Simian Virus 40-Rabbit β -Globin Recombinants Lacking Late mRNA Splice Sites Express Cytoplasmic RNAs with Altered Structures

R. TYLER WHITE,^{1†} PAUL BERG,¹ AND LUIS P. VILLARREAL²

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305,¹ and Department of Microbiology and Immunology, University of Colorado Health Science Center, Denver, Colorado 80262²

Received 8 June 1981/Accepted 23 November 1981

Deletions were introduced at exon-intron boundaries in the late region of a simian virus 40- β -globin cDNA recombinant to study the role of splicing in the formation of simian virus 40 late cytoplasmic RNAs. The recombinant was used as a wild type because it allowed characterization of mutant RNAs expressed from defective genomes in the presence of comparable RNAs contributed by the coinfecting helper virus. Removal of a 17-base pair segment at map position 0.76, which included a portion of the leader sequence implicated in the splicing of the major 16S mRNA, prevented expression of 16S-type mRNA. The same mutant accumulated cytoplasmic 19S-type mRNA, but the assortment of the 5' ends of these mRNAs differed from the assortment in the wild-type counterparts. Another mutant that lacks nucleotide sequences implicated in the splicing of the major 16S mRNA and one of the principal 19S-type RNAs accumulated a 16S-type mRNA with a previously undetected leader splice, an assortment of 19S mRNAs with new or normally underrepresented splices, and even a species of unspliced cytoplasmic 19S mRNA.

Simian virus 40 (SV40) late mRNAs are comprised of 5'-terminal leader segments spliced to protein-encoding body segments. The major 16S mRNA, which codes for virion structural protein VP1, consists of a leader segment transcribed clockwise from the L-strand between map coordinates 0.72 and 0.76 and a body segment from the region between map positions 0.94 and 0.17 (1, 6, 7, 9, 12, 13, 15, 17, 19, 21, 23). Late 19S mRNAs, which are believed to be translated into virion polypeptides VP2 and VP3 (28), consist of a body segment that is homologous to the DNA between map positions 0.765 and 0.17 spliced to a variety of leader segments derived from the region between map positions 0.67 and 0.76 (9, 19, 35). The pattern of leaders spliced to the 16S and 19S mRNA bodies is complex and could reflect one of the regulatory mechanisms that determine the relative abundances of 19S- and 16S-type mRNAs. Quite possibly, the splicing of different leader structures to the 19S mRNA body influences the relative translatabilities of the VP2 and VP3 coding sequences (9).

Some SV40 mutants that lack portions of the leader region are viable and produce mRNAs

with altered leader structures (31, 35). This suggests that the leader structures of the wild-type mRNAs are not essential for the formation or expression of spliced mRNAs; left open is the question of whether alterations in the leader structures affect the levels or proportions of the different late mRNAs. A deletion mutant of SV40 which lacks all of the splice boundaries that participate in the formation of late mRNAs (except the 16S mRNA 3' splice site) does not accumulate cytoplasmic late mRNA (20; R. T. White, Ph.D. thesis, Stanford University, Stanford, Calif., 1980), indicating that splicing of nuclear transcripts plays a role in the formation of stable SV40 cytoplasmic late mRNAs. Furthermore, the finding that mutants which lack the nucleotides that comprise the major 16S mRNA intervening sequence fail to produce mature 16S mRNA and VP1, even though transcription of the late region of this mutant would be expected to produce the mature 16S mRNA structure directly, also implicates proper splicing in SV40 mRNA biogenesis (11; B. H. Howard and P. Berg, unpublished data).

To analyze further how splicing influences late SV40 mRNA processing and to identify nucleotide sequences that are essential for RNA splicing, we examined structural features of SV40 cytoplasmic late mRNAs produced after

† Present address: Department of Biochemistry, University of California, Berkeley, CA 94720.

infection of monkey cells with an SV40- β -globin cDNA recombinant (26) and with variants of this recombinant lacking sequences previously implicated in late mRNA splicing. As expected, variants that lacked the sequences used to splice the major 16S mRNA leader (at map position 0.76) did not produce cytoplasmic 16S-type β -globin mRNA, but they did accumulate the 19S-type β -globin mRNA. Another mutant, which lacked the nucleotides that span the intervening and junction sequences of one of the major 19S mRNAs (the segment between map coordinates 0.76 and 0.765), had an unexpected phenotype; a 16S-type β -globin mRNA with a previously undetected leader splice was produced, which was translated to yield β -globin, and 19S-type β -globin mRNAs with leader segments having either new or scarce 5' termini and novel or previously underrepresented splice patterns accumulated in the cytoplasm.

MATERIALS AND METHODS

Cell culture and viruses. The history, growth, and protocols for virus infection of monkey kidney cell lines CV-1 and CV-1P have been described previously (25). The construction and characteristics of recombinant SVGT5- β G have also been described previously (26).

Bacterial strains and plasmids. A pBR322 recombinant containing rabbit β -globin cDNA (SVT5-BG) was provided by Richard Mulligan. Transformation of *Escherichia coli* K-12 strain HB101 with recombinants of pBR322 was performed as described previously (26).

Enzymes. Restriction endonucleases *Ava*II, *Hpa*I, *Hpa*II, *Bgl*I, and *Eco*RI were purchased from New England Biolabs, Bethesda Research Laboratories, or Boehringer Mannheim. Endonuclease *Taq*I and *E. coli* exonuclease VII were provided by Steven Goff, and λ 5'-exonuclease was provided by David Charney. S1 nuclease was purchased from Sigma Chemical Co., and T4 polynucleotide kinase was purchased from P-L Biochemicals.

Preparation of SVGT5dl1- β G, SVGT5dl2- β G, and SVGT5dl3- β G. SVGT5- β G circular DNA (5 μ g), prepared as described previously (14), was digested for 10 min at room temperature with 0.25 U of *Hpa*I endonuclease in 100 μ l of 10 mM Tris-hydrochloride (pH 7.4) containing 10 mM MgCl₂, 20 mM KCl, 1.0 mM dithiothreitol, and 100 μ g of bovine serum albumin per ml to yield SVGT5- β G linear DNA, which was purified by agarose gel electrophoresis. CV-1P cells were coinfectd with linear SVGT5- β G DNA and circular *tsA58* DNA, and viral DNAs were prepared from several independent plaques. Generally, the DNAs were mixtures of *tsA58* and unaltered SVGT5- β G DNA (probably the result of recircularization of the linear DNA). However, one plaque gave rise to a mixture of *tsA58* and an SVGT5- β G variant (SVGT5dl1- β G) which lacked the *Hpa*I restriction site at position 0.758.

SVGT5- β G linear DNA obtained after partial *Hpa*I endonuclease cleavage was digested at 37°C for 10 min

with 2 μ g of λ exonuclease per ml in a solution containing 70 mM potassium glycinate (pH 9.4), 2.5 mM MgCl₂, 0.1 mg of gelatin per ml, and 0.5 mM dithiothreitol; this treatment removed an average of 50 nucleotides from each end of the DNA. CV-1P cells were coinfectd with *tsA58* DNA and the exonuclease-treated SVGT5- β G DNA, and viral DNAs were obtained from individual plaques. One modified SVGT5- β G DNA, designated SVGT5dl2- β G, contained a deletion of the *Hpa*I restriction site at position 0.758.

A third variant of SVGT5- β G was obtained by using linear DNA produced by partial cleavage with *Ava*II endonuclease (5 μ g of SVGT5- β G DNA was incubated with 1 U of *Ava*II for 1 h at 25°C in 100 μ l of 6 mM Tris-hydrochloride [pH 8.0] containing 60 mM NaCl, 6 mM mercaptoethanol, 12 mM MgCl₂, and 100 μ g of bovine serum albumin per ml). Full-length linear SVGT5- β G DNA was isolated by agarose gel electrophoresis and was used to infect CV-1P cells along with *tsA58* DNA. The DNA isolated from one plaque lacked the *Ava*II site at map position 0.765; this variant was designated SVGT5dl3- β G.

Determination of nucleotide sequences. Limited chemical degradation of labeled DNA fragments (see below) was performed as described by Maxam and Gilbert (22), and the resulting reaction products were analyzed by electrophoresis in thin polyacrylamide gels (29).

Preparation of labeled protein extracts of infected cells. Confluent CV-1 monolayers were infected for 48 h with SVGT5- β G or its variants (multiplicity of infection, 50 PFU/cell) and were labeled with [³H]leucine and [³H]valine, and the proteins were extracted as previously described (26). β -Globin protein was immunoprecipitated with purified goat anti-rabbit β -globin (provided by S. Boyer, Johns Hopkins University) by a procedure in which *Staphylococcus aureus* cells (Calbiochem) were used as an adsorbent (27). Immunoprecipitated protein was subjected to electrophoresis in sodium dodecyl sulfate-polyacrylamide gels (18).

Extraction of viral RNA. Cytoplasmic RNA was isolated in two ways. In method 1 mainly cytoplasmic RNA was extracted because the nuclear RNA remained in unlysed nuclei. Semiconfluent cell monolayers were infected for 60 h with SVGT5- β G or its mutants (multiplicity of infection, 50 PFU/cell) and were washed twice with cold Tris-buffered saline and once briefly with a cold solution of 10 mM Tris-hydrochloride (pH 7.5) containing 10 mM NaCl and 5 mM MgCl₂. Then 1 ml of the latter buffer containing 0.5% Nonidet P-40 and 0.05% deoxycholate was added to each plate. The cells were harvested with a rubber policeman and mixed with an equal volume of 100 mM Tris-hydrochloride (pH 8.5) containing 300 mM NaCl and 15 mM EDTA. The cell aggregates were dispersed by forcing the mixture through a Pasteur pipette several times. The suspension was extracted twice with an equal volume of water-saturated phenol-chloroform (1:1), and the RNA was precipitated from the aqueous phase with 2 volumes of ethanol.

A second extraction procedure (method 2), which was designed to minimize contamination of cytoplasmic RNA with nuclear species, involved disruption of the infected cells in isotonic media and separation of the intact nuclei from cytoplasmic contents by centrifugation (34). About 48 h after infection, the semiconfluent monolayers were washed twice with cold Tris-

buffered saline and then once with a cold solution of 10 mM Tris-hydrochloride (pH 7.5) containing 0.15 M NaCl and 5 mM MgCl₂. After this, the cells were covered with 1 ml of the latter buffer containing 0.2% Nonidet P-40 and allowed to swell and lyse on ice for 5 min. The cells were harvested with a rubber policeman and dispersed with a Pasteur pipette. Nuclei were removed by centrifugation (3,000 rpm, 3 min). The cytoplasmic supernatant was decanted and centrifuged again to remove residual nuclei. The supernatant was mixed with an equal volume of 0.4 M NaCl–0.1 M Tris-hydrochloride (pH 8)–20 mM EDTA–1% sodium dodecyl sulfate and extracted twice with an equal volume of water-saturated phenol–chloroform (1:1), and the RNA was precipitated with 2 volumes of ethanol.

RNA isolated by either procedure was washed with 70% ethanol, dried slightly, and suspended in a small volume of 20 mM Tris-hydrochloride (pH 7.5) containing 10 mM NaCl and 10 mM MgCl₂. DNase I that was freed of RNase by treatment with iodoacetate (36) was added (20 µg of DNase per mg of nucleic acid), and the preparation was incubated at room temperature for 30 min. Then 0.5 volume of 80 mM Tris-hydrochloride (pH 8.0) containing 50 mM EDTA, 400 mM NaCl, and 0.5% sodium dodecyl sulfate was added, and the mixture was extracted with an equal volume of phenol chloroform, followed by precipitation with 2 volumes of ethanol.

Polyadenylated [poly(A)⁺] RNA was prepared by chromatography on oligodeoxythymidylic acid cellulose as described previously (3).

Electrophoresis of RNA in agarose gels. Poly(A)⁺ RNA isolated by method 1 (see above) was denatured with glyoxal in dimethyl sulfoxide and subjected to electrophoresis in a horizontal 1.5% agarose slab gel (10 µg of RNA per gel track) in 10 mM sodium phosphate buffer (pH 7.0) (24). After electrophoresis, the gel was immersed in 50 mM NaOH for 45 min to dissociate the RNA-glyoxal adduct. The gel was neutralized by two 20-min washes in 100 mM acetate buffer (pH 4.0), and the RNA was transferred to a diazobenzoyloxy cellulose sheet by blotting (2) and then hybridized for 48 h at 42°C with a nick-translated pBR322 recombinant containing β-globin cDNA. The cellulose sheet was washed for 2 h at 42°C with four changes of a solution containing 50% formamide and 5× SSC (1× SSC is 0.15M NaCl plus 0.015 M sodium citrate), followed by one wash at room temperature with Tris-buffered saline. The sheet was air dried and autoradiographed with Kodak XR-5 film at –70°C for 24 h with a Du Pont Cronex Lightning-Plus intensifying screen.

Preparation of DNA probes for structural analysis of viral RNA. DNA of SVGT5-βG or its mutants was cleaved with *TaqI* endonuclease or with *BglI* plus *EcoRI* (see below), and the 5' ends of the resulting fragments were labeled with [γ-³²P]ATP by an exchange reaction catalyzed by T4 polynucleotide kinase (5). The [γ-³²P]ATP used in the exchange reaction was prepared from carrier-free ³²PO₄ (New England Nuclear Corp.), as described previously (16). Labeled DNA fragments were purified by agarose gel electrophoresis.

Analysis of spliced structures of RNA. The following modification of the procedure of Berk and Sharp (4) was used to analyze the structure of spliced RNA. ³²P-

end-labeled DNA fragments (1 µg, 10⁶ cpm) were mixed with cytoplasmic RNA extracted by method 2 (see above) from 10⁷ cells (approximately 20 µg of total RNA) and ethanol precipitated. The pellets were dissolved in 40 µl of deionized formamide, heated to 65°C for 10 min, mixed with 10 µl of a solution containing 2.0 M NaCl, 0.2 M PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid)] buffer (pH 6.4), and 5 mM EDTA, and incubated at 49°C for 4 h. The samples were cooled on ice, diluted 10-fold with the appropriate buffer, and digested at 37°C for 45 min with either exonuclease VII or S1 nuclease (4). Samples were precipitated with ethanol and analyzed by electrophoresis in alkaline agarose gels (4). Dried gels were autoradiographed for 14 h as described above.

RESULTS

SV40 mutants lacking certain late region splice sites are defective and, therefore, require a helper virus for growth. To distinguish the mutant and helper RNA species, we used an SV40-β-globin recombinant genome which contains a rabbit β-globin cDNA segment in place of the VP1 coding sequences between map positions 0.945 and 0.145 (26) (Fig. 1) as the wild type. Although this recombinant genome is defective and must be propagated with *tsA58* helper virus, its β-globin mRNA is easily distinguished from the helper RNAs because the late viral 16S- and 19S-type mRNA species transcribed from the SVGT5-βG recombinant contain β-globin coding sequences instead of VP1 coding sequences (Fig. 1).

Nucleotide sequence analysis of mutants. SVGT5dl1-βG and SVGT5dl2-βG DNAs were cleaved with endonuclease *AvaII* and labeled with [γ-³²P]ATP by using T4 polynucleotide kinase. Digestion of the labeled fragments with endonuclease *HpaII* yielded one fragment from each mutant, extending from the *HpaII* site at position 0.725 to the labeled *AvaII* site at position 0.765 and spanning the region containing the SVGT5dl1-βG and SVGT5dl2-βG deletions. This fragment from map position 0.725 to map position 0.765 was subjected to chemical cleavage, followed by electrophoresis in thin polyacrylamide gels (Fig. 2). The nucleotide sequences read from the gels were compared with the previously published SV40 sequence (8, 32) to determine the endpoints of each deletion.

SVGT5dl1-βG lacks 17 nucleotides, including the splice site at map position 0.76 (Fig. 3). The deletion in SVGT5dl2-βG includes 17 nucleotides centered around the *HpaI* site at map position 0.758. The SVGT5dl2-βG genome contains an insertion of 4 deoxyribosyladenine-deoxyribosylthymine base pairs at the site of the deletion, since the mutant contains 11 deoxyribosyladenine-deoxyribosylthymine pairs between nucleotides 506 and 531 rather than the 7 deoxyribosyladenine-deoxyribosylthymine base pairs which should have been left by a 17-base

FIG. 1. SVGT5- β G and the structures of its late transcripts. (A) SV40 sequence between map positions 0.945 and 0.145 replaced with a partial cDNA copy of rabbit β -globin mRNA. The SV40 sequence retained in the recombinant is indicated by the thick black line. (B) Structures of the major cytoplasmic late RNAs produced by SVGT5- β G, which are analogous to those of wild-type SV40 mRNAs, containing both SV40 sequences (stippled areas) and β -globin sequences (thick black areas). The 5' ends of the mRNA leaders are left vague to indicate the variety of endpoints observed.

pair deletion of nucleotides 512 through 528. A sequence analysis of SVGT5dl3- β G (Villarreal, manuscript in preparation) revealed a 114-base pair deletion which includes the 5' splice site at map position 0.76 and the 19S 3' splice site at map position 0.765, as well as intervening sequences (Fig. 3).

β -Globin protein produced by SVGT5- β G mutants. Labeled extracts of cells infected with SVGT5- β G and each of the deletion mutants were immunoprecipitated with goat anti-rabbit β -globin and analyzed by electrophoresis in a sodium dodecyl sulfate-polyacrylamide gel (Fig. 4). Whereas SVGT5dl1- β G and SVGT5dl2- β G produced no detectable rabbit β -globin, SVGT5dl3- β G did express this protein at a low level compared with the SVGT5- β G parent.

β -Globin RNA produced by SVGT5- β G mutants. poly(A)⁺ cytoplasmic RNAs isolated from cells infected with SVGT5- β G and two of its variants were denatured with glyoxal and subjected to electrophoresis in an agarose gel. After transfer to diazotized cellulose paper, each RNA was hybridized with a ³²P-labeled, nick-translated plasmid containing the β -globin cDNA sequences (Fig. 5). SVGT5- β G infection yielded predominantly a 1.0-kilobase (kb) β -globin

FIG. 2. Polyacrylamide sequencing gels of end-labeled *HpaII-AvaII* fragments (map positions 0.725 to 0.765) from SVGT5dl1- β G (A) and SVGT5dl2- β G (B). Samples were subjected to electrophoresis at 1,400 V in an 8% polyacrylamide gel until the xylene cyanol marker dye had migrated to a position 4 cm from the top of the gel. The undried gels were exposed to Kodak XR-5 film at -70°C for 48 h with an intensifying screen. The numbers to the right of each gel refer to the nucleotide numbers (see legend to Fig. 3) of the bases at the boundaries of each deletion.

FIG. 3. Sequences of deletion mutants of SVGT5- β G. The sequence illustrated is part of the late sense strand, as determined by Fiers et al. (8) and Van Heuverswyn and Fiers (32). Numbering of the nucleotides begins at the middle of the pallindrome near the *Bgl*I cleavage site at map position 0.67.

FIG. 4. Electrophoresis of untreated and immuno-precipitated 3 H-labeled extracts from cells infected with SVGT5- β G and its mutants in a 15% sodium dodecyl sulfate-polyacrylamide gel. A protein extract from 2.5×10^6 cells infected with each virus was immunoprecipitated as described in the text, and the immune complexes were subjected to electrophoresis (B). One-tenth the amount of each cell extract was electrophoresed untreated (A). The gel was fluorographed by using EN 3 HANCE (New England Nuclear Corp.) for 1 week at -70°C with Kodak XR-5 film.

mRNA species and lesser amounts of a 1.8-kb RNA species. Since the β -globin cDNA segment in SVGT5- β G is 0.5 kb shorter than the SV40 segment which it replaces, the corresponding β -globin mRNA species should be 500 bases shorter than the corresponding SV40 mRNAs; hence, the β -globin RNAs correspond in size to the SV40 16S and 19S mRNAs.

SVGT5dl1- β G, which lacks the 16S 5' splice site at map position 0.76, produced no 16S-type β -globin mRNA, but we observed considerably higher levels of 19S-type RNA than with the SVGT5- β G parent. SVGT5dl2- β G infection yielded the same pattern of cytoplasmic β -globin RNAs as SVGT5dl1- β G infection (data not shown). However, SVGT5dl3- β G, which lacks the splice sites at map positions 0.76 and 0.765, produced β -globin RNA species that resembled the 16S and 19S types. The major β -globin mRNA produced by SVGT5dl3- β G appeared to be smaller and somewhat less abundant than the corresponding SVGT5- β G 16S-type mRNA; however, the larger, minor β -globin RNA species seemed to be about the same size and abundance as the 19S-type RNA species produced by the parent recombinant.

Structures of β -globin mRNAs produced by SVGT5- β G mutants. To analyze the spliced structures of the β -globin RNAs produced by SVGT5- β G and its deletion mutants, cytoplasmic RNA was annealed with two different hy-

FIG. 5. Agarose gel electrophoresis of β -globin RNAs expressed late after infection with SVGT5- β G and two variants. Cytoplasmic poly(A)⁺ RNA was denatured with glyoxal, subjected to electrophoresis in a 1.5% agarose gel, and transferred by blotting to diazotized cellulose. The cellulose sheet was incubated with 5×10^6 cpm of [³²P]DNA homologous to the rabbit β -globin sequence.

bridization probes. Both types of probe were labeled with ³²P at the 5' end of the L-strand; in one analysis the DNA probe obtained from each mutant extended from the *Bgl*I restriction site at map position 0.67 to the *Eco*RI cleavage site within the β -globin cDNA sequence. The labeled RNA-DNA hybrids formed after the poly(A)⁺ RNA from each infection was annealed with the appropriate labeled probe DNA were digested with S1 nuclease, and the digestion products were analyzed by alkaline agarose gel electrophoresis (Fig. 6). The second analysis was performed with labeled DNA fragments that extended from the *Taq*I restriction site at position 0.57 to the *Taq*I site within the β -globin cDNA sequence. The resulting RNA-DNA hybrids were digested with either S1 nuclease or exonuclease VII, and the digestion products were analyzed by alkaline agarose gel electrophoresis (Fig. 7). From the sizes of the fragments produced by S1 nuclease digestion of RNA-DNA hybrids, we mapped the 5' termini of the RNA body segments and, hence, one of the points of splicing of the RNA. Similar measure-

ments of the fragments formed by digestion of RNA-DNA hybrids with exonuclease VII revealed the positions of the 5' termini of the intact RNA species. Since with both types of probe the labeled end was within the β -globin sequence, any RNA-DNA hybrids formed between *tsA58* helper mRNA and the DNA probe would have left the labeled ends susceptible to S1 nuclease or exonuclease VII digestion. Consequently, the late transcripts produced by the *tsA58* helper virus were invisible and did not complicate the analysis.

S1 nuclease digestion of the RNA-DNA hybrids between SVGT5- β G RNA and the *Bgl*I-*Eco*RI DNA probe (Fig. 6) generally yielded fragments of 1.34, 0.43, and 0.37 kb. (The 1.23- and 0.315-kb fragments observed in this experiment resulted from hybridization of the helper 16S and 19S mRNAs with a contaminant in the probe, which originated from some helper DNA in the SVGT5- β G DNA.) The 1.34-kb fragment corresponded to an RNA with a continuous segment to map position 0.765 (i.e., the 19S mRNA 3' splice site) (Fig. 6B). The 0.43-kb fragment identified an RNA with a continuous sequence from the β -globin cDNA *Eco*RI site to the 16S mRNA 3' splice junction at position 0.94. The minor 0.37-kb fragment appeared to be derived from hybridization of the DNA probe with a 16S-like β -globin RNA whose leader segment was spliced at a site within the 5' untranslated region of the β -globin cDNA sequences (J. Sklar and P. Berg, unpublished data). The relative amounts of the spliced 19S- and 16S-type β -globin RNAs produced by SVGT5- β G infection were comparable to the amounts of the 19S and 16S mRNAs found by us and others in infections with wild-type SV40. A 16S-type RNA with an unusual splice is also produced by SVGT5- β G.

S1 nuclease digestion of RNA-DNA hybrids between SVGT5dl1- β G RNA and the DNA probe labeled at the *Taq*I site yielded one major fragment of 1.10 kb and two quite minor fragments of 0.62 and 0.55 kb (Fig. 7). The major fragment corresponded to an RNA that was continuous with the DNA to map coordinate 0.765; the minor fragments suggested discontinuities in minor RNAs at map positions 0.86 and 0.874. Exonuclease VII digestion of RNA-DNA hybrids from SVGT5dl1- β G yielded a major band of 1.5 kb, indicating the existence of an RNA with a 5' terminus at position 0.69. The two smaller fragments (0.62 and 0.55 kb) were also present, suggesting that the minor RNAs had their 5' ends at map positions 0.86 and 0.874, respectively.

Our data indicate that the major cytoplasmic RNA produced by SVGT5dl1- β G is a spliced 19S-type RNA species with a 5' leader beginning

FIG. 6. Alkaline agarose gel analysis of RNA-DNA hybrids with SVGT5-βG and mutants after digestion with S1 nuclease. (A) Samples were subjected to electrophoresis in a 1.0% alkaline agarose gel, which was dried and autoradiographed. The molecular length markers were obtained by digestion of ^{32}P -labeled SV40 DNA with *HincII* endonuclease. (B) Structures of DNA probes derived by digestion of SVGT5-βG, SVGT5dl1-βG, and SVGT5dl3-βG DNAs with *Bgl*I and *Eco*RI. The stippled bars define the β-globin sequence, which contained the labeled 3' end designated by the asterisk. The major S1-resistant DNA fragments observed are shown below the probes from which they were derived. The length determinations for DNA fragments in this gel system are accurate to ± 20 to 50 nucleotides. For this reason, the endpoints of fragments mapped by S1 nuclease or exonuclease VII analysis are accurate to ± 0.005 to 0.01 map units.

at map position 0.69 and a body whose 5' end maps at or near position 0.765. Two minor RNA species produced in infections with SVGT5dl1- β G have 5' ends within the late protein coding region and may be unspliced.

The same type of S1 nuclease analysis with the SVGT5dl2- β G transcripts revealed a major fragment of 1.10 kb and four minor fragments of 0.76, 0.69, 0.62, and 0.55 kb (Fig. 7). The major fragment represented a continuous RNA segment with a 5' terminus at position 0.765, whereas the minor species extended to map positions 0.834, 0.847, 0.86, and 0.874. Exonuclease VII digestion of these RNA-DNA hybrids yielded one major species of 1.5 kb, which corresponded to a transcript with a 5' terminus at position 0.69. This is consistent with an interpretation that the major transcript is spliced and that its 5' end is at map position 0.69 and its body segment begins at or near map position 0.765. Three additional exonuclease VII-derived fragments were present in lesser amounts; these were 1.28, 1.10, and 0.95 kb long, indicating that there were transcripts extending back to map positions 0.73, 0.765, and 0.8, respectively. Two additional rather faint fragments of 0.62 and 0.55 kb were also detected, and these corresponded to RNAs with 5' termini at positions 0.86 and 0.874, respectively. The minor "18S" RNA species produced by SVGT5dl2- β G are analogous to similar RNAs observed in the paranuclear fraction of cells infected with wild-type SV40 (33). The minor RNAs could represent degradation products or intermediates in the putative processing of 19S-type RNA to 16S RNA. From these data we conclude that as with SVGT5dl1- β G, the major cytoplasmic RNA synthesized by SVGT5dl2- β G is a spliced 19S-type RNA consisting of a leader with a 5' terminus at map coordinate 0.69 spliced from an as-yet-undetermined site to a body segment with its 5' end at or near map position 0.765. Minor transcripts with other 5' ends, both spliced and unspliced, also occur.

S1 nuclease digestion of RNA-DNA hybrids between SVGT5dl3- β G RNA and the *Bgl*I-*Eco*RI DNA probe yielded a poorly resolved set of fragments between 1.7 and 1.6 kb long, one 1.35-kb fragment, a predominant 0.43-kb fragment, and a lesser amount of a 0.37-kb fragment (Fig. 6). The 0.43-kb band corresponded to an RNA segment with a 5' terminus at map position 0.94, whereas the 0.37-kb fragment revealed a segment with a 5' endpoint within the untranslated region of the β -globin cDNA sequences. The existence of these two bands indicates that although the usual splicing from position 0.76 to form a 16S-type RNA species cannot occur, there must be some other type of leader splicing to yield the 16S-type β -globin RNA bodies. The

nature of this leader is unresolved at this time. The low amount of the 1.35-kb fragment indicates that there is an RNA with an endpoint at map position 0.75. When SVGT5dl3- β G RNA was hybridized to the *Taq*I-generated DNA probe, an S1 nuclease analysis revealed three major bands (Fig. 7). The 1.45-kb fragment and possibly the 1.20-kb fragment appeared to be the fragments that yielded the 1.6- to 1.7-kb bands with the other probe (Fig. 6); these would have corresponded to RNA species with homology to the probe up to map positions 0.68 and 0.728. The 1.1-kb band is derived from an RNA segment with a 5' end at position 0.75. Two smaller, less abundant species (0.62 and 0.52 kb) also showed up, suggesting bodies that end at map positions 0.86 and 0.874. No RNA body corresponding to the 16S-type β -globin mRNA would have been detected with this probe, as the expected 0.13-kb fragment would have run off the gel.

An exonuclease VII analysis of the RNA-DNA hybrids between SVGT5dl3- β G transcripts and the *Taq*I probe revealed one major fragment 1.45 kb long and a very minor fragment 1.2 kb long. Two smaller, less abundant species (0.62 and 0.52 kb) were also present. This means that the major 5' terminus for both 19S-type RNAs and the 16S-type RNA produced by SVGT5dl3- β G is at map position 0.68. These results indicate that one of the 19S-type RNAs produced by SVGT5dl3- β G is continuous from map position 0.68 to the 3' end of the probe. However, the more abundant species contain splices that yield bodies with 5' ends at or near map positions 0.728 and 0.75. The two smaller RNA species with 5' ends at map positions 0.86 and 0.874 also apparently are not spliced because the 5' end of the RNA corresponds to the 5' end of the bodies. The minor species with 5' endpoints at map position 0.728 could represent RNA containing no splices downstream from map position 0.728 or a minor component of spliced 19S RNAs with 5' termini at map position 0.728. Since map position 0.68 is the 5' end of the majority of the RNAs made by SVGT5dl3- β G, we surmise that the 5' end of the 16S-type RNA also maps at that position. This implies that in this mutant the 5' splice site for 16S-type RNA formation lies between map position 0.68 and the endpoint of the deletion at map position 0.745.

DISCUSSION

SVGT5- β G is a convenient model for studying the effect of deletions on the expression of viral RNAs because during infections with the recombinant and its helper, it is possible to monitor the structures of the mRNAs of the recombinant in the presence of the corresponding helper virus

FIG. 7. Alkaline agarose gel analysis of RNA-DNA hybrids with SVGT5- β G mutants after digestion with S1 nuclease or exonuclease VII (Exo VII). (A) Autoradiograph of a 1.0% alkaline agarose gel. (B) Structures of DNA probes derived by digestion of SVGT5- β G mutant DNAs with *TagI* and DNA fragments observed after S1 nuclease digestion of hybrids formed between mutant RNAs and these probes. (C) Exonuclease-derived fragments. For other details, see the legend to Fig. 6.

mRNAs. The one assumption that this experimental protocol makes is that transcription and processing of SVGT5- β G and SV40 late mRNAs are the same and that alterations in the structure of the leader region affect the expression of both genomes in the same way.

Figure 8 summarizes the relevant endpoints for the leader and body segments of the β -globin-containing late RNAs that are predicted on the basis of the known pattern of SV40 late mRNA splicing (6, 7, 9, 12, 13, 15, 17, 19, 21, 23, 35). The 5' ends of the 16S- and 19S-type β -globin RNA bodies determined by S1 nuclease analysis are the same as the 5' ends of 16S and 19S SV40 mRNAs. This technique does not permit identification of the 3' ends of the leaders. Mutants SVGT5dl1- β G and SVGT5dl2- β G both lack 17 base pairs that include the site for splicing the leader in the major 16S-type RNA, and as a consequence, each fails to make the 16S-type β -globin mRNA. Paradoxically, mutant SVGT5dl3- β G, which lacks 114 base pairs

that include the same leader splice site for 16S-type RNA formation, as well as the 3' splice site for 19S-type mRNA, produces substantial amounts of the 16S-type β -globin RNA and low levels of β -globin protein. The 16S-type β -globin mRNA produced by SVGT5dl3- β G has two notable features; first, the 5' end maps to position 0.69, a site not utilized for either SV40 major 16S mRNA or 16S-type β -globin mRNA of SVGT5- β G (9, 19, 35), and second, this unusual 16S-type RNA appears to be spliced from an as-yet-unidentified splice site to the splice junction at map position 0.94, which is normally used in SV40 16S mRNA formation. Conceivably, in SVGT5dl3- β G one of the already identified 19S leader splice sites (9, 19) is utilized to form 16S RNA. This suggests that the failure to produce 16S-type β -globin mRNA from mutants SVGT5dl1- β G and SVGT5dl2- β G may result from preferential splicing of the putative splice site to the 19S mRNA body, thereby preventing the formation of the unusual 16S-type β -globin

MAJOR TRANSCRIPTS

FIG. 8. Structures of major β -globin RNAs expressed by SVGT5- β G and its mutants. The 5' termini of leaders and body segments were mapped for each species. The 3' ends of the leader segments were not determined. SV40 wild type (wt*) and the SVGT5- β G wild type expressed mRNAs with analogous structures.

mRNA. It is also possible that because of the larger deletion in SVGT5dl3- β G, its primary RNA transcript has an altered secondary structure, thereby permitting splicing reactions with SVGT5dl3- β G RNAs but not with SVGT5dl1- β G or SVGT5dl2- β G RNAs.

Mutants SVGT5dl1- β G and SVGT5dl2- β G still retain nucleotide sequences that are known to participate in the splicing of 19S mRNAs. Hence, it is not surprising that each of these mutants makes 19S-type β -globin mRNAs. The spliced 19S RNAs expressed by these mutants contain 5' termini at or near map position 0.69, but no 5' termini at position 0.72, a major

terminus of cytoplasmic 19S mRNA produced by wild-type virus (9, 19, 35). Such an increase in the relative abundance of cytoplasmic RNAs with upstream termini has been noted previously with other mutants that have deletions in the late leader sequences (10, 35). Whether leader deletions cause increased transcription initiation at map position 0.69 or preferential processing and accumulation of transcripts with termini at map position 0.69 is not known.

In addition to the unexpected 16S-type β -globin mRNA, mutant SVGT5dl3- β G produces a population of novel 19S-type β -globin RNAs. Inasmuch as the extraction procedure used to

obtain these RNAs has been shown to exclude contaminating nuclear RNA (34), we surmise that these RNAs are cytoplasmic species. The Berk-Sharp analysis with exonuclease VII (4) suggests that these 19S-type β -globin RNAs have 5' ends that extend to map position 0.68, just on the late region side of the origin of DNA replication, but no 5' ends at the downstream termini found in wild-type RNAs (e.g., the major 19S 5' end is at map position 0.72). Therefore, it seems that the deletion in SVGT5dl3- β G, as well as other deletions that occur in the late leader region, alters the distribution of 5' ends of the stable 19S-type mRNAs. The two major 19S mRNAs produced by SVGT5dl3- β G have 5' splice junctions in the RNA bodies at map positions 0.728 and 0.75, locations that have not been implicated previously as splice junctions (9, 19). The least abundant of the 19S-type RNAs expressed by SVGT5dl3- β G appears not to be spliced since its sequence is contiguous with the DNA sequence up to map position 0.68, the same site as the 5' ends of the two spliced 19S-type mRNAs. This suggests that one of the 19S-type mRNAs is not spliced. Another possibility is that the RNA ending at map position 0.68 is spliced to a short segment that occurs very close to map position 0.68 or to a leader segment absent from the DNA probe. Unspliced 19S-type RNA has been noted previously in infections with wild-type SV40 (9), but subsequent studies (34) have indicated that this RNA is confined mainly to the nucleus. Therefore, mutant SVGTdl3- β G and some SV40 mutants with deletions in the late leader region (10) appear to accumulate unspliced 19S mRNAs in the cytoplasm.

The results of this study and previous experiments (35) emphasize the importance of splicing in the formation of normal quantities of SV40 late mRNAs. Prevention of normal splicing of the 16S-type mRNA results generally in the disappearance of that class of mRNA; however, under certain circumstances alternative splicing reactions result in the accumulation of reduced levels of structurally modified forms of 16S mRNA. Similarly, when the splice junctions used in producing the major species of mRNA are altered, alternative transcription initiations or splicing of 19S-type RNAs can occur. Further studies with modified genomes carrying specific structural modifications should be useful in determining the rules governing the sites of late transcription initiation, as well as the processing and accumulation of SV40 late mRNAs.

ACKNOWLEDGMENTS

This work was supported by Public Health Service research grants CA 25819 and CA 15513-07 from the National Cancer

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

LITERATURE CITED

1. Aloni, Y. 1975. Methylated SV40 mRNAs. *FEBS Lett.* 54:363-367.
2. Alwine, J. C., D. J. Kemp, B. A. Parker, J. Reiser, J. Renart, G. R. Stark, and G. M. Wahl. 1979. Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. *Methods Enzymol.* 68:220-242.
3. Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. U.S.A.* 69:1408-1412.
4. Berk, J. A., and P. A. Sharp. 1978. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 75:1274-1278.
5. Berkner, K. L., and W. R. Folk. 1977. Polynucleotide kinase exchange reaction. *J. Biol. Chem.* 252:3176-3184.
6. Celma, M., R. Dharr, J. Pan, and S. Weissman. 1977. Comparison of the nucleotide sequence of the messenger RNA for the major structural protein of SV40 with the DNA sequence encoding the amino acids of the protein. *Nucleic Acids Res.* 4:2549-2559.
7. Dhar, R., K. N. Subramanian, J. Pan, and S. M. Weissman. 1977. Nucleotide sequence of a fragment of SV40 DNA that contains the origin of DNA replication and specifies the 5' ends of "early" and "late" viral RNA. IV. Localization of the SV40 DNA complementary to the 5' ends of viral mRNA. *J. Biol. Chem.* 252:368-376.
8. Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Vourde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert. 1978. Complete nucleotide sequence of SV40 DNA. *Nature (London)* 273:113-120.
9. 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.
10. Ghosh, P. K., P. Roy, A. Barkan, J. E. Mertz, S. M. Weissman, and P. Lebowitz. 1981. Unspliced functional late mRNAs containing intervening sequences are produced by a late leader mutant of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 78:1386-1390.
11. Gruss, P., C. Lai, R. Dhar, and G. Khoury. 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 76:4317-4321.
12. Haegeman, G., and W. Fiers. 1978. Localization of the 5' terminus of late SV40 mRNA. *Nucleic Acids Res.* 5:2359-2371.
13. Haegeman, G., and W. Fiers. 1978. Evidence for 'splicing' of SV40 16S mRNA. *Nature (London)* 273:70-73.
14. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cultures. *J. Mol. Biol.* 26:365-369.
15. Hsu, M. T., and J. Ford. 1977. Sequence arrangement of the 5' ends of simian virus 40 16S and 19S mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* 74:4982-4985.
16. Johnson, R. A., and T. F. Walseth. 1979. The enzymatic preparation of [γ - 32 P]ATP, [γ - 32 P]GTP, [32 P]cAMP, and [32 P]cGMP, and their use in the assay of adenylate and guanylate cyclases and cyclic nucleotide phosphodiesterases. *Adv. Cyclic Nucleotide Res.* 10:135-167.
17. Khoury, G., B. J. Carter, F. Ferdinand, R. M. Howley, M. Brown, and M. Martin. 1976. Genome localization of simian virus 40 RNA species. *J. Virol.* 17:832-840.
18. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
19. Lai, C., R. Dhar, and G. Khoury. 1978. Mapping of spliced and unspliced late lytic SV40 RNAs. *Cell* 14:971-982.

20. Lai, C., and G. Khoury. 1979. Deletion mutants of simian virus 40 defective in biosynthesis of late viral mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **76**:71-75.
21. Lavi, S., and T. Groner. 1977. The 5'-leader sequence and coding region of late SV40 mRNAs are derived from non-contiguous segments of the viral genome. *Proc. Natl. Acad. Sci. U.S.A.* **74**:5323-5327.
22. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**:560-564.
23. May, E., H. Kopecka, and P. May. 1975. Mapping the transcription site of the SV40-specific late 16S mRNA. *Nucleic Acids Res.* **2**:1995-2005.
24. McMaster, G. K., and G. G. Carmichael. 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. U.S.A.* **74**:4835-4838.
25. Mertz, J., and P. Berg. 1974. Defective SV40 genomes. Isolation and growth of individual clones. *Virology* **62**:112-124.
26. Mulligan, R. C., B. H. Howard, and P. Berg. 1979. Synthesis of rabbit β -globin in cultured monkey kidney cells following infection with a SV40- β -globin recombinant genome. *Nature (London)* **277**:108-114.
27. Padgett, R. A., G. M. Wahl, P. F. Coleman, and G. R. Stark. 1979. N-(phosphonacetyl)-L-aspartate-resistant hamster cells overaccumulate a single mRNA coding for the multifunctional protein that catalyzes the first steps of UMP synthesis. *J. Biol. Chem.* **254**:974-980.
28. Prives, C. L., and H. Shure. 1979. Cell-free translation of simian virus 40 16S and 19S L-strand-specific mRNA classes to simian virus 40 major VP-1 and minor VP-2 and VP-3 capsid proteins. *J. Virol.* **29**:1204-1212.
29. Sanger, F., and A. R. Coulson. 1978. The use of thin acrylamide gels for DNA sequencing. *FEBS Lett.* **87**:107-110.
30. Shenk, T. E., C. Rhodes, P. Rigby, and P. Berg. 1975. Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in SV40. *Proc. Natl. Acad. Sci. U.S.A.* **72**:989-993.
31. Subramanian, K. N. 1979. Segments of simian virus 40 DNA spanning most of the leader sequence of the major late viral messenger RNA are dispensable. *Proc. Natl. Acad. Sci. U.S.A.* **76**:2556-2560.
32. 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.
33. Villarreal, L. P. 1981. A paranuclear extract contains a unique set of viral transcripts late in SV40 infection. *Virology* **113**:663-671.
34. Villarreal, L. P., and S. Jovanovich. 1980. Leakage of nuclear transcripts late in simian virus 40-infected CV-1 cells: quantitation of spliced and unspliced late 19S RNAs. *J. Virol.* **36**:595-600.
35. Villarreal, L. P., R. T. White, and P. Berg. 1979. Mutational alterations within the simian virus 40 leader segment generate altered 16S and 19S mRNA's. *J. Virol.* **29**:201-219.
36. Zimmerman, S. B., and G. Sandeen. 1966. The ribonuclease activity of crystallized pancreatic deoxyribonuclease. *Anal. Biochem.* **14**:269-277.