Expression of the Mouse Dihydrofolate Reductase Complementary Deoxyribonucleic Acid in Simian Virus 40 Vectors

SURESH SUBRAMANI, RICHARD MULLIGAN,[†] AND PAUL BERG*

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

Received 8 May 1981/Accepted 16 June 1981

A mouse complementary deoxyribonucleic acid segment coding for the enzyme dihydrofolate reductase has been cloned in two general classes of vectors containing simian virus 40 deoxyribonucleic acid: (i) those that can be propagated as virions in permissive cells and (ii) those that can be introduced into and maintained stably in various mammalian cells. Both types of vectors express the mouse dihydrofolate reductase by using signals supplied by simian virus 40 deoxyribonucleic acid sequences. Moreover, plasmid vectors carrying the complementary deoxyribonucleic acid segment can complement Chinese hamster ovary cells lacking dihydrofolate reductase.

The development of vectors to introduce isolated genes and their modified derivatives into cultured mammalian cells has provided a powerful tool for the analysis of gene expression and regulation (8, 9, 11-13, 23, 25, 26). This paper describes the expression of a complementary deoxyribonucleic acid (cDNA) coding for mouse dihydrofolate reductase (*Mo-dhfr*) (4) in two classes of vectors (23) in which the signals required for transcription initiation, termination, polyadenylation, and splicing are supplied by simian virus 40 (SV40) DNA sequences.

The first class of recombinants, which can be propagated as virions, contains the *Mo-dhfr* cDNA in two locations within the late region of SV40. The recombinants, SVGT5-*dhfr* and SVGT7-*dhfr*, contain the *dhfr* cDNA in place of the SV40 DNA segment from map positions 0.945 through 0.145 and 0.86 through 0.145, respectively. These recombinant viral genomes are propagated with a temperature-sensitive helper virus (*tsA58*) in permissive monkey cells (8, 9, 25) and produce cytoplasmic, polyadenylated messenger ribonucleic acids (mRNA's) of the expected structures containing *dhfr* sequences. Moreover, cultured monkey kidney cells infected with SVGT5-*dhfr* or SVGT7-*dhfr* synthesize at least 20-fold more dihydrofolate reductase (DHFR) than do mock-infected cells.

The *dhfr* cDNA has also been cloned in plasmid vectors whose construction will be described elsewhere (23; B. Howard, P. Southern, R. Mulligan, and P. Berg, manuscript in preparation; R.

Mulligan, P. Southern, M. Yaniv, A. Geller, and P. Berg, manuscript in preparation). In these vectors the cDNA is either under the control of the SV40 late promoter (e.g., pSV1GT5-*dhfr* and pSV1GT7-*dhfr*) or the SV40 early promoter (e.g., pSV2-*dhfr*, pSV3-*dhfr*, and pSV5-*dhfr*). These plasmid vectors also express the *dhfr* sequence as judged by their ability to transform Chinese hamster ovary (CHO) cells lacking DHFR to a normal phenotype.

MATERIALS AND METHODS

Restriction enzymes were purchased from Bethesda Research Laboratories or New England Biolabs. [³H]-folate (potassium salt; 5 Ci/mmol) was from Amersham Corp. Ham F12 medium (containing thymidine, glycine, and hypoxanthine for growth of CHO cells lacking DHFR) was obtained from GIBCO Laboratories. T4 DNA ligase was the gift of S. Scherer, Stanford University. *Escherichia coli* DNA polymerase I and S1 nuclease were from Boehringer Mannheim. Fetal and newborn calf serum were obtained from Microbiological Associates. HindIII linkers were purchased from Collaborative Research, Inc. The 1.6-kilobase (kb) mouse *dhfr* cDNA (clone *dhfr* 26) joined by GC bridges in the *Pst*I site of pBR322 (4) and the rabbit antibody directed against the mouse DHFR were the gifts of R. T. Schimke, Stanford University. The DHFR-negative and normal CHO cells were obtained from L. Chasin, Columbia University.

Animal cells and viruses. Conditions for the growth of CV1 and CV1P monkey cells and the procedure for propagation of SV40 mutants with helper virus have been described by Mertz and Berg (22), except that cells were grown in newborn calf serum rather than fetal calf serum. The DHFR-negative CHO line and its wild-type progenitor (CHO-K1) (5) were grown in Ham F12 and Dulbecco modified Eagle

[†] Current address: Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139.

medium, respectively, both media being supplemented with 10% newborn calf serum, penicillin, streptomycin, and 150 μ g of L-proline per ml. Selective medium for transformation of the mutant cell line was Dulbecco medium containing 10% dialyzed fetal calf serum, penicillin, streptomycin, and L-proline.

Isolation of the mouse *dhfr* cDNA segment. The 1.6-kb *dhfr* cDNA, which had been cloned in the *Pst*I site of pBR322 by using GC tails (4), was digested with *Pst*I endonuclease and isolated by gel electrophoresis (Fig. 1). The ends of the DNA fragment (1 μ g) were made blunt by incubation with S1 nuclease (4 U) for 20 min at 45°C in 30 mM sodium acetate (pH 4.5)–280 mM NaCl–4.5 mM zinc acetate. The DNA was then electrophoresed in an agarose gel (1%), extracted from the gel by a modification (B. Paterson, unpublished data) of the glass powder technique (36) and treated with DNA polymerase I for 1 h at 12°C in 40 mM tris(hydroxymethyl)aminomethane-hydrochloride (pH 7.5)–80 mM MgCl₂–10 mM β -mercaptoethanol and 2 mM each of the four deoxynucleotide triphosphates. The DNA polymerase I was inactivated by heating at 70°C for 5 min, and the DNA was then ligated with 50 pmol of *Hind*III linkers (which had been treated with polynucleotide kinase) and T4 DNA ligase for 10 h at 12°C (14). The fragment was separated from excess linkers by agarose gel electropho-

resis and digested first with *Hind*III and then with *Bgl*II endonucleases to obtain the 735-base-pair (bp) cDNA containing *Hind*III and *Bgl*II cohesive termini at the 5' and 3' ends, respectively (Fig. 1). This modified fragment could be cloned in various vectors (see Fig. 2 and 8).

Construction of vectors. The construction and properties of the viral vectors SVGT5 (25) and SVGT7 (26) have already been reported, and corresponding descriptions of the plasmid vectors pSV1GT5, pSV1GT7, pSV2, pSV3, and pSV5 have appeared recently (23, 24). The various *dhfr* cDNA derivatives were prepared as follows.

(i) **pSV1GT5-*dhfr* and pSV1GT7-*dhfr*.** The recombinant pSV1GT5-*gpt* (26) was digested partially with *Hind*III endonuclease, and full-length linear DNA was purified by electrophoresis on an agarose gel (0.5%). The linear DNA was digested with *Bam*HI and *Bgl*II endonucleases (to eliminate a contaminating fragment) and electrophoresed on an agarose gel (0.5%), and the 8.75-kb or 8.7-kb DNA fragments corresponding to the vectors pSV1GT5 and pSV1GT7, respectively, were extracted. These vectors, with *Hind*III and *Bam*HI cohesive ends, were ligated to the 735-bp *dhfr* cDNA. HB101 was transformed to ampicillin resistance with the ligation mixture, and the DNAs from several colonies were screened by digestion with restriction endonucleases to identify those whose plasmids conformed to the predicted pSV1GT5-*dhfr* and pSV1GT7-*dhfr* structures. These two plasmid DNA recombinants were prepared in large amounts and banded twice on CsCl-ethidium bromide gradients.

(ii) **SVGT5-*dhfr* and SVGT7-*dhfr*.** SVGT5-*dhfr* and SVGT7-*dhfr* were made by digestion of pSV1GT5-*dhfr* and pSV1GT7-*dhfr* DNAs with *Pst*I endonuclease and recyclicalization (at concentrations of less than 1 μ g of DNA per ml) of the SVGT5-*dhfr* or the SVGT7-*dhfr* DNA fragment (these fragments lack pBR322 sequences). Then CV1P cells were transfected with either SVGT5-*dhfr* or SVGT7-*dhfr* DNA (0.001 to 1 ng per plate) and *tsA58* DNA (10 ng per plate) to obtain plaques at 41°C. The virus in the plaques was used to produce the corresponding DNA and high-titer virus stocks (22). About 60% of the plaques picked with each recombinant contained virions with the predicted structures of SVGT5-*dhfr* and SVGT7-*dhfr* genomes as well as *tsA58*.

pSV2-*dhfr*, pSV3-*dhfr*, and pSV5-*dhfr*. The *dhfr* cDNA was cloned in the pSV2 vector containing *Hind*III and *Mbo*I sites (Howard et al., in preparation) to generate pSV2-*dhfr*. pSV3-*dhfr* was made from pSV2-*dhfr* by insertion of the 3-kb SV40 early region (the DNA fragment from the *Hpa*II site through the origin of replication to the *Bam*HI site, carrying *Bam*HI linkers at either end) into the unique *Bam*HI site of pSV2-*dhfr*. pSV5-*dhfr* was made in an analogous manner by insertion of the 3.5-kb DNA segment corresponding to the polyoma clone 411 (21) early region (DNA from the *Hinc*II site to the *Bam*HI site, with *Bam*HI linkers at both ends) into the *Bam*HI site of pSV2-*dhfr*.

Analysis of mRNA's. Cytoplasmic, polyadenylated mRNA's from CV1 cells infected at a multiplicity of 1 to 10 plaque-forming units per cell, with SVGT5-

FIG. 1. Isolation of the mouse *dhfr* cDNA segment. The 1.6-kb mouse *dhfr* cDNA which had been cloned by using GC tails in the *Pst*I site of pBR322 (4) was excised by using *Pst*I endonuclease and trimmed to a length of about 735 bp as described in the text. This 735-bp DNA has *Hind*III and *Bgl*II cohesive ends at its 5' and 3' termini, respectively. The 558-bp coding sequence (excluding the ATG and TAA codons) is indicated by the solid black segment, and the 5' and 3' noncoding regions are shown as clear segments.

dhfr and SVGT7-*dhfr* were isolated as described previously (25). Northern blots of these RNAs were done by the method of Alwine et al. (1), using nick-translated (32) DNA probes. Analyses of the RNAs by the method of Berk and Sharp (2) were done with uniformly labeled DNA probes. After hybridization of the RNAs with an excess of the labeled DNA fragments, the hybrids were digested with S1 nuclease, precipitated with ethanol, denatured in NaOH, and analyzed by electrophoresis on 1.5% alkaline agarose gels (30 mM NaOH, 1 mM ethylenediaminetetraacetic acid).

Protein extracts from infected cells. CV1 cells infected with SVGT5-*dhfr* or SVGT7-*dhfr* (1 to 10 plaque-forming units per cell) were washed twice (60 h postinfection) with ice-cold tris(hydroxymethyl)aminomethane-saline. The cells were scraped into a tube after the addition of 0.25 ml of cold tris(hydroxymethyl)aminomethane-saline containing 15% glycerol and 10 mM dithiothreitol and ruptured by sonication for 2 min, and then samples were made 0.1% in Nonidet P-40 and 0.05% in deoxycholate. After 10 min at 0°C the extracts were spun at 15,000 rpm for 30 min in a Sorvall centrifuge, and the supernatants were stored at -80°C. Little or no loss of DHFR activity occurred upon five cycles of freezing and thawing the extracts.

Assays for DHFR. DHFR activity was measured by the conversion of [³H]folate to [³H]tetrahydrofolate as described by Rothenberg (34) and modified by Nakamura and Littlefield (27). Alternatively, [³H]folate, reduced by dithionite to [³H]dihydrofolate (3), was used as substrate. In some experiments, the enzyme was assayed spectrophotometrically by following the utilization of reduced nicotinamide adenine dinucleotide phosphate (measured at 340 nm) (33). Protein concentrations were determined by the method of Lowry et al. (20). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was done by the method of Laemmli (18).

Transfection of CHO cells lacking DHFR. Between 12 to 18 h after seeding 10⁶ DHFR-negative CHO cells on 100-mm plates, the cultures were transfected with 10 µg of the appropriate pSV-*dhfr* DNA per plate by the procedure of Graham and Van der Eb (10) as modified by Parker and Stark (29). After growth of the cells for 3 days in nonselective medium (Ham F12), one-half of the cells from each confluent plate were redistributed in selective medium on five 100-mm plates. The selective medium was changed every 3 days; after 4 to 7 days, colonies were visible. About 2 weeks after the cells were seeded in selective medium, the colonies were fixed with methanol, stained with Giemsa stain, and counted.

RESULTS

Features of the mouse *dhfr* cDNA. The 735-bp cDNA used in these studies was isolated as described above and in the legend to Fig. 1. Preceding the ATG that codes for the first amino acid of the DHFR is a sequence of 84 nucleotides (28) containing no other ATGs. Four translation termination codons occur in the 5' untranslated region; three of these are in the same translation

frame, whereas the fourth is in another reading frame. The coding sequence for DHFR is 564 bp long and ends at the TAA termination codon located 81 bp upstream from the *Bgl*II site (28).

Expression of mouse DHFR from viral vectors. The *dhfr* cDNA was inserted into two locations of the late region of SV40. In SVGT5-*dhfr* and SVGT7-*dhfr* the cDNA replaces viral DNA sequences from the *Hind*III sites at map units 0.945 and 0.86, respectively, to the *Bam*HI site at map coordinate 0.145 (Fig. 2). Thus, the 735-bp *dhfr* cDNA replaces 1,040 bp of SV40 DNA in SVGT5-*dhfr* and 1,487 bp of SV40 DNA in SVGT7-*dhfr*. SVGT5-*dhfr* and SVGT7-*dhfr* DNA are, therefore, 305 bp and 752 bp smaller, respectively, than SV40 DNA. The structures of these recombinant genomes have been confirmed by digestion with various restriction enzymes (data not shown). The two recombinants, which are defective in the late region of SV40, can be propagated as virions in permissive monkey cells in the presence of temperature-sensitive mutants defective in the early function, e.g., *tsA58*. The recombinants grow as well as the helper virus and produce virus stocks containing nearly equal numbers of recombinant and helper genomes.

mRNA's made by SVGT5-*dhfr* and SVGT7-*dhfr*. The mRNA's expected to be synthesized from the two recombinants are shown in Fig. 2. In the SVGT5 recombinant the *dhfr* cDNA is 30 nucleotides downstream from the acceptor splice site (at map unit 0.945) that normally functions in the production of late 16S viral RNA. As a result, SVGT5-*dhfr* would be expected to make a 16S mRNA, i.e., one containing a 200 to 250-nucleotide leader spliced to sequences at map position 0.94 and the *dhfr* sequence to map position 0.145 and followed by SV40 nucleotides and a polyadenylated tail. Beside this RNA, SVGT5-*dhfr* should also yield 19S mRNA's with alternative leaders (schematized in Fig. 2) and spliced out sequences originating from either map positions 0.73 through 0.765 or 0.76 through 0.765. These two 19S RNAs have the same bodies and terminate at the same SV40 polyadenylation site as the 16S mRNA's. Since the acceptor splice site for the 16S RNA has been deleted in SVGT7-*dhfr*, this recombinant should produce only the 19S mRNA's (Fig. 2).

To test these predictions CV1 cells were infected with each of the recombinants, SV40 or a mock sample. The cytoplasmic, polyadenylated RNA was isolated 60 h postinfection, treated with glyoxal, electrophoresed in parallel tracks on agarose gels, and transferred to diazobenzyl-oxymethyl paper. The RNAs made in the var-

FIG. 2. Structure of SVGT5-dhfr and SVGT7-dhfr mRNA's expected from the recombinants. The dhfr cDNA from Fig. 1 was cloned between the HindIII sites and the BamHI site in the late region of SV40, replacing viral DNA from map units 0.945 through 0.145 in SVGT5-dhfr and from map coordinates 0.86 through 0.145 in SVGT7-dhfr. The origin of SV40 DNA replication is labeled ori. The thick circles represent SV40 regions retained; the dotted portions span sequences deleted from SV40 and the DNA replacing the deleted segment is the dhfr cDNA. Two size classes of mRNA's (19S and 16S) are expected for SVGT5-dhfr, and only one (19S) is expected for SVGT7-dhfr. These are based on the known structures of the SV40 late mRNA's. The 19S dhfr mRNA's are shown as having two types of leaders (7, 19), whereas the smaller message has the major 16S late mRNA leader. The 5' ends are unspecified because of the multiplicity of reported 5' termini (7, 19). Wavy lines represent regions spliced out of each RNA. AUGs, in clockwise order, denote translation initiation codons for VP2 and VP3, respectively. The A at the ends of the mRNA's signifies the polyadenylate tail.

ious infections were identified by hybridization of the blots with nick-translated DNA probes (Fig. 3).

The late mRNA's made by the helper were

detected in the tracks on the left by using as probe the 825-bp HindIII-BamHI SV40 DNA fragment from map positions 0.985 through 0.145, cloned between the HindIII and BamHI sites in pBR322. In cells infected with either recombinant this probe hybridizes to the 19S and 16S RNAs made by the tsA58 helper.

The RNAs containing dhfr sequences were detected in the tracks on the right of Fig. 3 by using labeled dhfr cDNA as probe. Although RNA from mock- or SV40-infected cells contained no dhfr sequences, there were clearly 19S and 16S mRNA's containing dhfr sequences in cells infected with SVGT5-dhfr. These were about 300 nucleotides shorter than the corresponding mRNA's of SV40 because SVGT5-dhfr is about that much smaller than SV40 in the late region. As anticipated, infections with SVGT7-dhfr yielded only 19S mRNA's, which were the expected 750 nucleotides shorter than the SV40 19S late RNA.

Analysis of the spliced structure of dhfr mRNA's. The method of Berk and Sharp (2) was used to analyze the spliced structure of the dhfr mRNA's made by the above recombinants. DNA probes, labeled uniformly with ^{32}P , were made from each recombinant (Fig. 4A and B). The probes were made by digesting each recombinant DNA with BglI (a site within the SV40 origin of replication) and TaqI (a site about 90 nucleotides into the dhfr cDNA and immediately after the dhfr initiator ATG). These BglI-TaqI DNA probes from SVGT5-dhfr (1,583 bp, Fig. 4A) and SVGT7-dhfr (1,136 bp, Fig. 4B) were purified by agarose gel electrophoresis and hybridized (in DNA excess) to cytoplasmic, polyadenylated RNA from cells infected with either SV40 (to identify DNA fragments protected from S1 nuclease digestion by tsA58 RNA) or the corresponding recombinants. The RNA-DNA hybrids were digested with S1 nuclease and then electrophoresed on alkaline agarose gels (Fig. 5).

As shown schematically in Fig. 4A, late mRNA made by the tsA58 helper should protect a 200- to 240-nucleotide fragment corresponding to the 16S and 19S RNA leaders. Two other fragments, 935 and 30 nucleotides in length, should also be protected by portions of the bodies of the 19S and 16S late RNAs, respectively. Of these, the 200- to 250- and the 935-nucleotide bands are shown in Fig. 5, track 3.

With the same probe, assuming that the splicing patterns are as outlined in Fig. 2, RNAs from SVGT5-dhfr should protect the same 200- to 250-nucleotide fragment and the 1,025- and 120-nucleotide fragments (Fig. 4A and Fig. 5, track 2), corresponding to segments of the 19S and

FIG. 3. mRNA's made in CV1 cells infected with SVGT5-dhfr and SVGT7-dhfr recombinants. Cytoplasmic, polyadenylated mRNA's were isolated from CV1 cells 48 to 60 h after infection at a multiplicity of 1 to 10 plaque-forming units per cell with mixed virus stocks of *tsA58* and either SVGT5-dhfr or SVGT7-dhfr as described in the text. RNAs were also obtained from mock- and SV40-infected cells. The RNAs were treated with glyoxal, electrophoresed on agarose gels, and then transferred to DBM paper. The four tracks on the left were probed with 32 P-labeled helper-specific DNA (see text) that hybridizes only to the late mRNA's made by SV40 and *tsA58*, whereas the four tracks on the right were probed with the *dhfr* cDNA. The positions of the 19S and 16S mRNA's made by SV40 and the recombinants (based on analogy to corresponding RNAs of SV40) are indicated at the sides.

16S RNA bodies. The 1,583-nucleotide band seen in track 2 and 3 of Fig. 5 is the renatured probe. Thus, the RNAs made by SVGT5-dhfr appeared to be spliced as indicated in Fig. 2.

With RNA from SVGT7-dhfr-infected cells, hybridization of the helper RNA with the SVGT7-dhfr *Bgl*I-*Taq*I DNA fragment yields the expected 200- to 250-nucleotide fragment (16S and 19S RNA leaders) and a 488-nucleotide fragment (protected by the 19S RNA body) (Fig. 4B and Fig. 5, track 5). The 488-nucleotide band produced by the 19S RNA body of the helper is not visible in the exposure in Fig. 5, lane 4, because the amount of 19S RNA made by *tsA58* was lower than that of the 19S RNA made by SVGT7-dhfr. The band at 1,136 nucleotides is the renatured probe. With RNA from SVGT7-dhfr-infected cells, there should be, besides the bands produced by the helper RNA, an additional 578-nucleotide fragment (19S RNA body; see Fig. 4B); this band is prominent in Fig. 5, track 4. However, there were also lesser amounts of two larger fragments (about 860 and 720 nucleotides). These DNA fragments correspond to RNAs lacking an interruption in the region between map positions 0.74 and 0.86 as judged by additional S1 nuclease analyses with other DNA probes from *tsA58* (data not shown). These *dhfr* RNAs occur at only 0.1 to 0.2 times the abundance of the spliced *dhfr* mRNA's made by SVGT7-dhfr. They are either spliced at se-

quences upstream of map positions 0.72 through 0.74 or have 5' ends at this location.

Cells infected with SVGT5-dhfr and SVGT7-dhfr make DHFR. To examine the proteins produced in CV1 cells infected at a multiplicity of 1 to 10 plaque-forming units with SVGT5-dhfr and SVGT7-dhfr virions, cell extracts (see above) were electrophoresed in a sodium dodecyl sulfate-polyacrylamide gel; comparable analyses were done with mock- or SV40-infected cells. The major viral capsid protein VP1, easily identifiable in SV40-infected cell extracts, is also made by the *tsA58* helper virus in SVGT5-dhfr- and SVGT7-dhfr-infected cells (Fig. 6). Note that in cells infected with recombinants there is a protein which is absent in mock- and SV40-infected cells. The electrophoretic mobility of this protein indicates a molecular weight of 22,000, that of mouse DHFR (4).

To determine whether this protein band corresponded to mouse DHFR, the extracts were electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel containing a cleavable cross-linker; then the cross-linker, *N,N'*-diallyltartardiamide, was cleaved with periodic acid as described by Renart et al. (31). The proteins were transferred to DBM paper and reacted with antibody directed against mouse DHFR, and the antigen-antibody complexes were labeled by incubation of the DBM paper with 125 I-labeled *Staphylococcus aureus* protein A (Fig. 7). Quite

FIG. 4. Analysis of the spliced structure of the *dhfr* mRNA's. The DNA probes and the fragments expected on alkaline gels after S1 nuclease digestion are shown schematically (see Fig. 2). The uniformly labeled DNA probes from (A) SVGT5-*dhfr* and (B) SVGT7-*dhfr* span the region from the BglI site on SV40 to the TaqI site about 90 nucleotides into the *dhfr* cDNA. The open segments of the DNA probes represent SV40 sequences, whereas the shaded portions represent *dhfr*. The map positions and cleavage sites for restriction endonucleases of interest are shown above and below the DNA probes, respectively. Gaps in the mRNA's above the probe represent splices and the numbers above the mRNA's denote the length of each segment in nucleotides.

clearly, the 22,000-molecular-weight protein made in SVGT5-*dhfr*- and SVGT7-*dhfr*-infected cells shared antigenic determinants with mouse DHFR.

Extracts of SVGT5-*dhfr*- and SVGT7-*dhfr*-infected CV1 cells also had at least 20-fold more DHFR enzyme activity than did mock-infected cells. Measurements of the conversion of [³H]dihydrofolate to [³H]tetrahydrofolate indicate that SVGT5-*dhfr* and SVGT7-*dhfr* had 275

and 273 U of DHFR activity (1 U equals the formation of 1 pmol of tetrahydrofolate per 15 min per μ g of protein), respectively, whereas mock-infected CV1 cells had only 12.5 U of enzyme activity and SV40-infected cells contained 19 U of DHFR activity under similar conditions. Comparable results were obtained when the conversion of [³H]folate to [³H]tetrahydrofolate was assayed.

These results indicate that substantial quantities of functional mouse DHFR are made when SVGT5-*dhfr* or SVGT7-*dhfr* infects monkey cells.

Expression of *dhfr* from plasmid vectors. The viral vectors described above produce substantial quantities of DHFR protein because they are amplified more than 10⁵-fold during the vegetative growth phase, but this mode of transduction is limited to cells permissive for SV40 replication and results in the death of the infected cells.

To overcome these disadvantages the *dhfr* cDNA was introduced into a set of plasmid vectors (Fig. 8), several of which can be stably maintained in mammalian cells. The structure

FIG. 5. Berk-Sharp analysis of *dhfr* mRNA's. RNAs made by SV40, SVGT5-*dhfr*, and SVGT7-*dhfr* were hybridized (in DNA excess) to uniformly labeled DNA probes (from the BglI site in SV40 to a TaqI site just past the ATG in the *dhfr* cDNA) from SVGT5-*dhfr* and SVGT7-*dhfr*. The hybrids were treated with S1 nuclease, and the protected DNAs were analyzed on alkaline agarose gels. The figure shows the protected DNA fragments. The track on the left shows size markers.

FIG. 6. Proteins made in SVGT5-*dhfr*- and SVGT7-*dhfr*-infected CV1 cells. Cells were either mock infected or infected with SV40, SVGT5-*dhfr*, or SVGT7-*dhfr* recombinants. Cell extracts obtained 48 to 60 h postinfection were electrophoresed on sodium dodecyl sulfate-polyacrylamide gels and stained with Coomassie blue. Molecular weight markers for the proteins are at the left. The positions of VP1 and mouse DHFR (arrow) are shown at the right.

and properties of these vectors have been described elsewhere (23; Howard et al., in preparation; Mulligan et al., in preparation). Briefly, with pSV2-*dhfr*, pSV3-*dhfr*, and pSV5-*dhfr* DNAs, the *dhfr*-coding region is expressed from the early region promoter of SV40, and the resulting transcript can be spliced by using the small t antigen intervening sequence and polyadenylated at the SV40 early region polyadenylation site. With pSV1GT5-*dhfr* and pSV1GT7-*dhfr* the *dhfr* segment is transcribed from the SV40 late promoter by using the SV40 late region splicing and polyadenylation signals.

To determine whether these plasmid *dhfr* DNAs could transform mammalian cells for the *dhfr* function, we determined whether they could complement DHFR-negative CHO cells. DHFR-negative CHO cells can not synthesize tetrahydrofolate and, therefore, require thymidine, glycine, and purines to grow. Consequently, after infection of the mutant CHO cells with the *dhfr* plasmids, transformation for the *dhfr* marker can be scored by the growth of cells into colonies in the absence of thymidine, glycine, and purines (Table 1). No colonies were detected on mock-infected plates, indicating that reversion of the mutant CHO cells to a wild-type phenotype is rare. pSV2-*dhfr* and pSV5-*dhfr*

DNAs gave the same transformation frequency (1×10^{-4}). pSV3-*dhfr*, pSV1GT5-*dhfr*, and pSV1GT7-*dhfr* had 5- to 10-fold-lower frequencies (2×10^{-5}).

The mutant CHO cells were devoid of measurable DHFR activity (less than 0.5% of that of the wild type) by several different assays (5). By contrast, the pSV2-*dhfr*-transformed CHO clones had twice the level of DHFR activity of wild-type (CHO-K1) cells (13 versus 6 U of activity, respectively). Thus, expression of the newly introduced *dhfr* genes yields functional DHFR activity and complements the DHFR-negative phenotype of these cells.

Studies to ascertain the stability of the transformed *dhfr* marker and associated DNA segments in selective and nonselective culture conditions as well as characterizations of the copy number, physical state, and associations of the transduced DNA are still in progress.

DISCUSSION

Expression of *dhfr* in SVGT5 and SVGT7 vectors. The present work demonstrates that SV40 recombinant genomes containing a mouse *dhfr* cDNA segment in place of portions of the late region of SV40 can be propagated faithfully as virions with an appropriate helper and, moreover, that the exogenous genetic information is

FIG. 7. Immunological relatedness of DHFR made by SVGT5-*dhfr* and SVGT7-*dhfr* with mouse DHFR. Cell extracts from mock-, SV40-, SVGT5-*dhfr*-, and SVGT7-*dhfr*-infected CV1 cells were electrophoresed on sodium dodecyl sulfate-polyacrylamide gels containing a cleavable cross-linker, and the proteins were transferred from the gel to DBM paper (31). The DBM paper containing the putative DHFR was reacted with antibody (directed against mouse DHFR) and ^{125}I -labeled *Staphylococcus aureus* protein A (15) and then autoradiographed.

FIG. 8. Structures of the plasmid vectors containing the *Hind*III-*Bgl*II *dhfr* cDNA fragment. The *dhfr* cDNA is denoted by the hatched regions. Solid black segments are pBR322 sequences which contain the pBR322 origin of DNA replication and the ampicillinase gene. In the pSV1 vectors the lightly stippled areas are SV40 sequences (map positions 0.325 through 0.14) present in the pBR322-SV40 recombinant used to construct the pSV1 derivatives (see the text), and the clear regions indicate the SV40 DNA segments present in SVGT5-*dhfr* and SVGT7-*dhfr*. SV40 sequences in pSV2, pSV3, and pSV5 are shown as lightly stippled regions.

TABLE 1. Transformation of DHFR-negative CHO cells to a DHFR-positive phenotype with various vector *dhfr* DNAs

Recombinant	Transformation frequency ^a
pSV1GT5- <i>dhfr</i>	2×10^{-5}
pSV1GT7- <i>dhfr</i>	2×10^{-5}
pSV2- <i>dhfr</i>	1×10^{-4}
pSV3- <i>dhfr</i>	2×10^{-5}
pSV5- <i>dhfr</i>	2×10^{-4}

^a Transformation frequency is expressed as the fraction of cells plated in selective medium that give rise to a colony after transfection with saturating levels of DNA.

expressed as mRNA's and DHFR protein. As was the case with recombinant genomes containing a rabbit β -globin cDNA (25, 26) or an *E. coli* gene (Ecogpt or *gpt*) coding for xanthine-guanine phosphoribosyl transferase (23, 24), the *dhfr* mRNA's and DHFR protein are produced at rates comparable to those of the corresponding SV40 gene products. This permits a comparison of several aspects of the expression of the three exogenous DNA segments placed at iden-

tical locations in the SV40 DNA. The major hybrid mRNA's produced by SVGT5 and SVGT7 recombinants with β -globin-, *gpt*-, or *dhfr*-coding segments contain spliced 5' leaders characteristic of SV40 16S or 19S late mRNA's (or both); hence, we surmise that transcription initiates at the SV40 late promoter. The SVGT5-*dhfr* recombinant, which retains the SV40 sequences involved in late mRNA splicing (7) and polyadenylation (6, 30), generates the expected structural analogs of the SV40 16S and 19S late mRNA's. Similarly, SVGT7-*dhfr*, which lacks the splice junction for 16S mRNA formation, but retains the 19S mRNA splice boundaries (Fig. 2), generates substantial quantities of the correctly spliced 19S *dhfr* mRNA analog; small amounts (<10%) of other mRNA's were also detected, but their structures were not analyzed. The results summarized above contrast with findings made with comparable recombinants in which the β -globin cDNA (25, 26) or *gpt* segments (23, 24) were inserted into SVGT5 and SVGT7 vectors. In SVGT5- β -globin the expected 16S and 19S analog mRNA's were predominant (>95% of β -globin mRNA), but

SVGT7- β -globin yielded, in addition to the expected 19S type β -globin mRNA, aberrantly spliced species. These RNAs contained a SV40 late leader segment spliced to nucleotides in the 5' untranslated portion of the β -globin cDNA sequence (J. Sklar and P. Berg, unpublished results). Aberrant splicing has also been noted with SVGT5-*gpt* and SVGT7-*gpt* recombinants. Each of these recombinant genomes generates, in addition to the expected 16S and 19S *gpt* mRNA's, novel mRNA species in which an SV40 leader segment is spliced to nucleotide(s) within the *gpt*-coding region (26). Thus, whereas aberrant splicing from SV40 leaders to "foreign" DNA sequences occurred with β -globin and *gpt* segments, no comparable missplicing with *dhfr* sequences was noted.

There are several notable features regarding the formation of DHFR protein in cells infected with SVGT5-*dhfr* or SVGT7-*dhfr*. Figure 6 shows that DHFR, like the SV40 virion protein VP1, is detectable in infected cell extracts by protein staining of the electrophoresed gels. This suggests that there is substantial expression of the *dhfr* segment, perhaps as efficient as the viral segment it replaces. Moreover, DHFR protein appears to be more stable in the infected monkey cells than is β -globin; the latter, although synthesized at rates comparable to VP1, does not accumulate and was detected only by pulse-labeling because the half-life is less than 30 min (25).

A surprising aspect of *dhfr* expression concerns the production of DHFR protein in SVGT7-*dhfr*-infected cells. As mentioned above, the 19S *dhfr* mRNA was the predominant species produced in the infection. This 19S mRNA contains an SV40 leader segment spliced to a nucleotide just upstream of the initiator AUG for VP2 (7). Consequently, the 19S *dhfr* mRNA analog has two AUG codons upstream of the *dhfr* initiator AUG: one coding for initiation of VP2 and the other coding for VP3 (6, 30). If the large quantity of DHFR is being translated from the major 19S spliced mRNA, it appears that translation of the *dhfr* coding sequence occurs from an "internal" AUG codon. A similar finding was made by Mulligan and Berg (24), who observed that in the recombinant SVGT7-*gpt* the initiator AUG coding for xanthine-guanine phosphoribosyl transferase is the third AUG in the *gpt* segment and, therefore, the fifth one in the sequence downstream from the leader in the 19S mRNA produced by that recombinant. Equally perplexing translations from internal AUGs of pseudo dicistronic mRNA's (M. Fromm, P. J. Southern, and P. Berg, unpublished results) will be reported in forthcoming papers.

The problem of translation initiation in eucaryote mRNA's has been considered extensively by Kozak (16). To account for the supposed specificity of translation initiation at the first AUG downstream from the 5'-cap sequence (17), she proposed a "scanning" mechanism in which a ribosomal subunit binds at or near the 5' end of mRNA's, moves to the first AUG downstream from the cap structure, and initiates translation at that site. A possibility raised by the results with SVGT7-*dhfr* and with the *gpt* recombinants (24) is that ribosomes that can terminate translation from the first AUG can reinitiate translation at a nearby AUG; or ribosomes may, in certain circumstances, pass over the first or subsequent AUGs and initiate translation at a downstream AUG. It is interesting to note that the *dhfr* cDNA segment contains four terminator codons upstream of the initiator AUG codon, of which three are in the same reading frame, A, and the fourth is in another translation frame, B. An examination of the sequence at the SV40-*dhfr* cDNA join at map position 0.86 (Fig. 9) indicates that translation of VP2 or VP3 (or both) from their respective initiator codons on the 19S type *dhfr* mRNA would be terminated at the UGA in frame B of the *dhfr* cDNA. Whether terminator codons upstream of an internal AUG permit translation to initiate at those sites is presently under investigation. In this regard, we note that recombinants containing a β -globin cDNA segment that lacks terminator codons in the 5' untranslated region, within the VP1 coding sequence (26) or within the large T antigen-coding sequence (P. Southern and P. Berg, unpublished) of SV40 DNA, fail to produce β -globin protein, although they make the expected hybrid mRNA's.

Expression from plasmid vectors. The ability of pSV1GT5-*dhfr* and pSV1GT7-*dhfr* to

Fig. 9. mRNA sequence corresponding to the 5' noncoding region of the *dhfr* cDNA in SVGT7-*dhfr*. Shown is the base sequence across the SV40-*dhfr* cDNA joint and through the 5' end of the *dhfr* coding region. The AUG corresponding to the start of the DHFR protein and the four termination codons preceding it are indicated.

transform DHFR-negative CHO cells to DHFR positive (Table 1) was unanticipated since we expected that the *dhfr* segment would not be expressed from the late promoter in those semi-permissive cells. Quite possibly, the transforming activity of these recombinants is due to low-level expression of the late region; this could explain why the frequency of transformation is lower than with *dhfr* plasmids transcribed from the SV40 early promoter (see below); or integration of these recombinants into the cellular DNA could occur in such a way as to permit or activate expression of the late region. It is also conceivable that SV40 late region expression is ordinarily lethal, and that by substituting the *dhfr* segment for normal SV40 late genes, expression of the late region is innocuous.

The transformation frequencies observed for pSV2-*dhfr*, pSV3-*dhfr*, and pSV5-*dhfr* DNAs are comparable to those obtained with the *gpt* gene in these vectors (24a). pSV3-*dhfr* transforms cells at a 5- to 10-fold-lower frequency than do the other two, possibly because the SV40 T antigen (35) made by pSV3-*dhfr* represses transcription of the *dhfr* segment from the SV40 early promoter.

ACKNOWLEDGMENTS

We thank Jeff Sklar for his help in determining the sequence shown in Fig. 9.

This research was funded by Public Health Service grant GM-13235 from the National Institute of General Medical Sciences, by Public Health Service grant CA 15513-07 from the National Cancer Institute, and by grant MV-35J from the American Cancer Society.

S.S. is a fellow of the Jane Coffin Childs Fund for Medical Research.

LITERATURE CITED

- Alwine, J. C., D. J. Kemp, and G. R. Stark. 1977. Methods for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. U.S.A.* 74:5350-5354.
- Berk, A. J., and P. Sharp. 1978. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 75:1274-1278.
- Blakeley, R. L. 1960. Crystalline dihydropteroylglutamic acid. *Nature (London)* 188:231-232.
- Chang, A. C. Y., J. H. Nunberg, R. J. Kaufman, H. A. Erlich, R. T. Schimke, and S. N. Cohen. 1978. Phenotypic expression in *E. coli* of a cDNA sequence coding for mouse dihydrofolate reductase. *Nature (London)* 275:617-624.
- Chasin, L., and G. Urlaub. 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U.S.A.* 77:4216-4220.
- Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Woorde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert. 1978. Complete nucleotide sequence of SV40 DNA. *Nature (London)* 273:113-119.
- Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman. 1978. Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. *J. Mol. Biol.* 126:813-846.
- Goff, S. P., and P. Berg. 1976. Construction of hybrid viruses containing SV40 and λ phage DNA segments and their propagation in cultured monkey cells. *Cell* 9:695-705.
- Goff, S. P., and P. Berg. 1979. Construction and expression of simian virus 40 recombinant genomes containing the *Escherichia coli* gene for thymidine kinase and a *Saccharomyces cerevisiae* gene for tyrosine transfer RNA. *J. Mol. Biol.* 133:359-383.
- Graham, R., and A. Van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52:456-467.
- Hamer, D. H., and P. Leder. 1979. SV40 recombinants carrying a functional RNA splice junction and polyadenylation site from the chromosomal mouse β -globin gene. *Cell* 17:737-747.
- Hamer, D. H., and P. Leder. 1979. Splicing and formation of stable RNA. *Cell* 18:1299-1302.
- Hamer, D. H., K. D. Smith, S. H. Boyer, and P. Leder. 1979. SV40 recombinants carrying rabbit β -globin gene coding sequences. *Cell* 17:725-735.
- Heyneker, H. L., J. Shine, H. M. Goodman, H. Boyer, J. Rosenberg, R. E. Dickerson, S. A. Narang, K. Itakura, S. Y. Lin, and A. D. Riggs. 1976. Synthetic lac operator DNA is functional *in vivo*. *Nature (London)* 263:748-752.
- Kessler, S. W. 1975. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody adsorbent: parameters of the interaction of antibody-antigen complexes with protein A. *J. Immunol.* 115:1617-1624.
- Kozak, M. 1978. How do ribosomes select initiation regions in messenger RNA? *Cell* 15:1109-1123.
- Kozak, M. 1980. Evaluation of the scanning model for initiation of protein synthesis in eukaryotes. *Cell* 22:7-8.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
- Lai, C. J., R. Dhar, and G. Khoury. 1978. Mapping the spliced and unspliced late lytic SV40 RNAs. *Cell* 14:971-982.
- Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193:265-275.
- Magnusson, G., and M. Nilsson. 1977. Multiple free viral DNA copies in polyoma virus transformed mouse cells surviving productive infection. *J. Virol.* 22:646-653.
- Mertz, J. E., and P. Berg. 1974. Defective simian virus 40 genomes: isolation and growth of individual clones. *Virology* 62:112-124.
- Mulligan, R. C., and P. Berg. 1980. Expression of a bacterial gene in mammalian cells. *Science* 209:1423-1427.
- Mulligan, R. C., and P. Berg. 1981. Factors governing the expression of a bacterial gene in mammalian cells. *Mol. Cell. Biol.* 1:449-459.
- 24a. 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.
- Mulligan, R. C., B. H. Howard, and P. Berg. 1979. Synthesis of rabbit β -globin in cultured monkey kidney cells following transfection with a SV40 β -globin recombinant genome. *Nature (London)* 277:108-114.
- Mulligan, R. C., R. T. White, and P. Berg. 1980. Formation of β -globin following infection with recombinants containing rabbit β -globin cDNA at different locations of SV40's late region. Miami Winter Symposium—Mobilization and reassembly of genetic infor-

- mation. 17:201-215.
27. Nakamura, H., and J. W. Littlefield. 1972. Purification, properties and synthesis of dihydrofolate reductase from wild-type and methotrexate resistant hamster cells. *J. Biol. Chem.* **247**:179-187.
 28. Nunberg, J. H., R. J. Kaufman, A. C. Y. Chang, S. N. Cohen, and R. T. Schimke. 1980. Structure and genome organization of the mouse dihydrofolate reductase gene. *Cell* **19**:355-364.
 29. Parker, B. A., and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infection by virus or virus DNA. *J. Virol.* **31**:360-369.
 30. 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 simian virus 40. *Science* **200**:494-502.
 31. Renart, J., J. Reiser, and G. R. Stark. 1979. Transfer of proteins from gels to diazobenzyloxymethyl paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proc. Natl. Acad. Sci. U.S.A.* **76**:3116-3120.
 32. 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.
 33. Roberts, D., and T. C. Hall. 1967. The reduction of folate and of dihydrofolate by homogenates of leukocytes from patients with leukemia or with myeloid metaplasia. *Cancer Res.* **27**:994-999.
 34. Rothenberg, S. P. 1966. A rapid radioassay for folic acid reductase and amethopterin. *Anal. Biochem.* **16**:176-179.
 35. Tegtmeyer, P., M. Schwartz, J. K. Collins, and K. Rundell. 1975. Regulation of tumor antigen synthesis by simian virus gene A. *J. Virol.* **16**:168-178.
 36. Vogelstein, B., and D. Gillespie. 1979. Preparation and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. U.S.A.* **76**:615-619.