MOLECULAR AND CELLULAR BIOLOGY, Dec. 1982, p. 1628–1632
0270-7306/82/121628-05$02.00/0

Vol. 2, No. 12

NOTES

Expression of Bacterial β-Galactosidase in Animal Cells

GYNHEUNG AN,† KATSUHIKO HIDAKA,‡ AND LOUIS SIMINOVITCH*

Department of Medical Genetics, University of Toronto, and The Hospital for Sick Children, Toronto, Ontario, Canada, M5G 1X8

Received 4 June 1982/Accepted 26 August 1982

A recombinant plasmid containing the gene for bacterial β-galactosidase, situated close to the simian virus 40 early promoter, has been constructed. Transfection of CHO, L, and COS-1 cells with this plasmid led to the expression and appearance of the enzyme. Using this system, we have developed a series of promoter cloning vehicles capable of accepting promoter signals for animal genes.

Although recent advances in recombinant DNA technology have made many eucaryotic genes available for studies in eucaryotic cells, progress in the examination of regulation in such systems has been limited (4, 8, 9). In principle, the methods used for procaryotic cells, in which a control sequence is fused to a foreign gene for which the gene product is readily measurable (1, 5), should be applicable to eucaryotic systems. In this paper we describe the construction of a recombinant plasmid containing the gene for bacterial β-galactosidase in which the gene is under the control of an animal promoter and is expressed at high levels after transfection of CHO, L, or COS-1 cells.

The procedures used to insert the *Escherichia coli* gene for β-galactosidase into an animal vector that contains the signals necessary for the expression of bacterial genes are outlined in Fig. 1. In pGA291, the simian virus 40 (SV40) early promoter is located about 1,300 base pairs (bp) upstream from the coding sequence for the β-galactosidase gene (12). The polyadenylation signal and intervening sequence of the SV40 early transcript are located about 4 kilobases downstream from the structural gene (16).

Since there is a long leader sequence between the SV40 early promoter and the translation initiation site for the *tufB-lacZ* region in pGA291, it was anticipated that translation of the gene might not be efficient in animal cells. The first AUG in eucaryote mRNAs usually serves as an initiation codon for translation, and AUG sequences found in the leader mRNA would probably decrease the translation efficiency (10, 15). Another plasmid, pGA293, was therefore constructed from GA291 by removing most of the leader region (see Fig. 1). The DNA sequences at the leader region of *tufB-lacZ* in pGA293 are shown in Fig. 1. The major cap site is located about 67 bp upstream from the translation start codon for the *tufB-lacZ* and there is no ATG sequence in between.

The levels of β-galactosidase were found to be low in L and COS-1 cells and almost immeasurable in CHO cells when the cells were transfected with CHO DNA as controls. These hosts were therefore used to examine whether the β-galactosidase gene from pGA291 and pGA293 was expressed after transfection. Cells were transfected with plasmid DNA by the calcium phosphate technique and assayed for β-galactosidase activity.

As may be seen in Table 1, β-galactosidase activity in CHO cells was about 50-fold higher when the cells were transfected with pGA293 as compared to controls transfected with the same amount of either CHO DNA, pSV2, or pGA164. The data in Table 1 also show that the activity obtained with pGA291 was about 20% of that found with pGA293. The presence of the long leader sequence in pGA291 therefore decreased the expression of the enzyme but did not abolish it. Since β-galactosidase activity in cells transfected with pGA164, which contains the *lac* gene but lacks the SV40 promoter and the processing signals, was similar to that in controls, our results demonstrate that expression of the bacterial β-galactosidase is dependent on the SV40 early promoter.

The data in Table 1 show that the pGA293 plasmid β-galactosidase gene was also expressed in L cells and COS-1 cells. The background levels in these systems are higher than in CHO cells, but there are obviously large increases in activity after transfection. The very

† Present address: Department of Microbiology and Immunology, University of Washington, Seattle, WA 98195.

‡ Present address: Department of Surgery, Saga-Medical School, Nabeshima-Machi, Saga-City 840-01, Japan.

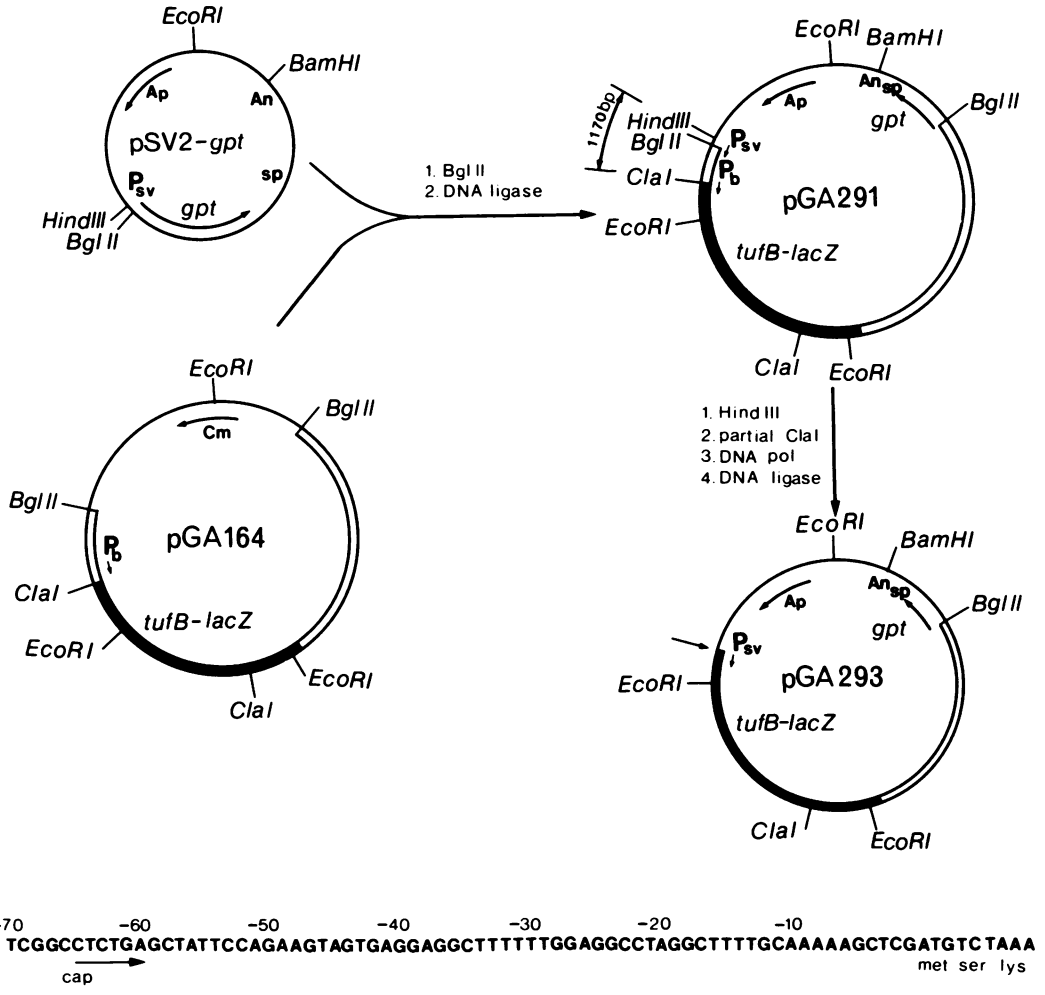

FIG. 1. Structures of plasmids PGA291 and pGA293. A DNA fragment containing the bacterial gene for β-galactosidase was obtained from the plasmid pGA164 (12). The plasmid contains the *tufB-lacZ* fusion in which the distal part of *tufB*, a gene for elongation factor EF-Tu, was replaced with the *lac* structural gene (12). In this fusion, the reading frame of *lacZ*, the gene for β-galactosidase, is the same as that of *tufB*. Although the fusion protein produced from the gene is larger than the normal *lacZ* gene product, it functions as a β-galactosidase in *E. coli*. pGA164 was used as the source of the β-galactosidase gene since it contains a *Cla*I site at the initiation codon of the *tufB* (see later). To place the bacterial gene under an animal promoter, the *Bgl*II fragment (double horizontal lines) of pGA164 containing the *tufB-lacZ* fusion (filled box) was inserted at the *Bgl*II site in pSV2-gpt (16). Plasmid pGA293 was constructed as follows. Plasmid PGA291 was digested completely with *Hin*dIII and partially with *Cla*I. The restricted DNA was treated with DNA polymerase large fragment in the presence of the four nucleotides to generate flush ends. It was then sealed with T4 DNA ligase and used to transform *E. coli* MC1000 (5), which carries a *lac* deletion. PGA291 contains the bacterial promoter for the fusion on the 1,170-bp *Hin*dIII-*Cla*I fragment (2). Therefore, MC1000 harboring pGA291 formed blue colonies in X-gal (5-bromo-4-chloro-3-indoryl-β-D-galactoside) indicator plates. However, MC1000 harboring a plasmid containing a deletion of the 1,170-bp *Hin*dIII-*Cla*I fragment would form pale blue colonies. The latter colonies contained the plasmid pGA293. The DNA sequence of the transcript strand at the leader region of *tufB-lacZ* in pGA293 is shown in the lower part of the figure. The sequence was deduced from the SV40 sequence (20) and the *tufB* sequence (2). The major cap site for the SV40 early transcript is located near −67 (11). The numbers indicate the distance from the translation initiation codon. Symbols: Ap, ampicillin resistance gene; Cm, chloramphenicol acetyltransferase; *gpt*, a bacterial gene for xanthine-guanine phosphoribosyl-transferase; P_B, the region containing the bacterial promoter for the *tufB-lacZ* gene; Psv, the region containing the animal promoter for the *tufB-lacZ* gene and the origin of replication for SV40; Sp, intervening sequence of the SV40 early transcript; An, polyadenylation signal for SV40 early transcripts.

TABLE 1. β-Galactosidase expression in transfected cells[a]

DNA donor	Enzyme activity (units) in recipient cells:		
	CHO TK⁻	COS-1	Lmtk⁻
pGA293	308.4 (100)[b]	1,250.0	206.0
CHO	6.5 (2.1)	29.7	15.7
pSV2	5.5 (1.8)		
pGA164	9.8 (3.2)		
pGA291	60.1 (19.5)		
pGA296	124.1 (40.2)		
pGA300	50.5 (16.4)		
pGA307	10.6 (3.4)		

[a] CHO thymidine kinase-deficient (TK⁻) cells, mouse Lmtk⁻ cells, and simian COS-1 cells were maintained in minimal essential medium α containing penicillin, streptomycin, and 10% fetal bovine serum (19). Approximately 10^6 cells grown in a 75-cm^2 flask were transfected with 15 μg of DNA plus 10 μg of CHO DNA as a carrier, according to Graham and Van der Eb (7), as modified by Wigler et al. (21), with minor modifications as described below. Briefly, 1 ml of 250 mM $CaCl_2$ containing DNA was added to 1 ml of twice-concentrated HEPES-buffered saline (250 mM NaCl, 1.5 mM Na_2HPO_4, 40 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid], pH 7.1 ± 0.05) with vigorous mixing using bubbling air. After 20 min at room temperature, the precipitate was added to the cells. A 100-μl volume of amphotericin B, dissolved freshly in dimethyl sulfoxide at 1 mg/ml, was added to the flask. The addition of amphotericin B consistently enhanced transformation frequency (Hidaka, An, and Siminovitch, manuscript in preparation). After incubation under CO_2 at 37°C for 20 h, the medium was removed, and fresh medium was added. After 2 days at 37°C, the transfected cells were trypinized and assayed for β-galactosidase. The assay of β-galactosidase activity in animal cells was essentially similar to the bacterial assay as described by Miller (14). Cells were harvested, washed with phosphate-buffered saline, and suspended in 1 ml of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4 [pH 7.0], 10 mM KCl, 1 mM $MgSO_4$, and 50 mM β-metcaptoethanol). One drop of toluene was added, and the mixture was incubated for 1 h at 37°C. A 0.2-ml volume of *o*-nitrophenyl-β-D-galactoside was added, the mixture was incubated at 28°C for 1 to 2 h, and the reaction was stopped by adding 0.5 ml of 1 M sodium carbonate. The cell debris was pelleted by centrifugation in an Eppendorf centrifuge for 10 min. The supernatant was used to measure optical density (OD) at 420 and 550 nm. One unit of β-galactosidase is defined as $10^9 \times (OD_{420} - 1.1 \times OD_{550})/(\text{cell number} \times \text{reaction time [h]})$.

[b] The figures in parentheses refer to the β-galactosidase activity expressed as a percentage of the value of pGA293, taken as 100%. The L-cell value (top line) was obtained by using 20 μg of pGA293 and 10 μg of CHO DNA.

large value observed in COS-1 cells is probably due to the fact that the plasmid replicates in these hosts because of the presence of endogenous T antigen (6).

TABLE 2. Kinetics of expression of β-galactosidase activity in CHO cells[a]

Expression time (h)	No. of cells ($\times 10^5$)	Sp act (U)	Total activity (U × 10^5)
0	5.5	10.3	0
12	5.2	136.2	654.7
24	6.0	146.7	819.0
48	10.5	88.2	819.0
72	34.5	66.5	1,942.4
96	70.0	22.6	868.0

[a] Experiments were carried out as described in Table 1, footnote *a*. Total activity was calculated as the number of cells times the specific activity after subtraction of the CHO background.

It was next of interest to determine the kinetics of enzyme formation and the dose response. In the first experiment, CHO cells were transfected with 10 μg of pGA293, and 16 h later the medium was replaced with fresh medium. The transfected cells were then harvested at various times and assayed for β-galactosidase activity. The amount of β-galactosidase per cell reached a maximum 24 h after removal of the DNA used for transfection, and thereafter the activity per cell decreased (Table 2). Total activity continued to increase in the culture for 3 days, however. Data on the dose response for phenotypic expression of plasmid activity are given in Table 3. β-Galactosidase activity increased as a function of DNA concentration to a maximum when 10 μg of plasmid DNA per flask was used.

The above results demonstrate that the plasmids we have constructed (pGA291 and pGA293) are vehicles which could be used to examine phenotypic expression of a bacterial gene in eucaryotic cells. We then examined the role of polyadenylation sites and splice sequences which have been considered necessary for the expression of mammalian genes (16). pGA293 was originally constructed in the pSV2-

TABLE 3. Dose response for expression of β-galactosidase activity in CHO cells[a]

pGA293 DNA (μg)	CHO DNA (μg)	No. of cells ($\times 10^5$)	Sp act (U)
0	20.0	13.5	3.2
0.1	20.0	15.0	3.7
1.0	19.0	15.0	9.2
5.0	15.0	18.0	45.9
10.0	10.0	14.4	237.1
20.0	10.0	9.5	142.6

[a] Experiments were carried out as described in Table 1, footnote *a*, except for the variations in amounts of plasmid DNA. Total activity was calculated as the number of cells times the specific activity after subtraction of the CHO background.

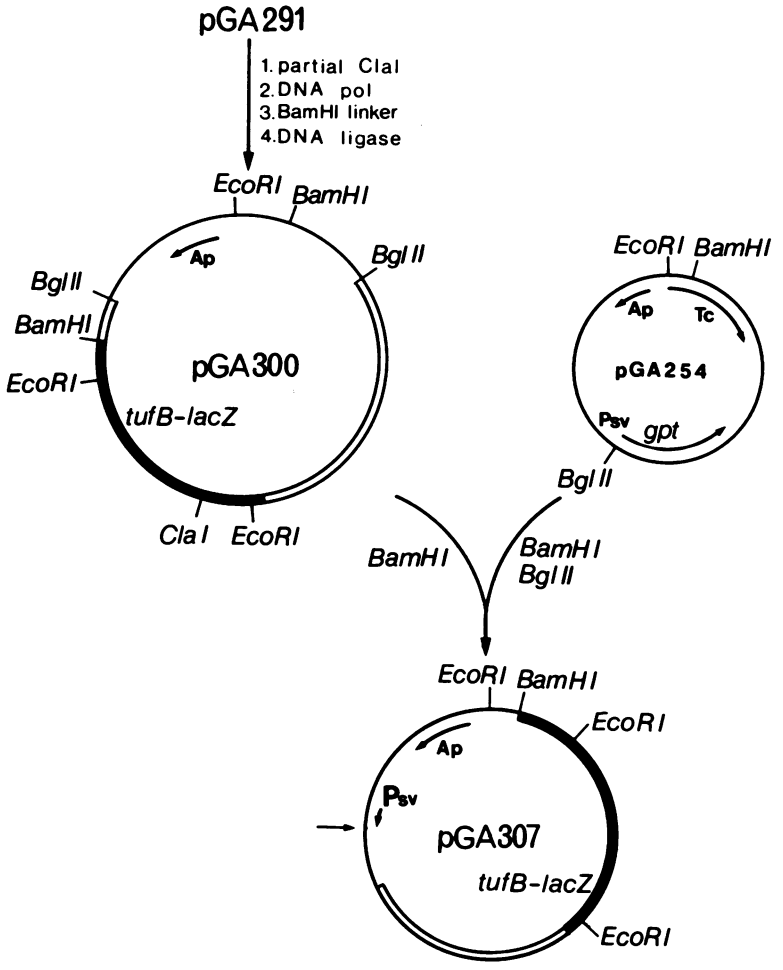

FIG. 2. Construction of plasmids pGA300 and pGA307. Even though a foreign fragment containing an animal promoter could be introduced at the *Cla*I site of the translation initiation site for the *lac* gene in pGA291, there is another *Cla*I site in the *lacZ* structural gene. It was therefore not convenient to use pGA291 for this purpose. Also the SV40 early promoter located upstream from the *Cla*I site would influence the expression of the bacterial gene, as shown by our earlier results. The *Cla*I site at the initiator region was therefore converted to a *Bam*HI site as follows. PGA291 was partially digested with *Cla*I in the presence of 150 μg of ethidium bromide per ml (17). The DNA was subjected to electrophoresis in an agarose gel, and the full-length fragment was isolated from the gel. Blunt ends were generated on the fragment by DNA polymerase, and an oligonucleotide-containing *Bam*HI site was ligated to the fragment. One of the resulting plasmids contained a *Bam*HI site in place of the *Cla*I at the initiator region. This plasmid was designated as pGA300. To make pGA307, plasmid pGA254 was constructed first by replacing the *Eco*RI-*Bam*HI fragment of pSV2 with a 1,459-bp *Eco*RI-*Sau*3A fragment containing the tetracycline resistance genes (Tc) from pBR322. pGA254 therefore contains both ampicillin and tetracycline resistance genes. It has lost the *Bam*HI site of pSV2 and has acquired a new *Bam*HI site in the tetracycline resistance gene (An, unpublished data). The *Bam*HI-*Bgl*II fragment containing the bacterial *gpt* gene of pSV2 from plasmid pGA254 was replaced with the large *Bam*HI fragment containing the *lac* gene pGA300. The new plasmid, pGA307, contains only one *Bam*HI site, located at the start region of the *lac* gene, since the *Bam*HI site at the 3′ end of the gene has been removed by ligation to the *Bgl*II site in pSV2.

gpt plasmid so that these signals would be present in the expression vector. The *Bam*HI-*Bgl*II fragment containing the polyadenylation site and splice signals was subsequently removed from pGA293, and the β-galactosidase activity produced in CHO cells transfected with the resulting deletion plasmid, pGA296, was assayed. Surprisingly, there was considerable expression of bacterial β-galactosidase under these conditions (Table 1). The amount of β-

galactosidase produced by pGA296 was about 40% of that from pGA293. These results indicate that the processing signals at the 3′ end of the gene may not be necessary for the expression of the bacterial gene in CHO cells. Schumperli et al. observed that removal of the splice sequences did not influence the expression of the bacterial gene for galactokinase in animal cells (18). Since several animal genes do not contain intervening sequences, it is perhaps not surprising to have obtained expression of the *lac* gene in the absence of intervening sequences. However, expression in the absence of the polyadenylation signal for the animal gene is more difficult to understand. It is possible that the DNA sequence in pGA296 downstream from *lacZ* contains a polyadenylation signal.

The construction of plasmids pGA300 and pGA307 to facilitate the insertion of animal promoters is shown in Fig. 2. pGA300 contains two *Bam*HI sites, one at the initiator region of the *lac* gene and the other downstream from the processing signals. Therefore, the *Bam*HI fragment from pGA300 containing the bacterial *lac* gene as well as the processing signals could be inserted downstream from any desired animal promoter. The bacterial *lac* gene would then be under the control of the animal promoter. The activity of this plasmid was about 15% of that of pGA293.

However, because there is no unique restriction endonuclease site in the *lac* leader region, it would be difficult to insert foreign fragments into pGA300 itself. The manipulations shown in Fig. 2 were carried out to prepare pGA307, a plasmid containing a unique *Bam*HI site at the initiator region of the *lac* gene. pGA307 contains a DNA sequence of SV40 origin which includes 72-bp repeats and which is known to improve transformation (3, 13). However, there might be a potential effect of having the SV40 promoter present during transcription in the opposite direction of that of the fusion transcription unit. As may be seen in Table 1, pGA307 gives a very low background activity when it is used for transfection and therefore it should be an excellent vehicle for insertion and assay of various promoter-like sequences.

We thank P. Berg for pSV2-*gpt* and Y. Gluzman for the COS-1 cell line.

G.A. holds a fellowship from the Medical Research Council of Canada. This work was funded by the Medical Research Council and National Cancer Institute of Canada.

LITERATURE CITED

1. **An, G., and J. D. Friesen.** 1979. Plasmid vehicles for direct cloning of *Escherichia coli* promoters. J. Bacteriol. **140**:400–407.
2. **An, G. and J. D. Friesen.** 1980. The nucleotide sequence of *tufB* and four nearby tRNA structural genes of *Escherichia coli*. Gene **12**:33–39.
3. **Banerji, J., S. Rusconi, and W. Schaffner.** 1981. Expression of a β-globin gene is enhanced by remote SV40 DNA sequence. Cell **27**:299–308.
4. **Benoist, C., and P. Chambon.** 1981. The SV40 early promoter region: sequence requirements *in vivo*. Nature (London) **290**:304–310.
5. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
6. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
7. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
8. **Grosveld, G. C., E. deBoer, C. K. Shewmaker, and R. A. Flavell.** 1982. DNA sequence necessary for transcription of the rabbit β-globin gene *in vivo*. Nature (London) **295**:120–126.
9. **Hamer, D. H., M. Kaehler, and P. Leder.** 1980. A mouse globin gene promoter is functional in SV40. Cell **21**:697–708.
10. **Kozak, M.** 1980. Evaluation of the "scanning model" for initiation of protein synthesis in eukaryotes. Cell **22**:7–8.
11. **Lebowitz, P., and P. K. Ghosh.** 1982. Initiation and regulation of simian virus 40 early transcription in vitro. J. Virol. **41**:449–461.
12. **Lee, J. S., G. An, J. D. Friesen, and N. P. Fiil.** 1981. Location of the *tufB* promoter of *E. coli*: cotranscription of *tufB* with four transfer RNA genes. Cell **25**:251–261.
13. **Levinson, B., G. Khoury, G. vande Woude, and P. Gruss.** 1982. Activation of SV40 genome by 72-base pair tandem repeats of Moloney sarcoma virus. Nature (London) **295**:568–572.
14. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Mulligan, R. C., and P. Berg.** 1981. Factors governing the expression of a bacterial gene in mammalian cells. Mol. Cell. Biol. **1**:449–459.
16. **Mulligan, R. C., and P. Berg.** 1981. Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyl-transferase. Proc. Natl. Acad. Sci. U.S.A. **78**:2072–2076.
17. **Parker, R. C., R. M. Watson, and J. Vinograd.** 1977. Mapping of circular DNAs by cleavage with restriction endonucleases and calibration by agarose gel electrophoresis. Proc. Natl. Acad. Sci. U.S.A. **74**:851–855.
18. **Schumperli, D., B. H. Howard, and M. Rosenberg.** 1982. Efficient expression of *Escherichia coli* galactokinase gene in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **79**:257–261.
19. **Stanners, C. P., G. L. Elicleri, and H. Green.** 1971. Two types of ribosomes in mouse-hamster hybrid cells. Nature (London) New Biol. **230**:52–54.
20. **van Heuverswyn, H., and W. Fiers.** 1979. Nucleotide sequence of the Hind-C fragment of simian virus 40 DNA. Eur. J. Biochem. **100**:51–60.
21. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated transfer of the adenine phosphoribosyl-transferase locus into mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **76**:1373–1376.