

- G. Urlaub, L. Chasin, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1373 (1979).
9. L. H. Graf, G. Urlaub, L. Chasin, *Somat. Cell Genet.* **5**, 1031 (1979).
 10. S. C. Lester, S. K. LeVan, C. Steglich, R. DeMars, *ibid.* **6**, 241 (1980).
 11. M. Perucho, D. Hanahan, L. Lipsich, M. Wigler, *Nature (London)* **285**, 207 (1980).
 12. I. Lowy, A. Pellicer, J. Jackson, G. K. Sim, S. Silverstein, R. Axel, in preparation.
 13. J. Littlefield, *Proc. Natl. Acad. Sci. U.S.A.* **50**, 568 (1963).
 14. W. Szybalski, E. H. Szybalska, G. Ragni, *Natl. Cancer Inst. Monogr.* **7**, 75 (1962).
 15. S. Kit, D. Dubbs, L. Piekarski, T. Hsu, *Exp. Cell Res.* **31**, 297 (1963).
 16. W. H. Lewis, P. R. Srinivasan, N. Stokoe, L. Siminovitch, *Somat. Cell Genet.* **6**, 333 (1980).
 17. M. Wigler, M. Perucho, D. Kurtz, S. Dana, A. Pellicer, R. Axel, S. Silverstein, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3567 (1980).
 18. F. L. Graham and A. J. van der Eb, *Virology* **52**, 456 (1973).
 19. M. Wigler, R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, R. Axel, *Cell* **16**, 777 (1979).
 20. J. Spizizen, B. E. Reilly, A. H. Evans, *Annu. Rev. Microbiol.* **20**, 321 (1966).
 21. P. J. Kretschmer, A. C. Y. Chang, S. N. Cohen, *J. Bacteriol.* **124**, 225 (1975).
 22. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
 23. J. C. Fiddes, P. H. Seeburg, F. M. DeNoto, R. A. Hallowell, J. D. Baxter, H. M. Goodman, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4294 (1979).
 24. K. Yamamoto, V. L. Chandler, J. Riug, D. Ucker, *Cold Spring Harbor Symp. Quant. Biol.*, in press.
 25. W. F. Flintoff, S. V. Davidson, L. Siminovitch, *Somat. Cell Genet.* **2**, 245 (1976).
 26. See R. Mulligan and P. Berg, *Science* **209**, 1422 (1980).
 27. A. Pellicer, M. Wigler, R. Axel, S. Silverstein, *Cell* **14**, 133 (1978).
 28. M. Perucho, D. Hanahan, M. Wigler, *Cell*, in press.
 29. D. Robins, A. Henderson, S. Ripley, R. Axel, in preparation.
 30. J. Smiley, D. A. Steege, D. K. Juricek, W. D. Summers, F. H. Ruddle, *Cell* **15**, 455 (1978).
 31. M. Perucho and M. Wigler, in preparation.
 32. D. Hanahan, D. Lane, L. Lipsich, M. Wigler, M. Botchan, *Cell* **21**, 127 (1980).
 33. B. Wold, M. Wigler, E. Lacy, T. Maniatis, S. Silverstein, R. Axel, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5694 (1979).
 34. N. Mantei, W. Boll, C. Weissman, *Nature (London)* **281**, 40 (1979).
 35. R. C. Mulligan, B. H. Howard, P. Berg, *ibid.* **277**, 108 (1979).
 36. D. Hamer and P. Leder, *ibid.* **281**, 35 (1979).
 37. C. Weissman, personal communication.
 38. E. C. Lai, S. L. Woo, M. E. Bordelon-Riser, T. H. Fraser, B. W. O'Malley, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 244 (1980).
 39. R. Breathnach, N. Mantei, P. Chambon, *ibid.*, p. 740.
 40. P. N. Goodfellow, E. A. Jones, V. van Hengnigen, E. Solomon, M. Bobrow, V. Miggiano, W. F. Bodmer, *Nature (London)* **254**, 267 (1975).
 41. Z. Schlegel and T. L. Benjamin, *Cell* **14**, 587 (1978).
 42. T. Maniatis, R. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, G. K. Sim, A. Efstratiadis, *ibid.* **15**, 687 (1978).
 43. W. D. Benton and R. W. Davis, *Science* **196**, 180 (1977).
 44. M. Grunstein and D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961 (1975).
 45. G. A. Galan, W. H. Klein, R. J. Britten, E. H. Davidson, *Arch. Biochem. Biophys.* **179**, 584 (1977).
 46. M. Wigler, R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, R. Axel, in *Eucaryotic Gene Regulation*, R. Axel, T. Maniatis, C. F. Fox, Eds. (Academic Press, New York, 1979), p. 457.
 47. For review see T. Maniatis, E. F. Fritsch, J. Lauer, R. M. Lawn, *Annu. Rev. Genet.*, in press.
 48. B. Steinberg, R. Pollack, W. Topp, M. Botchan, *Cell* **13**, 19 (1978).
 49. For a discussion of this problem for endogenous cellular genes see L. Siminovitch, in *Eucaryotic Gene Regulation*, R. Axel, T. Maniatis, C. F. Fox, Eds. (Academic Press, New York, 1979), pp. 433-443.
 50. R. Sweet, J. Jackson, A. Pellicer, M. Ostrander, S. Silverstein, R. Axel, in preparation.
 51. C. Waalmijk and R. A. Flavell, *Nucleic Acids Res.* **5**, 4031 (1978).
 52. A. P. Bird, M. H. Taggard, B. A. Smith, *Cell* **17**, 889 (1979).
 53. J. L. Mandel and P. Chambon, *Nucleic Acids Res.* **7**, 2081 (1980).
 54. A. H. Hinnen, J. B. Hicks, G. R. Fink, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1929 (1978).
 55. R. DeMars, *Mutat. Res.* **24**, 335 (1974).
 56. A. S. Henderson, M. T. Yu, K. C. Atwood, *Cytogenet. Cell Genet.* **21**, 231 (1978).
 57. This work was supported by grants from the U.S. Public Health Service, National Institutes of Health, to R.A. (CA 16346 and CA 23767) and to S.S. (CA 17477).

7 July 1980

Expression of a Bacterial Gene in Mammalian Cells

R. C. Mulligan* and P. Berg

The most powerful tools now available for studying the molecular anatomy of eukaryote genes and chromosomes, particularly those of higher vertebrates, are restriction enzymes, simple physical methods for separating and visualizing DNA molecules, molecular cloning, and rapid DNA sequencing. In just a few years, the application of these techniques has produced a qualitative change in our views of gene structure and organization in mammalian and other eukaryote organisms. Such newly coined terms as gene libraries, split genes, pseudogenes, and transposons and picturesque references to "shotgunning," or even "walking and jumping along

chromosomes," reflect this revolution and enrich the lexicon of modern genetics. Solving the nucleotide sequence of the chromosomal locus encoding the entire human β -globin-like gene cluster, a region encompassing about 65 kilobase pairs (kbp) (1), would have been considered visionary 10 years ago, but now that prospect looms on the horizon as a feasible undertaking.

Nevertheless, in spite of this impressive progress it is worth considering whether knowing the molecular anatomical details of genes can, by itself, explain the subtleties of gene expression and regulation during growth and development. Put another way, can we deduce the mechanisms of transcription initiation, splicing, and polyadenylation from the nucleotide sequences of isolated genes? Can we expect to learn from the nucle-

otide sequence of the human γ - and β -globin genes why the former is expressed only during fetal life and the latter only in adulthood? Most likely not; at least not without an assay for the biological activity of the genes in question.

About 10 years ago we began to consider how the biological activity of isolated genes could be assayed. That interest coincided with another preoccupation concerned with devising a virus-mediated transducing system for cultured mammalian cells. The overlap of these two interests culminated in a general approach for introducing isolated genes and their modified derivatives into the genomes of cultured mammalian cells (2). Our immediate goal was, and still is, to characterize the physical state, expression, and regulation of the new genes in their transduced hosts.

Because bacteriophages had proved to be so versatile for transducing genes into bacterial cells (3), simian virus 40 (SV40), a mammalian virus, was adopted as the vector to mediate the gene transfer. SV40 was chosen because its minichromosome propagates vegetatively or becomes stably integrated into selected host cell genomes. SV40 was also attractive because its genes and their corresponding functions had been identified and experiments were under way to map the genes to specific regions of the virus's DNA. Subsequently, the entire 5243-base pair (bp) sequence of the circular viral DNA was solved (4, 5), and

R. C. Mulligan is a predoctoral fellow and P. Berg is Willson Professor of Biochemistry in the Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305.

extensive information about the replication, expression, and regulation of the viral genome in different cells has become available (6).

To avoid a dependence on biologically generated transducing viruses and to increase the probability of obtaining specific transducing genomes, we elected to construct, *in vitro*, recombinants between SV40 DNA and the gene of interest. This was accomplished by ligating appropriate DNA fragments to whole (2) or subgenomic segments (7, 8) of viral DNA via enzymatically synthesized homopolymeric cohesive termini (2, 9). Today, ligation of two DNA molecules is usually accomplished via ends that are generated by restriction enzyme cleavage of natural or engineered restriction sites (10).

Initially, the recombinant genomes were propagated as virions and transduction occurred concomitantly with virus infection (7, 8, 10). Since this experimental design requires that the recombinant genomes replicate, they must contain the origin of SV40 DNA replication (*ori*). Furthermore, to encapsidate the recombinant genome into a virion, the DNA molecule must be smaller than 5.3 kbp, that is, about one mature viral DNA length. Since the vector lacks genetic functions coded by the excised DNA segment, the recombinant genomes are defective and, therefore, must be propagated with a helper virus that supplies the missing gene product or products. In our protocol (7, 8, 10) the recombinant genome retains at least one functioning virus gene, and consequently, it can complement a defective gene in the helper virus. For example, recombinants in which the DNA insert replaces all or part of SV40's late region can be propagated with SV40 mutants that have a defective early region (for example, tsA mutants at high temperature); similarly, recombinants in which early region segments are replaced by the DNA implant can be propagated with a helper that is defective in its late region (for example, tsB mutants at high temperature).

Our early attempts to obtain expression of cloned segments as distinct messenger RNA's (mRNA) and proteins were negative (7, 8). Subsequently, recombinant genomes containing either a rabbit β -globin complementary DNA (cDNA) (11), a *Drosophila melanogaster* gene for histone H2b (12) or a cDNA coding for mouse dihydrofolate reductase (DHFR) (13), in place of portions of SV40's late region, were constructed. After infection of monkey cells, each of the recombinants expresses the implanted gene sequence as novel hybrid

mRNA's; moreover, the proteins β -globin (10), histone H2b (14), and mouse DHFR (15) are synthesized at levels comparable to those of SV40 late proteins. Similar successes in obtaining expression of cloned genes have been achieved by Hamer and Leder with recombinants carrying the mouse genomic β -globin (16) or α -globin (17) genes, a result that established proper splicing of the globin intervening sequences and translation of the resulting mRNA's in a heterologous host.

Summary. Transfection of cultured monkey kidney cells with recombinant DNA constructed with a cloned *Escherichia coli* gene that codes for xanthine-guanine phosphoribosyltransferase and several different SV40 DNA-based vectors, results in the synthesis of readily measurable quantities of the bacterial enzyme. Moreover, the physiological defect in purine nucleotide synthesis characteristic of human Lesch-Nyhan cells can be overcome by the introduction of the bacterial gene into these cells.

Until recently, our principal focus has been to exploit the ability of the recombinant genomes to replicate in the virus's permissive host. During vegetative multiplication, the transduced genomes are amplified about 10^4 - to 10^5 -fold, thereby ensuring high yields of the gene products. Such studies have been informative about the necessity and mechanistic subtleties of RNA splicing, the rules governing expression of coding sequences inserted at different loci in SV40 DNA, and about many facets of SV40 gene expression itself. But this experimental design also has a distinct shortcoming: The cell is killed during the course of the infection, thereby precluding the opportunity to monitor the transduced gene's expression in continuously multiplying cells. Moreover, our current collection of recombinants can only be studied in permissive cells, that is, those that can amplify the recombinant genomes. This constraint excludes many specialized and differentiated animal cells as hosts for the transduced genes.

To circumvent these disadvantages, we sought to develop transducing vectors that could be introduced and maintained in a variety of cells. Our initial attempts in this direction indicated that our approach would be facilitated by the availability of a gene whose function could be selected for. Since our experience indicated that the proper positioning of protein coding sequences in SV40 vectors would ensure their expression in transduced cells (18), we explored the possibility of using a bacterial gene for that purpose. The gene chosen was the *Escherichia coli* gene (*Ecogpt*) coding for the enzyme xanthine-guanine phosphoribosyltransferase (XGPRT). The

E. coli XGPRT differs from the analogous mammalian enzyme, hypoxanthine-guanine phosphoribosyltransferase (HGPRT), in that xanthine is considerably more active as a substrate than hypoxanthine in nucleotide synthesis (19). By contrast, mammalian HGPRT does not utilize xanthine efficiently as a substrate; indeed, mammalian cells do not convert xanthine to xanthylic acid or to guanylic acid at a significant rate.

In this article we describe the isolation of *Ecogpt* and its recombination with ap-

propriate SV40 DNA based transducing vectors. Transfection of a variety of cultured mammalian cells with such recombinant DNA's results in the formation of readily measurable quantities of *E. coli* XGPRT. Moreover, human Lesch-Nyhan cells that have been transfected with vectors containing *Ecogpt* DNA grow under conditions in which the parental cells do not survive. This result indicates that *E. coli* XGPRT can overcome the Lesch-Nyhan cell's physiological defect in purine nucleotide synthesis, and suggests that *Ecogpt* may be a generally useful selectable marker for mammalian cells.

Isolation of *Ecogpt* for Introduction into SV40 DNA Vectors

Ecogpt was obtained by a series of manipulations and subcloning operations (Fig. 1). The existence and availability of the transducing phage λ gpt (see legend to Fig. 1) greatly facilitated the isolation in that it provided a more enriched source of the *gpt* gene than *E. coli* DNA itself. Equally crucial to the success of the gene isolation was the availability of suitable Gpt⁻ mutants of *E. coli* (20) whose transformation to Gpt⁺ could be readily monitored. With one notable exception, the subcloning procedure outlined in Fig. 1 employs the same strategy used in obtaining *Ecotdk*, the gene coding for thymidine kinase (8). In this instance, however, the 5' and 3' ends of the *gpt*-containing DNA fragment has been modified to contain either a Hind III or a Bam HI cohesive end. As a result, ligation of the *gpt* fragment to the corresponding cohesive ends of the pBR322 plasmid and

SV40 vectors is more efficient, and isomeric recombinants, which differ only in the orientation of the Gpt relative to the vector, are equally frequent products.

Since both orientational isomers of pBR322-*gpt* were equally efficient in complementing the Gpt⁻ defect of the transformed *E. coli* host, and since the same amounts of XGPRT were induced in the transformants, we surmise that the 1-kbp *Ecogpt* DNA contains all the genetic signals needed to express XGPRT, that is, the transcriptional promoter and terminator as well as the protein coding sequence.

Expression of *Ecogpt* in

Cultured Monkey Cells

We have already described (10) an SV40 vector (SVGT5) which lacks the DNA segment between map position 0.945 and 0.145. Vector SVGT5 still retains the putative SV40 late promoter and polyadenylation sites, as well as the late mRNA leader sequences and splice junctions; it lacks most of the region

specifying the major virus capsid protein. Since SVGT5 was active in promoting the expression of cDNA's, the cloned *Ecogpt* DNA was ligated to SVGT5 via the complementary Hind III and Bam HI cohesive ends and the resulting recombinant, SVGT5-*gpt*, was cloned in monkey cells as previously described (10). As expected, two SVGT5-*gpt* derivatives were obtained, differing only in the orientation of the *gpt* segment. Figure 2 shows the structure of SVGT5-*gpt*, emphasizing the location of the *gpt* segment (the solid black segment). The two orientational isomers of the SVGT5-*gpt* are referred to hereafter as SVGT5-*gpt*(l) and SVGT5-*gpt*(e) to indicate that one recombinant has the *gpt* coding sequence linked to SV40's late strand (l) and the other to the early strand (e), respectively. The structures of the putative *gpt* mRNA's are also shown to emphasize the significance of vector DNA sequences that specify the 5' and 3' ends and the splice junctions of the predominant SV40 mRNA families (21, 22).

Extracts of CV1 cells obtained 36 to 48

hours after infection with SVGT5-*gpt*(l) contain at least 40 times more XPRT activity than uninfected CV1 extracts (inset in Fig. 3A). Insignificant XPRT activity was detected in comparable extracts from cells infected with SV40 or SVGT5-*gpt*(e). Also, Fig. 3A shows that in cells infected with SVGT5-*gpt*(l) nearly half of the enzymatic activity that converts ¹⁴C-labeled guanine to ¹⁴C-labeled guanosine monophosphate (GMP) resembles *E. coli* XGPRT in its insensitivity to inhibition by unlabeled hypoxanthine. By contrast, the same reaction with uninfected cell extracts or extracts obtained after infections with either SV40 or SVGT5-*gpt*(e) is completely inhibited by unlabeled hypoxanthine, a property characteristic of the cellular HGPRT.

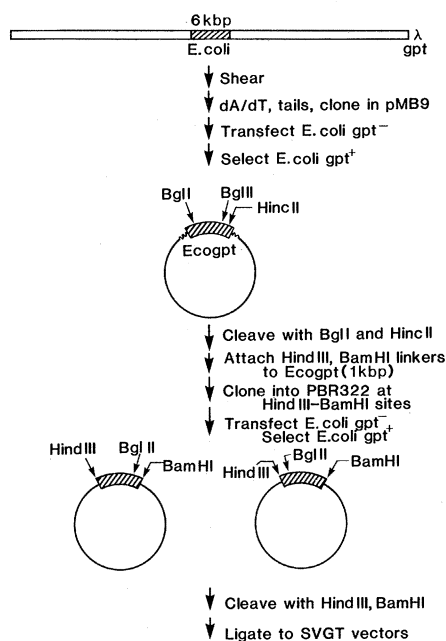
Electrophoretic evidence also supports the contention that *E. coli* XGPRT is formed after infection with SVGT5-*gpt*(l) (Fig. 3B). After polyacrylamide gel electrophoresis of extracts from cells infected with SV40, SVGT5-*gpt*(e), and SVGT5-*gpt*(l) or uninfected cells, assays of GPRT activity in situ show that only SVGT5-*gpt*(l) induces an enzyme activity that comigrates with *E. coli* XGPRT. Each of the other extracts contain monkey cell HGPRT but no detectable *E. coli* XGPRT. From these experiments, we infer that the new enzymatic activity is *E. coli* XGPRT. Note that SVGT5-*gpt*(e) does not produce XGPRT. This might reflect either a failure to transcribe the *gpt* sequence from the *E. coli* promoter or an inability to process that transcript. More likely, however, the *gpt* sequence is expressed only from SV40's promoter, and the transcript is processed by means of SV40 splicing signals. Thus, mammalian cells can express the bacterial gene provided its coding sequence is properly positioned in a suitable vector. An analysis of the precise structure of the mRNA's made by SVGT5-*gpt* as well as other recombinants should enable us to ascertain what features of the recombinant structures are significant for obtaining expression of the implanted DNA segments.

New Vectors for Transducing

Genes into Cells

As indicated earlier, a primary objective of this work is the development of vectors that can be used to introduce and maintain genes of interest in mammalian cells. Such vectors could be integrated into the cell's chromosome or be maintained as part of an autonomously replicating DNA molecule. If such vectors

Fig. 1. Scheme for isolation of a small DNA fragment containing the *E. coli gpt* gene. The DNA used for the isolation of *Ecogpt* was a λ phage containing an insertion of 6 kbp of *E. coli* DNA spanning the *pro gpt lac* region of the *E. coli* chromosome (35). This phage, λ gpt, efficiently transduces a variety of Gpt⁻ strains, including *pro gpt lac* deletion strains, to Gpt⁺. λ gpt DNA was sheared to an average size of 2 kbp and short dT (deoxythymidylate) chains were added to the termini of the sheared fragments as described (2, 7-9). The poly(dT) tailed phage DNA was annealed with poly(dA) tailed Eco RI endonuclease-cleaved pMB9 DNA. Then the recombinant molecules were transfected into *E. coli* JW2 (19), a Gpt⁻ *purE* strain (the *purE* mutation prevents de novo synthesis of purine nucleotides). Tetracycline-resistant (*tet*^r) clones were screened for Gpt⁺ colonies by their ability to grow on a minimal salts medium (M9) containing glucose, vitamin-free casamino acids, thiamine, tryptophan, and 60 μ g of guanine per milliliter. Plasmid DNA isolated from one such colony was found by restriction endonuclease analysis to contain a 1.6-kbp segment of *E. coli* DNA. Transfection of JW2 and *E. coli* GP120—a Gpt⁻ deletion strain (36)—with this plasmid DNA (pPT-1) yielded *tet*^r transformants that were all Gpt⁺. Recombinant plasmids containing the *Ecogpt* segment in the opposite orientation were also obtained and these transformants were also Gpt⁺. The XGPRT activity in extracts from Gpt⁻ or wild-type strains carrying pPT-1 DNA was 20-fold higher than that found in wild-type *E. coli* strains. To obtain a smaller segment encoding the *Ecogpt* gene, pPT-1 DNA was cleaved with Bgl I and Hinc II endonucleases, and the 1-kbp fragment from within the *Ecogpt* insert was isolated by gel electrophoresis. The Bgl I endonuclease generated end of this fragment was made blunt by treatment with *E. coli* DNA polymerase and the four deoxyribonucleoside triphosphates (37). A mixture of decanucleotide linkers containing Hind III and Bam HI endonuclease recognition sequences (Collaborative Research) (each in 40-fold molar excess) were ligated to the blunt ends of the *Ecogpt* fragment with T4 DNA ligase (10). After restriction endonuclease digestion with both Bam HI and Hind III endonucleases, the fragment was ligated to Hind III and Bam HI cleaved pBR322 DNA with T4 DNA ligase (10). Transfection of the Gpt⁻ strain of *E. coli* with the hybrid DNA molecules yielded Gpt⁺ clones that contained both orientational isomers of the *Ecogpt* insert.



contained a selectable genetic marker, retention of the vectors and any non-selectable genes associated with them would be more likely and readily monitored. To accommodate genes of interest, the vectors should contain suitably positioned restriction sites for cloning appropriate DNA fragments. Last, but not least, considerable economies of time and money would be achieved if the manipulations needed to construct the desired recombinants could be performed in bacteria rather than cultured mammalian cells.

The general features of the construction and properties of the first generation of such vectors are shown in Fig. 4 (23). Each vector DNA can propagate as a plasmid in *E. coli*, hence the prefix p before their designations. Both pSV1GT5 and pSV1GT7 contain the SV40 DNA vectors SVGT5 and SVGT7 (18) shown here with the *Ecogpt* segment inserted at the nominal locations between SV40 map positions 0.945 or 0.86 and 0.14, respectively. Joined to the SVGT5 or SVGT7 DNA is pBR322 DNA containing a segment of DNA from SV40's early region (map position 0.14 to 0.325), ligation having occurred at their common Pst I restriction site at SV40 map position 0.28. This construction is convenient because, after preparation in *E. coli*, the SVGT5 or SVGT7 recombinant can be excised with Pst I endonuclease and propagated as a virus chromosome by transfection into monkey cells with the appropriate helper. Since pSV1GT5 or pSV1GT7 contains an intact early region, these recombinants can also replicate in primate cells, the permissive host for SV40. Moreover, because of the approximately 1-kbp duplication of SV40 DNA sequences in the pSV1GT molecules (the segment between map positions 0.145 and 0.325), homologous recombination during replication generates the respective SVGT5 or SVGT7 recombinants in situ.

Shown here with an *Ecogpt* segment, pSV2 consists of a 2.3-kbp portion of pBR322 DNA containing the plasmid's origin of DNA replication and ampicillinase gene joined to a segment of SV40 DNA containing the SV40 early promoter upstream of *Ecogpt*; the *Ecogpt* segment is followed by another SV40 DNA segment containing the small t antigen intervening sequence (24), the early region polyadenylation site and the 3' terminal segment of SV40's late region. Since pSV2-*gpt* DNA lacks the gene coding for SV40 T antigen, it does not replicate in monkey cells (25) although it can be propagated in *E. coli*.

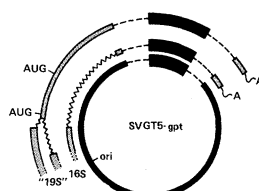


Fig. 2. Structure of SVGT5-*gpt* recombinant and expected structure of the hybrid SV40-*gpt* mRNA's. The pBR322-*Ecogpt* recombinant DNA's containing the 1-kbp *Ecogpt* segment in both orientations (see Fig. 1) were cleaved with Hind III and Bam HI endonucleases to obtain the cloned inserts. These were ligated to SVGT5 DNA and propagated as viruses in CV1 cells as described (10). The circle represents retained SV40 DNA, the dotted portion spans the deleted SV40 late segment, and the solid black segment within the deleted region is the 1-kbp *Ecogpt* DNA fragment in either of the two orientations. Two size classes of hybrid SV40-*gpt* mRNA's are expected. The smaller and larger mRNA's have leaders characteristic of SV40 16S and 19S mRNA's, respectively. The location of the 5' ends of the leaders are left unspecified because of the multiplicity of 5' ends (21, 38). The wavy lines represent regions that are spliced out of the mRNA's.

Plasmids pSV3 and pSV5 also shown here with the *Ecogpt* DNA segment are derivatives of pSV2 that contain at pSV2's single Bam HI restriction site DNA segments which include either (i) a complete early region from SV40 DNA

(pSV3) or (ii) polyoma's early region with two origins of DNA replication (26) (pSV5). Since pSV3 and pSV5 vectors contain viral genes that promote DNA replication at their respective origins, pSV3 recombinants replicate in monkey cells and pSV5 derivatives replicate in mouse cells (25).

Synthesis of *E. coli* XGPRT in CV1 Cells Transfected with pSV-*Ecogpt* DNA's

CV1 cells were transfected with either pSV1GT5-*gpt*, pSV1GT7-*gpt*, pSV2-*gpt*, pSV3-*gpt*, or pSV5-*gpt* DNA (see legend to Fig. 5); after 72 hours cell extracts were subjected to electrophoresis in polyacrylamide slab gels and assayed in situ for GPRT (see legend to Fig. 3B). Each recombinant induced a GPRT activity whose electrophoretic mobility was characteristic of the *E. coli* enzyme (Fig. 5). Since only about 5 to 10 percent of the CV1 cells are transfected by this procedure (27), the amount of *E. coli* GPRT activity is correspondingly lower relative to the amount of cellular GPRT and other proteins in the extract (see Fig. 3B for comparison of the results of infections with virus).

It should be noted that GPRT activity can be detected in DNA transfections

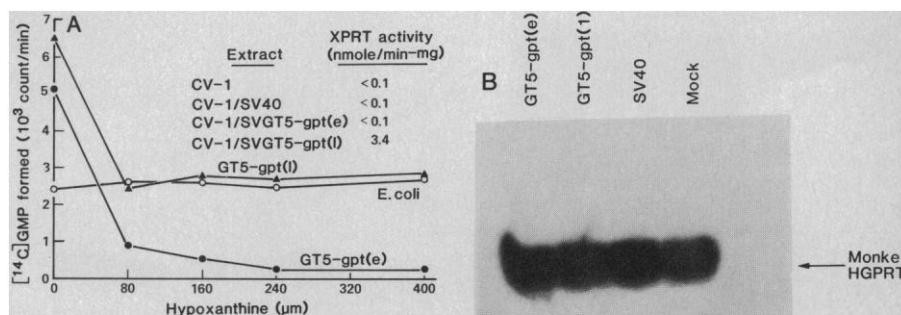


Fig. 3. *E. coli* XGPRT in monkey cells infected with SVGT5-*gpt*. The designations SVGT5-*gpt*(e) and SVGT5-*gpt*(l) are explained in the text. (Inset) XPRT was measured by the conversion of ^{14}C -labeled xanthine (X) to ^{14}C -labeled GMP as follows. Reaction mixtures (50 μl) contained 100 mM tris, pH 7.5, 4 mM MgCl_2 , bovine serum albumin (2 mg/ml), 2 mM phosphoribosylpyrophosphate and 180 μM ^{14}C -labeled xanthine (58 mCi/mmol; Schwarz/Mann). After 15 minutes at 37°C, a portion (5 μl) of the reaction was mixed with unlabeled XMP and applied to cellulose thin-layer plates (Polygram CEL 300, Brinkmann Instruments) and chromatographed in 1M ammonium acetate until the solvent had moved 8 cm. After the plates were dried, the area containing XMP was located under ultraviolet light, cut out, and counted. (A) The activity of the GPRT induced by SVGT5-*gpt* plotted as a function of the concentration of hypoxanthine. The GPRT activity in infected and uninfected CV1 extracts was assayed by the conversion of ^{14}C -labeled guanine to ^{14}C -labeled GMP; the reactions were as described above, except that 20 μM ^{14}C -labeled guanine was used. After the incubation, the reaction mixture was spotted on DE81 filter disks, washed with 10 mM tris, pH 7.4, and counted. (B) Separation of monkey and *E. coli* purine phosphoribosyltransferase activities by polyacrylamide gel electrophoresis and assay of the two enzymes in situ. Extracts of SVGT5-*gpt* infected cells were subjected to electrophoresis on a 7.5 percent polyacrylamide gel containing 0.2M tris, pH 8.5; GPRT activity was assayed in situ as described (38). The ^3H -labeled GMP formed was detected by fluorography and exposure to x-ray film for 24 hours. The locations of *E. coli* and monkey GPRT are indicated by arrows.

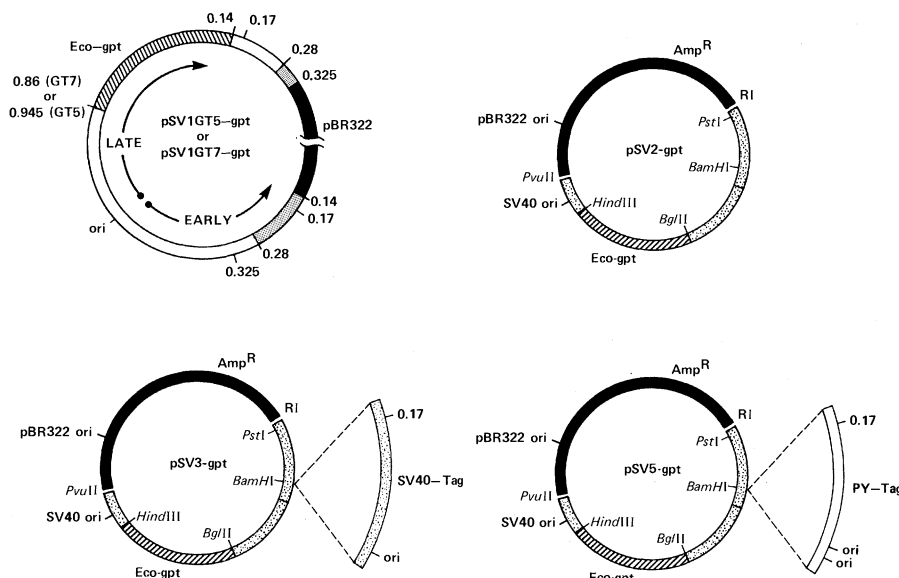


Fig. 4. Structure of plasmid vectors. All the vectors are shown with the *gpt* segment hatched. The solid black segments in each of the diagrams represent pBR322 DNA sequences: 4 kbp in the pSV1 vectors and 2.3 kbp in pSV2, -3, and -5. In each instance, the pBR322 DNA segment contains the origin of pBR322 DNA replication and the ampicillinase gene. In the diagram of the pSV1 vector, the lightly stippled regions are SV40 sequences (from map position 0.14 to 0.325) present in the pBR322 DNA used to construct the pSV1 derivatives; the clear regions indicate the SV40 segment in SVGT5 or -7-*gpt* DNA as originally isolated from infected CV1 cells (see Fig. 2 for SVGT5-*gpt*). The SV40 sequences in pSV2, -3 and -5 are shown as lightly stippled regions.

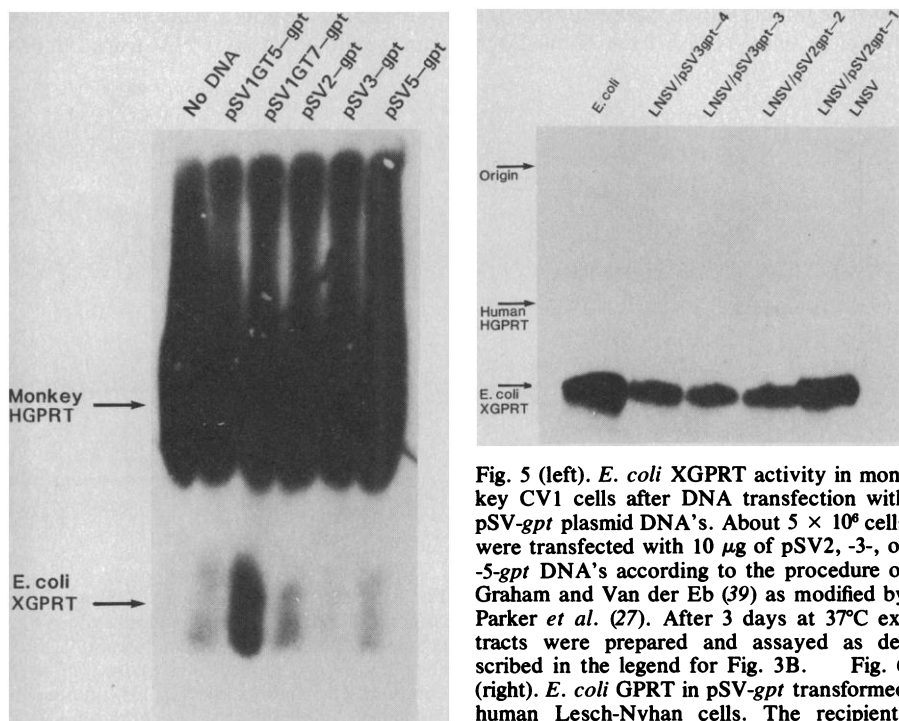


Fig. 5 (left). *E. coli* XGPRT activity in monkey CV1 cells after DNA transfection with pSV-*gpt* plasmid DNA's. About 5×10^6 cells were transfected with 10 μ g of pSV2, -3, or -5-*gpt* DNA's according to the procedure of Graham and Van der Eb (39) as modified by Parker *et al.* (27). After 3 days at 37°C extracts were prepared and assayed as described in the legend for Fig. 3B. Fig. 6 (right). *E. coli* GPRT in pSV-*gpt* transformed human Lesch-Nyhan cells. The recipients were SV40 transformed human fibroblasts derived from a skin biopsy of a patient with Lesch-Nyhan syndrome (40). Approximately 10^6 cells were transfected with 10 μ g of either pSV2-*gpt* or pSV3-*gpt* DNA's and grown for 3 days in Dulbecco-modified Eagle medium (DME) with 5 percent fetal calf serum. Then, 5×10^5 cells were seeded per 100-mm plate in DME containing 10 percent dialyzed fetal calf serum, 10^{-4} M hypoxanthine, 10^{-6} M aminopterin, and 10 μ g/ml each of thymidine and glycine. In the pSV2 and pSV3 transfected cultures, surviving colonies were apparent by the tenth day; the frequency of such colonies was 1×10^{-4} to 2×10^{-4} based on the number of cells plated. Two representative colonies from each transfection were subcultured with cloning cylinders and grown under selective conditions for approximately 20 additional generations prior to analysis. In situ GPRT assays of the extracts of selected clones were performed as described in Fig. 3B.

riated from a skin biopsy of a patient with Lesch-Nyhan syndrome (40). Approximately 10^6 cells were transfected with 10 μ g of either pSV2-*gpt* or pSV3-*gpt* DNA's and grown for 3 days in Dulbecco-modified Eagle medium (DME) with 5 percent fetal calf serum. Then, 5×10^5 cells were seeded per 100-mm plate in DME containing 10 percent dialyzed fetal calf serum, 10^{-4} M hypoxanthine, 10^{-6} M aminopterin, and 10 μ g/ml each of thymidine and glycine. In the pSV2 and pSV3 transfected cultures, surviving colonies were apparent by the tenth day; the frequency of such colonies was 1×10^{-4} to 2×10^{-4} based on the number of cells plated. Two representative colonies from each transfection were subcultured with cloning cylinders and grown under selective conditions for approximately 20 additional generations prior to analysis. In situ GPRT assays of the extracts of selected clones were performed as described in Fig. 3B.

with pSV2-*gpt* DNA, a molecule that cannot replicate in monkey cells. Clearly, additional experiments are needed to relate the efficiency of *gpt* expression to the number of *gpt* gene copies and to the extent of autologous repression that T antigen exerts on *gpt* transcription from the SV40 promoter (28).

Transformation of Human Lesch-Nyhan Cells with pSV2-*gpt* and pSV3-*gpt* DNA's

Human Lesch-Nyhan cells lack HGPRT (29); HGPRT-negative cells, unlike normal ones, cannot grow in a culture medium containing hypoxanthine, aminopterin, and thymidine (HAT medium) (30). Therefore, it was of interest to know whether *E. coli* XGPRT would enable Lesch-Nyhan cells to survive and multiply in HAT medium. In order to test this point, skin fibroblasts, biopsied from a Lesch-Nyhan patient and subsequently immortalized by SV40 transformation (31), were transfected with pSV2-*gpt* or pSV3-*gpt* DNA's, as described in the legend to Fig. 3; transformants were cultured and recovered as in Fig. 6.

Cells that had not received DNA failed to survive in HAT medium, and no clones were detected after 15 days. However, in cultures transfected with pSV2-*gpt* or pSV3-*gpt* DNA, there were detectable colonies within 7 to 10 days, and these were picked and subcultured after 15 days. Approximately 10 to 20 colonies per 10^5 cells were obtained in the two sets of transfections, indicating a transformation frequency, under these conditions, of approximately 10^{-4} . Two isolates each from the pSV2-*gpt* and pSV3-*gpt* transformed cells were tested for the presence of *E. coli* XGPRT after approximately 40 cell divisions in the selective medium (Fig. 6). Quite clearly, both sets of transformants contain the bacterial form of GPRT and, as expected, lack the human GPRT.

pSV2-*gpt* and pSV3-*gpt* transformed Lesch-Nyhan cells have also been subcultured in normal (nonselective) medium for about 12 generations. Extracts of such cells show no detectable diminution of the amount of bacterial GPRT. Although more extensive tests for the continued maintenance and expression of *Ecogpt* are needed, the preliminary indications are that the transformants are genetically stable; that is, segregation of *Ecogpt* is infrequent. The copy number and physical state of the *Ecogpt* DNA in these transformants have not yet been characterized.

Comments and Perspectives

The expression of the prokaryote gene *Ecogpt* in mammalian cells using the lytic vectors SVGT5 and SVGT7 and their plasmid vector analogs, pSV1GT5 and pSV2GT7, probably depends on transcription from the SV40 late promoter; expression of the *gpt* sequence with pSV2, pSV3, and pSV5 vectors probably relies on the function of the SV40 early promoter. The 3' end of the *gpt* mRNA's most likely terminate with polyadenylate [poly(A)] as a result of the SV40 signals that follow the bacterial DNA segment in the vector. Splicing, an essential feature of most mammalian mRNA formation (8, 10, 32, 33), also relies on vector DNA sequences. The late region vector, SVGT5, permits splices in the *gpt* transcript that are characteristic of the SV40 16S and 19S mRNA's, while splicing characteristic of the 19S type mRNA's (21, 22) can occur with SVGT7. Similarly, the *gpt* segment in pSV2, pSV3, and pSV5 precedes a sequence containing the small t antigen intron (24), thereby permitting splicing of those *gpt* transcripts. From our data on the location of the *gpt* coding sequence within the 1-kbp cloned segment and on the characterization of the *gpt* mRNA's produced from the various recombinants (34), we can say that the location of the inserted DNA segment in the vector and the position and nucleotide sequence preceding the inserted gene's initiator AUG determines the efficiency of the new gene's expression.

Emphasized earlier was the fact that recombinant genomes propagated as SV40 virions must be smaller than 5.3 kbp; consequently, cloning with such vectors is limited, in principle, to foreign DNA segments smaller than 5 kbp, but in practice to those smaller than 2.5 kbp. However, because the plasmid vectors pSV1, -2, -3, and -5 are propagated in *E. coli*, there is no constraint on the size of the cloned DNA segment. Moreover, since relatively large quantities of the desired recombinant DNA can be obtained from bacterial cultures, transfections of

mammalian cell cultures with 10 to 20 μ g of recombinant DNA are possible.

Since the introduction and expression of *Ecogpt* permits the survival and growth of Lesch-Nyhan cells in HAT medium, we surmise that *E. coli* XGPRT can provide at least the purine salvage function of the mammalian HGPRT. Apart from the complicated and contentious considerations of potential gene therapy, the successful complementation of an HGPRT deficiency by the *Ecogpt* recombinant DNA's creates interesting opportunities for their use as multipurpose vectors for normal cells. A procedure has been worked out that permits the selection of cells expressing *Ecogpt* from among normal nontransformed cells (34). The selection relies on the unique specificity of the bacterial enzyme for xanthine and the inability of normal cells to use this purine for guanylic acid formation.

Since there is no reason to believe that *Ecogpt* is alone in its ability to be expressed in mammalian cells, it is possible that bacteria, their viruses, and simple eukaryotes provide a rich source of genes specifying suitable products whose expression could be successfully engineered in mammalian cells. In this regard it should be emphasized again that pSV2, -3, and -5 vectors contain suitable restriction sites for recombination with one or more additional DNA segments, for example, at the Eco RI, Pst I, and Bam HI sites in the segment derived from SV40's late region. Accordingly, appropriately modified genes or clusters of genes from any source could be co-transduced with *Ecogpt* in acute transfection or by selection for stable *gpt* transformants.

References and Notes

1. E. F. Fritsch *et al.*, *Cell* **19**, 959 (1980).
2. D. A. Jackson, R. H. Symons, P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2904 (1972).
3. G. S. Stent and R. Calendar, *Molecular Genetics* (Freeman, San Francisco, 1978).
4. W. Fiers *et al.*, *Nature (London)* **273**, 113 (1978).
5. V. B. Reddy, B. Thimmappaya, R. Dhar, K. M. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, S. M. Weissman, *Science* **200**, 494 (1978).

6. J. Tooze, Ed., *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1980), chaps. 2 to 5.
7. S. P. Goff and P. Berg, *Cell* **9**, 695 (1976).
8. ———, *J. Mol. Biol.* **133**, 359 (1979).
9. P. E. Lobban and A. D. Kaiser, *ibid.* **78**, 453 (1973).
10. R. C. Mulligan, B. H. Howard, P. Berg, *Nature (London)* **277**, 108 (1979).
11. A. Efstratiadis, F. C. Kafatos, T. Maniatis, *Cell* **10**, 571 (1977).
12. R. P. Lifton, M. L. Goldberg, R. W. Karp, D. S. Hogness, *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1047 (1978).
13. A. C. Y. Chang, J. H. Nunberg, R. J. Kaufman, H. A. Ehrlich, R. T. Schimke, S. N. Cohen, *Nature (London)* **275**, 617 (1978).
14. H. Hofstetter and P. Berg, unpublished result.
15. S. Subramani, R. C. Mulligan, P. Berg, unpublished result.
16. D. H. Hamer and P. Leder, *Nature (London)* **281**, 35 (1979).
17. D. H. Hamer, M. Kaehler, P. Leder, *Cell*, in press.
18. R. C. Mulligan, R. T. White, P. Berg, *Miami Winter Symp.*, in press.
19. R. L. Miller, G. A. Ramsey, T. H. Krenitsky, G. B. Elion, *Biochemistry* **11**, 4723 (1972).
20. J. A. Holden, P. D. Harriman, J. D. Wall, *J. Bacteriol.* **126**, 1141 (1976).
21. P. K. Ghosh, V. B. Reddy, J. Swinscoe, P. Lebowitz, S. M. Weissman, *J. Biol. Chem.* **253**, 3643 (1978); *J. Mol. Biol.* **126**, 813 (1978).
22. C.-J. Lai, R. Dhar, G. Khoury, *Cell* **14**, 971 (1978).
23. B. H. Howard, R. C. Mulligan, P. Berg, a detailed description of the structure and properties of these vectors is in preparation.
24. V. B. Reddy, P. K. Ghosh, P. Lebowitz, M. Piatak, S. M. Weissman, *J. Virol.* **30**, 279 (1979).
25. P. J. Southern and P. Berg, unpublished data.
26. G. Magnusson and M.-G. Nilsson, *J. Virol.* **22**, 646 (1977).
27. B. A. Parker and G. R. Stark, *ibid.* **31**, 360 (1979).
28. J. C. Alwine, S. I. Reed, G. R. Stark, *J. Virol.* **24**, 22 (1977).
29. J. E. Seegmiller, F. M. Rosenbloom, W. H. Kelley, *Science* **155**, 1682 (1967).
30. J. Littlefield, *ibid.* **145**, 709 (1964).
31. C. M. Croce, B. Bakay, W. L. Nyhan, H. Kowrowski, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2590 (1973).
32. D. H. Hamer and P. Leder, *Cell* **18**, 1299 (1979).
33. C.-J. Lai and G. Khoury, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 71 (1979).
34. R. C. Mulligan and P. Berg, in preparation.
35. Obtained from P. D. Harriman, University of Missouri, and characterized by R. C. Mulligan and Harriman.
36. Obtained from J. Gots, University of Pennsylvania, Philadelphia.
37. P. H. Seeburg, J. Shine, J. A. Martial, J. D. Baxter, H. M. Goodman, *Nature (London)* **270**, 486 (1977).
38. L. A. Chasin and G. Urlaub, *Somat. Cell Genet.* **2**, 453 (1976).
39. F. L. Graham and A. J. van der Eb, *J. Virol.* **52**, 455 (1973).
40. The LNSV cells were sent to us by B. Bakay and W. L. Nyhan, University of California, San Diego.
41. Supported by grants from NIH (GM 13235), NCI (CA 15513), and the American Cancer Society (MV-351) and by an NSF predoctoral fellowship (to R.C.M.).

* Present address: Center for Cancer Research, Massachusetts Institute of Technology, Cambridge 02139, and Endocrine Unit, Massachusetts General Hospital, Boston.

20 June 1980