206/SV010, one containing the SV40 and the other the polyoma virus processing and polyadenylation signals (Table 1). We conclude that the fraction of cells that survive HAT selection is determined by the level of TK in recipient LTK$^-$ cells during a critical interval. Presumably this critical interval, during which the *tk* gene is transiently expressed, requires factors present only in the nucleus. Therefore, the critical interval begins after the transforming DNA reaches the nucleus and continues at least until the selective agent is added. These conclusions rely on the assumption that the level of TK in Cos-1 cells can be related to the complex phenomenon of biochemical transformation of LTK$^-$ cells. These plasmids attained a similar copy number in Cos-1 cells (16), eliminating replication as a factor in the quantitation of cytoplasmic RNA levels. The species difference between mouse L cells and monkey CV-1 cells does not affect the correlation, since both polyoma virus and SV40 insertion constructs showed the correlation between level of expression in Cos-1 cells and transformation frequency of LTK$^-$ cells. The normal host species for polyoma virus and SV40 are murine and simian cells, respectively. It is likely that any DNA element that acts to increase the level of selectable gene mRNA will increase the transformation frequency.

By using the approach followed in the experiments presented in this report, it should be possible to isolate processing and polyadenylation signals from virtually any source. The assays used are sensitive to both very low levels and changes in the level of *tk* gene expression. Therefore, this approach should help in determining what types of DNA sequences or secondary structures are capable of increasing the efficiency of processing and polyadenylation or altering the stability of the mRNA. In addition, the same methods can be used with different *tk* gene constructs (e.g., lacking promoter activity) in attempts to isolate cellular DNA fragments containing promoters or enhancers.

ACKNOWLEDGMENTS

We thank Wilma Summers, Joanne Tornow, MaryEllen Polvino-Bodnar, and Roger Kahn for helpful discussions, Katherine Miceli for excellent technical assistance, and Kathryn Belmont for preparation of the manuscript.

This work was supported by Public Health Service grant CA-16038 from the National Institutes of Health and by BRSG grant RR 0538, awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health.

LITERATURE CITED

1. **Adesnik, M., and J. E. Darnell.** 1972. Biogenesis and characterization of histone mRNA in HeLa cells. J. Mol. Biol. **67**:397–406.
2. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidilic acid-cellulose. Proc. Natl. Acad. Sci. U.S.A. **69**:1408–1412.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell **27**:299–308.
4. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
5. **Bergmann, I. E., and G. Brawerman.** 1977. Control of breakdown of the polyadenylate sequence in mammalian polyribosomes: role of poly(adenylic acid) protein interactions. Biochemistry **16**:259–264.
6. **Bergmann, I. E., and G. Brawerman.** 1980. Loss of the polyadenylate segment from mammalian messenger RNA: selective cleavage of this sequence from polyribosomes. J. Mol. Biol. **139**:439–454.
7. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
8. **Birchmeier, C., R. Grosschedl, and M. L. Birnstiel.** 1982. Generation of authentic 3′ termini of an H2A mRNA *in vivo* is dependent on a short inverted DNA repeat and on spacer sequences. Cell **28**:739–745.
9. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline

extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1523.
10. **Brawerman, G.** 1981. The role of the poly(A) sequence in mammalian messenger RNA. Crit. Rev. Biochem. 10:1–38.
11. **Chen, M. S., and W. H. Prusoff.** 1979. Phosphorylation of 5-iodo-5'-amino-2',5'-dideoxyuridine by herpes simplex virus type 1-encoded thymidine kinase. J. Biol. Chem. 254:10449–10452.
12. **Chen, M. S., J. Walker, and W. H. Prusoff.** 1979. Kinetic studies of herpes simplex virus type 1-encoded thymidine and thymidylate kinase, a multifunctional enzyme. J. Biol. Chem. 254:10747–10753.
13. **Cheng, Y.-C., and M. Ostrander.** 1976. Deoxythymidine kinase induced in HeLa Tk⁻ cells by herpes virus type 1 and type 2. J. Biol. Chem. 251:2605–2610.
14. **Clewell, D. B., and D. R. Helinski.** 1970. A membrane filter technique for the detection of complementary DNA. Biochemistry 9:4428–4440.
15. **Colbere-Garapin, F., S. Chousterman, F. Horodniceanu, P. Kourilsky, and A.-C. Garapin.** 1979. Cloning of the active thymidine kinase gene of herpes simplex virus type 1 in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. U.S.A. 76:3755–3759.
16. **Cole, C. N., and G. Santangelo.** 1983. Analysis in Cos-1 cells of processing and polyadenylation signals by using derivatives of the herpes simplex virus type 1 thymidine kinase gene. Mol. Cell. Biol. 3:267–279.
17. **Dierks, P., A. van Ooyen, N. Mantei, and C. Weissman.** 1981. DNA sequences preceding the rabbit β-globin gene are required for formation in mouse L cells of β-globin RNA with the correct 5' terminus. Proc. Natl. Acad. Sci. U.S.A. 78:1411–1415.
18. **Favaloro, J., R. Treisman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNAs: analysis by two-dimensional nuclease S1 mapping. Methods Enzymol. 65:718–749.
19. **Fitzgerald, M., and T. Shenk.** 1981. The sequence 5'-AAUAAA-3' forms part of the recognition site of polyadenylation of late SV40 mRNAs. Cell 24:251–260.
20. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of SV40 early mutants. Cell 23:175–182.
21. **Gonda, T. J., D. K. Sheiness, and J. M. Bishop.** 1982. Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. Mol. Cell. Biol. 2:617–624.
22. **Hagenbüchle, C., R. Bovey, and R. A. Young.** 1980. Tissue-specific expression of mouse α-amylase genes: nucleotide sequence of isoenzyme mRNAs from pancreas and salivary glands. Cell 21:179–187.
23. **Hentschel, C. C., and M. L. Birnstiel.** 1981. The organization and expression of histone gene families. Cell 25:301–313.
24. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cultures. J. Mol. Biol. 26:365–369.
25. **Jacob, E., G. Malacinski, and M. L. Birnstiel.** 1976. Reiteration frequency of the histone genes in the genome of the amphibian, *Xenopus laevis*. Eur. J. Biochem. 69:45–54.
26. **Jung, A., A. E. Sippel, M. Grez, and G. Schutz.** 1980. Exons encode functional and structural units of chicken lysozyme. Proc. Natl. Acad. Sci. U.S.A. 77:5759–5763.
27. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions: a critical reexamination. Biochemistry 16:4743–4751.
28. **Littauer, U. Z., and H. Soreq.** 1982. The regulatory fucntion of poly(A) and adjacent 3' sequences in translated RNA. Prog. Nucleic Acid Res. Mol. Biol. 27:53–83.
29. **Mandel, M., and A. Higa.** 1970. Calcium-dependent bacteriophage-DNA interaction. J. Mol. Biol. 53:159–162.
30. **Maxam, A., and W. S. Gilbert.** 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.
31. **McKnight, S. L.** 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucleic Acids Res. 4:241–246.
32. **Mellon, P., V. Parker, Y. Gluzman, and T. Maniatis.** 1981. Identification of DNA sequences required for transcription of the human α-globin gene in a new SV40 host-vector system. Cell 27:279–288.
33. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science 209:1422–1427.
34. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit βglobin in cultured monkey kidney cells following infection with a SV40 β-globin recombinant genome. Nature (London) 277:108–114.
35. **Myers, R. M., D. C. Rio, A. K. Robbins, and R. Tijian.** 1981. SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell 25:373–384.
36. **Ostrander, M., S. Vogel, and S. Silverstein.** 1982. Phenotypic switching in cells transformed with the herpes simplex virus thymidine kinase gene. Mol. Cell. Biol. 2:708–714.
37. **Perricaudet, M., J.-M. le Moullec, P. Tiollais, and U. Pettersson.** 1980. Structure of two adenovirus type 12 transforming polypeptides and their evolutionary implications. Nature (London) 288:174–176.
38. **Perucho, M., D. Hanahan, and M. Wigler.** 1980. Genetic and physical linkage of exogenous sequences in transformed cells. Cell 22:309–317.
39. **Polvino-Bodnar, M., and C. N. Cole.** 1982. Construction and characterization of viable deletion mutants of simian virus 40 lacking sequences near the 3' end of the early region. J. Virol. 43:489–502.
40. **Proudfoot, N., and G. G. Brownlee.** 1976. 3' non-coding region sequences in eucaryotic messenger RNA. Nature (London) 263:211–214.
41. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, and S. M. Weissman.** 1978. The genome of SV40. Science 200:494–502.
42. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113:237–251.
43. **Soeda, E., J. R. Arrand, N. Smolar, J. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) 283:445–453.
44. **Sompayrac, L. M., and K. M. Danna.** 1981. Efficient infection of monkey kidney cells with DNA of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. 78:7575–7578.
45. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
46. **Summers, W. C., and W. P. Summers.** 1977. [^{125}I]deoxycytidine used in a rapid, sensitive, and specific assay for herpes simplex virus type 1 thymidine kinase. J. Virol. 24:314–318.
47. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. 77:5201–5205.
48. **Tooze, J., ed.** 1980. Molecular biology of tumor viruses, part II: DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
49. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. U.S.A. 78:1441–1445.
50. **White, R. T., P. Berg, and L. P. Villarreal.** 1982. Simian virus 40-rabbit β-globin recombinants lacking late mRNA splice sites express cytoplasmic RNAs with altered structures. J. Virol. 42:262–274.
51. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated gene transfer of the adenine phosphoribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. U.S.A. 76:1373–1376.
52. **Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel.** 1979. Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell 16:777–785.