Regulation of SV40 Gene Expression through RNA Splicing

G. KHOURY, J. C. ALWINE, R. DHAR, P. GRUSS, C.-J. LAI, S. SEGAL, AND I. SEIF

Laboratory of Molecular Virology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Split genes and RNA splicing appear to be universal among eukaryotic cells and their viruses (see Gilbert 1978; Darnell 1978; Crick 1979). Although we know very little about the mechanism of splicing and its role in the regulation of gene expression, the eukaryotic viral systems such as SV40 or adenovirus are particularly useful for studying these problems. The advantages include (1) small, easily manipulated genomes for which many of the gene products are known and can be assayed; (2) an extensively studied genetic background with a number of well-characterized temperature-sensitive and deletion mutants; and (3) a large amount of primary sequence data on viral DNA and spliced viral RNAs (Reddy et al. 1978a,b; Fiers et al. 1978; Ghosh et al. 1978; Zain et al. 1979).

Recent studies on SV40 (Lai et al. 1978) and adenovirus type 2 (Ad2) (Darnell 1978) suggest that splicing is a late event in RNA processing, which occurs subsequent to polyadenylation. It appears that the primary nucleotide sequence at splice sites is crucial to the splicing reactions (Breathnach et al. 1978); however, there is little information regarding the nucleotides that constitute preferred sites for enzyme recognition.

We have been interested for some time in the signals that regulate the splicing events. As a general approach to this question, we have employed temperature-sensitive mutants and deletion mutants of SV40 which affect either early or late viral genes. A quantitative and qualitative analysis of the viral RNA synthesized in cells infected by these mutants has permitted us to draw a number of conclusions regarding the potential mechanism and function of splicing events. In addition to the analysis of the primary RNA sequence of preferred donor and acceptor splice sequences (Seif et al. 1979), we have obtained evidence that suggests that RNA conformation plays an essential role in splicing reactions. We also present the results of studies suggesting that splicing is a required function for the stabilization of a number of viral and, presumably, eukaryotic mRNAs.

MATERIALS AND METHODS

Viruses. The wild-type SV40 was strain 776. Early SV40 mutants included the viable 54/59 deletion (*dl*) mutants from the following sources. Deletion mutants 883, 884, 885, 886, 888, 890, and 891 were kindly provided by T. Shenk (University of Connecticut, Farmington), and deletion mutants 2000, 2001, 2005, and 2007 were the kind gift of M. Sleigh and W. C. Topp (Cold Spring Harbor Laboratory). Nonviable early mutants *dl*1001 and *dl*1002 were grown with the helper virus *tsB*4. The early SV40 temperature-sensitive mutant *tsA*58 was originally obtained from P. Tegtmeyer (State University of New York, Stony Brook).

Cell cultures. Primary African green monkey kidney (AGMK) cells were grown in minimum essential medium (MEM) supplemented with 10% fetal calf serum (FCS) (MEM-10). After two to three passages, confluent cell cultures were infected with SV40 (strain 776), or SV40 mutants, at a multiplicity of 1–10 pfu/cell and maintained in MEM-2 for 40–48 hours prior to harvesting viral RNA. In the case of early 54/59 mutants, cells were infected in the presence of cytosine arabinoside (50 μg/ml) and RNA was harvested after 24 hours of infection.

The F9 line of embryonal carcinoma cells was kindly provided by A. Levine (State University of New York, Stony Brook). Its maintenance and growth properties have been previously described (Segal et al. 1979). Differentiation of this cell line in the presence of retinoic acid was done according to the method of Strickland and Mahdavi (1978). PCC3 cells were kindly provided by F. Jacob (Institute Pasteur, Paris).

Preparation of cytoplasmic and nuclear RNAs. Cytoplasmic RNA was prepared from infected cells late after SV40 infection according to the procedures previously described (Khoury et al. 1979). Poly(A)-containing RNA was separated from non-poly(A)-containing RNA by chromatography on an oligo(dT)-cellulose column (type 7; P-L Biochemicals).

Preparation of ^{32}P-labeled SV40 DNA restriction enzyme fragments and the separation of the SV40 DNA strands. SV40-infected cells were washed three times with phosphate-free medium prior to labeling with 0.1 mCi/ml of [^{32}P]orthophosphate (Amersham-Searle) in phosphate-free MEM-2, 24 hours after infection. After a labeling period of 48 hours, viral DNA was prepared according to the procedure of Hirt (1968). Form-I SV40 DNA with a specific activity of 1×10^6 to 2×10^6 cpm/μg was obtained by CsCl–ethidium-

bromide equilibrium sedimentation (see Lai and Nathans 1974b). Conditions for digestion of SV40 DNA with various restriction endonucleases (obtained from new England BioLabs, unless otherwise specified) were those indicated by the company. Cleaved DNA fragments were separated by electrophoresis on 1.4% agarose gels in buffer containing 0.04 Tris (pH 7.8), 0.005 M NaAc, and 0.001 M EDTA. Strand separation of SV40 DNA fragments was achieved by agarose gel electrophoresis (Khoury and May 1977) according to a modification of the method of Hayward (1972).

Hybridization of RNA with duplex DNA. Conditions for hybridization of viral RNA to double-stranded (DS) DNA were established at a temperature above the melting temperature (T_m) of the DNA duplex but below the T_m of RNA-DNA hybrids. We employed hybridization conditions of 0.4 M NaCl, 0.04 M PIPES (pH 6.5), and 0.001 M EDTA in 80% formamide at 48°C as described by Casey and Davidson (1977) and used by Berk and Sharp (1977). In brief, RNA was mixed with a 10–20-fold molar excess of the appropriate [^{32}P]DNA fragments produced by cleavage of SV40 DNA with various restriction enzymes; the mixtures were then precipitated in 70% alcohol. The precipitate was dried and then dissolved in 10–20 μl of formamide buffer. After dissociation of DNA duplexes at 68°C for 10 minutes, the hybridization mixture was incubated at 48°C for 2–3 hours. In the experiments where single-stranded (SS) DNA was used, hybridization was performed in 50% formamide, 0.4 M NaCl, and 0.04 M PIPES at 37°C for 20–40 hours.

Digestion of hybrid molecules with single-strand-specific nucleases. Conditions and steps involved in digestion of hybridization mixture with nuclease S1 were similar to those described by Berk and Sharp (1977). *Escherichia coli* strain M0676, used as a source of exonuclease VII, was the kind gift of S. Goff (Massachusetts Institute of Technology). Exonuclease VII was purified according to the method of Chase and Richardson (1974). For digestion, the reaction mixture was added to a 20× volume of buffer containing 0.01 M Tris (pH 8), 0.1 M NaCl, 0.01 M EDTA, 0.005 M β-mercaptoethanol, and an appropriate amount of exonuclease VII. After incubation at 37°C for 30 minutes, the reaction was stopped by adding a 2.5× volume of cold alcohol containing yeast tRNA as a carrier.

Gel electrophoresis. Labeled DNA segments obtained from digestion with nucleases were analyzed by electrophoresis on alkaline agarose gels (McDonell et al. 1977). In brief, the digest was precipitated in 70% alcohol by centrifugation in an Eppindorf PR/J centrifuge at 50,000*g* for 15 minutes. After dissolving in electrophoresis buffer (0.03 M NaOH, 0.001 M EDTA), the DNA samples were applied to 1.4% agarose in the electrophoresis buffer. Electrophoresis was conducted at 3 V/cm for 10–11 hours. The agarose gel was dried directly on DEAE paper and exposed to X-ray film using an intensifying screen (Dupont Cronex) for autoradiography.

Isolation and analysis of viral proteins. Confluent monolayers of secondary AGMK cells in 60-mm plastic petri dishes were inoculated with 20 to 50 pfu/cell of wild-type or mutant SV40. At 48 hours postinfection, cultures were starved for 1 hour in methionine-free medium labeled subsequently for 1 hour in methionine-free medium supplemented with 100 μCi/ml of [^{35}S]methionine. After methionine labeling, cellular proteins were extracted and immunoprecipitated with anti–SV40 T serum as described by Simmons and Martin (1978), subjected to 20% polyacrylamide gel electrophoresis in 15-cm slab gels (acrylamide : bisacrylamide = 60 : 1) for 8–10 hours at 75V (Tegtmeyer et al. 1977). The labeled protein bands were identified by fluorography at −70°C using Kodak SB-5 film that had been preflashed (Laskey and Mills 1975). The relative amounts of radioactivity in the protein bands were determined from densitometric tracings of autoradiographs of the gels.

RESULTS

Primary Nucleotide Sequence around Splice Sites

An analysis of the primary sequence of regions involving splice sites may provide an insight into the mechanism of RNA splicing. We have determined the DNA sequences that correspond to the early and late splice sites of human papovavirus BKV transcripts. A comparison of these sequences with those of other viral and eukaryotic mRNAs (for reference, see Seif et al. 1979) has indicated a number of similarities, which are discussed below. We refer to the RNA nucleotides (or their equivalents in the DNA of the genome) involved in the ligation process as donor (D) and acceptor (A) nucleotides. D and A bear a 5′→3′ orientation to each other in the spliced transcript. We refer to the regions that surround D and A in the genome or the primary transcript as donor site (DS) and acceptor site (AS).

Sequences Characteristic of Donor and Acceptor Sites

A list of the nucleotide sequences associated with various eukaryotic splice sites is presented in Figure 1. As the intervening sequences (IVS) vary considerably in size, they have been only partially represented in this figure, except for the smallest IVS of SV40 and BKV, which are only 31 nucleotides long and are presented in their entirety (marked by a central horizontal line). We have analyzed the 20–30 nucleotides at the donor and acceptor sites because deletion mutants of SV40 suggest that the nucleotides close to the excision points are sufficient for the splicing reactions (C.-J. Lai et al., in prep.). In the cases reported in Figure 1, the same sequences are present at both

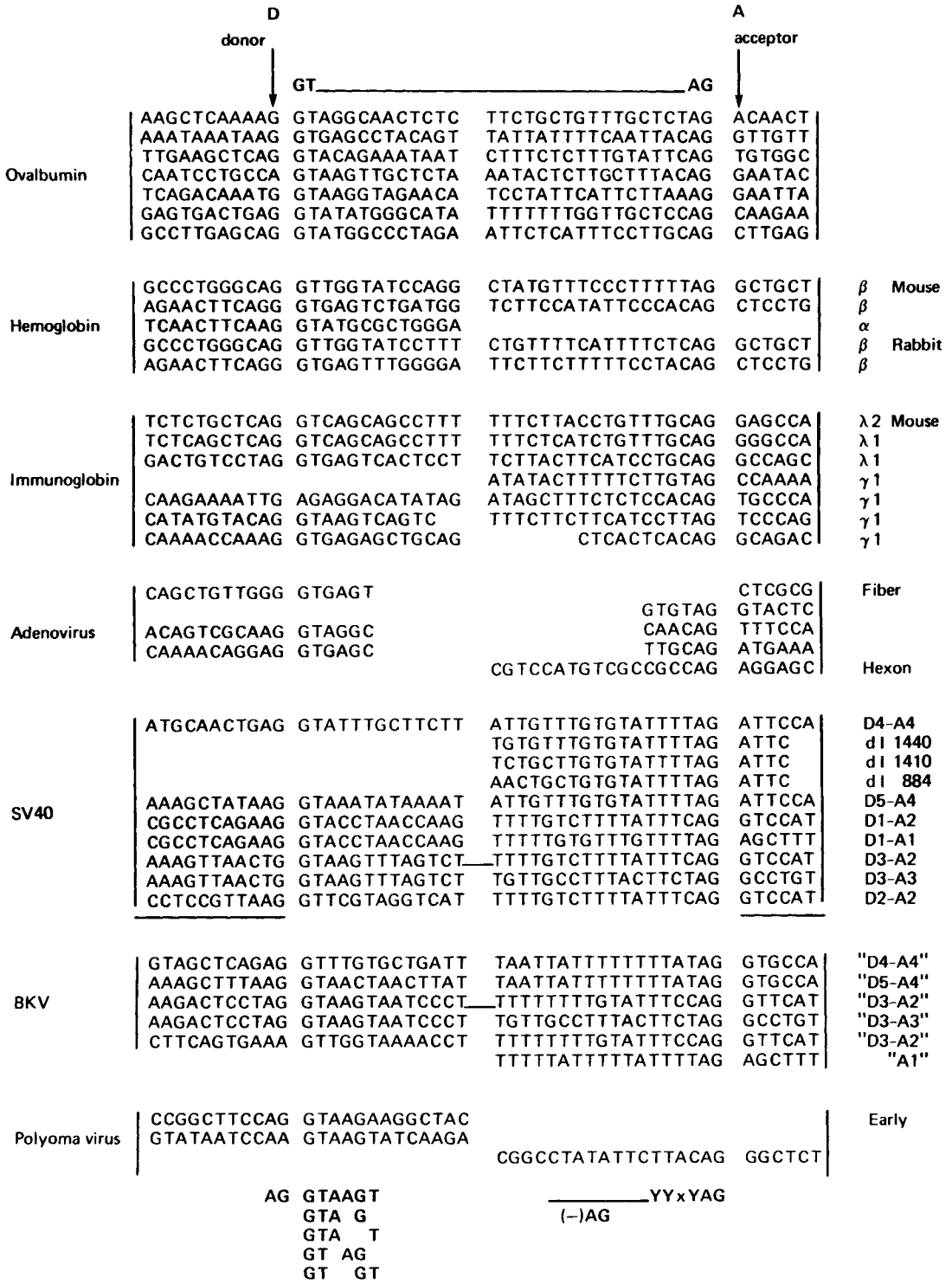

Figure 1. A comparison of eukaryotic and viral mRNA splice sites. The sequences surrounding the donor (D) and the acceptor (A) nucleotides are presented in their DNA form. The "GT-AG" rule of Breathnach et al. (1978) is illustrated at the top and our suggestions for preferred donor and acceptor sites, at the bottom (Y = pyrimidine). Because of their length or incomplete sequencing, most IVS are presented in part (illustrated by a central gap in the sequences); two IVS are presented in their entirety and the gap is filled by a horizontal bar (SV40 and BKV D3-A2 splices). The putative BKV and polyoma virus splice sites have been sequenced only at the DNA level. Sources for the splice site sequences are ovalbumin (Breathnach et al. 1978; Catterall et al. 1974; Gannon et al. 1979), hemoglobin (Efstradiatis et al. 1977; Konkel et al. 1978; Leder et al. 1978; van den Berg et al. 1978), immunoglobin (Tonegawa et al. 1978; Bernard et al. 1978; Sakano et al. 1979), adenovirus (Zain et al. 1979; Akusjärvi and Pettersson 1978), SV40 (Reddy et al. 1978a,b, 1979; Ghosh et al. 1978), BKV (Dhar et al. 1978; Yang and Wu 1979) (Fig. 2), and polyoma (Soeda and Griffin 1978; Hattori et al. 1979; Legon et al. 1978). Some sequences are in contradiction with others in the literature. In such cases, unless specified in the Discussion, our choice here has been arbitrary (see Seif et al. 1979).

donor and acceptor sites; these repeated sequences would allow splicing to occur in several different frames while still generating the same spliced product. For example, the spliced mouse β-globin RNA (Konkel et al. 1978) 5′-GC*AGG*C-3′ can be generated from the primary transcript

$$5'\text{-C}\overset{1}{\downarrow}A\overset{2}{\downarrow}G\overset{3}{\downarrow}G\overset{4}{\downarrow}\text{U}\ldots\text{U}\overset{1}{\downarrow}A\overset{2}{\downarrow}G\overset{3}{\downarrow}G\overset{4}{\downarrow}\text{C-}3'$$

by any one of the four pairs of endonucleolytic cleavage reactions designated by the pairs of arrows. Nevertheless, Breathnach et al. (1978), having compared a number of splice sites, propose that the IVS are bounded by GT at the 5′ end and by AG at the 3′ end. In other words, the donor nucleotide is followed by the dinucleotide sequence GT and the acceptor nucleotide is preceded by AG. The only exception to this rule at present is an IVS for a cloned mouse immunoglobulin heavy chain, which contains an AG instead of a GT at the 5′ end (Sakano et al. 1979). Proposed transcripts of BK virus and polyoma (Py) virus appear to conform to this rule (Fig. 1).

The nucleotide junctions formed after splicing include G-G, G-A, G-U, G-C, and A-G. In each case the donor nucleotide is a purine. It is furthermore a *G*, with the exceptions of an ovalbumin donor, a suggested donor of BKV and of polyoma virus.

From a more detailed comparison (Fig. 1), the *donor* site at the 5′ GT end of the IVS takes at least one of four forms: ↓GTAxG, ↓GTAxxT, ↓GTxAG, or ↓GTxxGT, with only one exception (↓GTACAG in ovalbumin). We refer to these sequences as the four groups of donor sites. These four groups derive from two basic forms: ↓GTxxG and ↓GTA.

Considering that the dinucleotide that precedes the 5′ end of the IVS is generally an AG (see Fig. 1), and with the four groups in mind, an optimal donor site might be AG↓GTAAGT. If the dinucleotide prior to the excision point (arrow) is not an AG, then there is a greater tendency to include a maximum of remaining sequence ↓GTAAGT; if the 5′ IVS deviates significantly from ↓GTAAGT, the donor AG is more likely to be present as in the ovalbumin splice sequence described above (AG↓GTACAG).

At the acceptor site, the 3′ end of the IVS is pyrimidine rich. With few exceptions it ends in PyPyx-PyAG↓. Within the 16 nucleotides preceding the 3′ dinucleotide AG↓, dipurines other than AA are rarely present. Furthermore, no AG appears within the 13 nucleotides preceding the 3′-terminal AG↓. The absence of dipurines (in particular AG) preceding the AG↓ at the 3′ end of the IVS may well be designed to minimize the likelihood of erroneous splices.

From the description of donor and acceptor sequences outlined above, some similarities of one to another are apparent. This is the case within the frequently observed junction repeat AG↓GT. Nevertheless, none of the acceptor sites, listed in Figure 1, can be categorized in any of the four groups of donor sites, and none of the donor sites listed has an AG preceded by a pyrimidine stretch, which is characteristic of an acceptor site. This suggests that the repeats may be present because of similarity in recognition of donors and acceptors by the splicing enzymes, but that there are additional nucleotide signals that establish a clear distinction between donor and acceptor. These signals are located primarily within the IVS. Consequently, they are removed during the splicing process. These preferred sequences for splicing are based on the limited number of splice sites presently available. It is clear that in the near future a large number of additional sequences will become available, and these will be used both to test and extend the "rules" presented above.

In the following section, we present evidence suggesting that, although preferred nucleotide sequences exist at which the ligation reaction can eventually join D to A at a splice junction, such a sequence itself is not enough to program splicing events. We belive that the results we have obtained make a case for the role of RNA conformation in determining the frequency with which a particular splicing event will occur.

Analysis of RNA from SV40 Deletion Mutants

In addition to donor and acceptor sequences, other segments of precursor RNA molecules, e.g. the IVS that are deleted by a splicing event and the flanking sequences that border the newly formed splice junction, may contribute to the splicing reaction. In one approach to this problem we have analyzed the RNAs from SV40 deletion mutants occurring in these regions. Alterations in the DNA structure at defined sites will be reflected by analogous deletions in the primary viral transcripts. The effect of specific changes in the splicing templates can then be assayed by quantitative and qualitative evaluation of stable, cytoplasmic, mutant RNA species. We have analyzed the mRNAs of both early and late SV40 deletion mutants. Since analogous results have, in general, been obtained with both sets of mutants, we will concentrate in this paper on mutants from the early region.

The early region of SV40 extends counterclockwise on the map of the genome from a position near the origin of viral DNA replication (approximately 0.67 map units) to the terminus of the major early transcripts at 0.15 map units (see Fig. 2). Given a knowledge of the entire SV40 nucleotide sequence (Reddy et al. 1978b; Fiers et al. 1978), a number of biological and biochemical observations are more easily interpreted. There are at least two early RNAs that appear to be processed from a single primary early transcript (Fig. 2); they differ in the size of their internal spliced out segment (Reddy et al. 1979; Berk and Sharp 1978) (Fig. 2). The smaller spliced transcript (T-mRNA), from which a 346-nucleotide IVS is removed, encodes the SV40 large T antigen (~90K protein) (Rundell et al. 1977; Prives et al. 1977). From the larger transcript (t-mRNA) a 65-nucleotide IVS is deleted, using the same distal splice junction at 0.53 map units (Reddy et

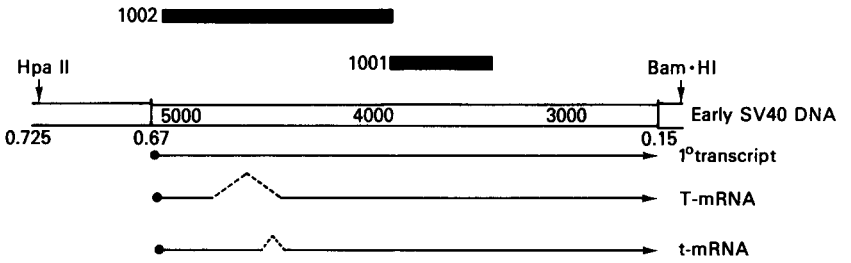

Figure 2. SV40 early transcription. A colinear primary transcript is synthesized from the early region of the SV40 genome (0.67–0.15 map units) (unpubl.). This transcript is spliced from approximately 0.60 to 0.53 map units for the large-T-antigen mRNA or from 0.54 to 0.53 map units for the small-T-antigen mRNA (Berk and Sharp 1978; Reddy et al. 1978a,b). Also presented are the positions of two early nonviable deletion mutants (*dl*1001 and *dl*1002) and the restriction enzyme cleavage sites used to produce the early region DNA probes (*Hpa*II = 0.725, *Bam*I = 0.145). The nucleotide numbers (3000, 4000, 5000) correspond to the approximate positions of SV40 nucleotides sequenced by Reddy et al. (1978a,b).

al. 1979). This t-mRNA is translated in the same reading frame and its product contains the same N-terminal amino acids and N-terminal [^{35}S]methionine tryptic peptides as does large T antigen (Paucha and Smith 1978; Simmons and Martin 1978; Crawford et al. 1978). However, near the end of the 5′ segment of this spliced molecule (~0.54 map units), prior to the splice junction, a termination codon is encountered that limits the polypeptide product (small T antigen) to 17,000 daltons (Prives et al. 1977; Crawford et al. 1978; Sleigh et al. 1978). Thus, the "body" (distal or 3′ portion) of t-mRNA represents a large untranslated segment, approximately 1970 nucleotides long.

We present data indicating that nonviable deletions well removed from splice sites generate stable truncated mRNAs, whereas those that delete splicing sequences do not. We also describe the analysis of early viral mRNA and viral polypeptides from a number of viable early SV40 deletion mutants that remove IVS and/or the small-T early mRNA splice junctions. The data suggest that deletions in the T-mRNA IVS as close as 29 nucleotides to the donor splice site or 11 nucleotides to the acceptor splice site lead to the production of normal amounts of mRNA and large T antigen. These same mutants represent deletions of nucleotides in the sequences that include and/or flank the 5′ end of the t-mRNA splice site. The deletion of sequences that flank the small T splice junctions, even when removed as far as 200 nucleotides from the donor site, can significantly influence the accumulation of a particular spliced molecule, presumably by determining the splicing frequency.

Early Deletion Mutants

We have examined the RNAs transcribed from two types of early SV40 deletion mutants. The first is represented by nonviable (defective) mutants *dl*1001 and *dl*1002 (Fig. 2). Mutant *dl*1001 contains a deletion located entirely within the coding sequence of large T antigen (0.42–0.325 map units); *dl*1002 (0.65–0.42 map units) is missing coding sequences, early splice junctions, and intervening sequences. The second category of early SV40 mutants, the early viable deletions or 54/59 mutants (Fig. 3), map between 0.535 and 0.60 SV40 map units (Shenk et al. 1976; Sleigh et al. 1978; Feunteun et al. 1978). These 54/59 deletions vary in size, but are located entirely within the intervening sequences that are spliced out of the T-mRNA. At the same time, they represent deletions within the coding sequences unique to the small T protein (Fig. 2). Since SV40 small T antigen appears to be dispensable, at least in lytic infections, this category of deletion mutants is viable and requires no helper virus. The location and the DNA sequence of a number of 54/59 deletion mutants are now known (Thimmappaya and Shenk 1979; Volckaert et al. 1979; I. Seif and R. Dhar, unpubl.) and are presented in Figure 3 and Table 1.

Mapping Cytoplasmic RNA from Defective Deletion Mutants

The construction and biological properties of SV40 mutant *dl*1001 (missing a DNA segment from 0.43 to 0.325 map units) have been previously described (Lai and Nathans 1974a; Rundell et al. 1977). *dl*1001 is defective in the early (gene-*A*) function and has been grown with a late temperature-sensitive helper virus, *tsB*4. Before analyzing the RNA synthesized by this mutant, we needed to determine the early transcriptional pattern of wild-type SV40 (for comparison with the mutant RNAs and because the wild-type early RNA is present in all cells infected with mutant virus populations containing a helper virus). The analysis of early SV40 RNA annealed with a wild-type SV40 SS DNA probe and subjected to nuclease-S1 digestion (see Materials and Methods) is presented in Figure 3a. As previously described by Berk and Sharp (1978), three major DNA segments (0.38, 0.125, and 0.07 map units in size) represent the two major early viral mRNAs, both of which are spliced (see Fig. 2). This 0.36-map-unit segment that migrates just under the 0.38-map-unit "body" is seen in preparations that contain late viral RNA and results from the overlap of the 3′ ends of early and late viral RNA molecules (see Reddy et al. 1978), leading to displacement of the 3′ ends of some early mRNA molecules from the hybridization probe (J. Alwine, unpubl.).

Table 1. Properties of Early Viable SV40 Mutants

Mutant strain	Size of deletion (nucleotides)	Position of deletion[a]		Type of deletion[b]	Size of small T antigen	
		proximal	distal		predicted[c]	observed[d]
*dl*883	58	4589(572)	4532(629)	splice junction	none	none
*dl*884	247	4748(413)	4502(659)	splice junction	none	12K
*dl*885	17	4667(494)	4651(510)	frameshift	14K	14K
*dl*886	10	4672(489)	4663(498)	frameshift	14.5K	15K
*dl*888	68	4611(550)	4544(617)	splice junction	none	none
*dl*890	27	4746(415)	4720(441)	in-frame	16K	16K
*dl*891	25	4809(352)	4785(376)	frameshift	11K	11K
*dl*2000	~ 10	—	—	—	—	16.5K
*dl*2001	~150	—	—	—	—	12K
*dl*2005	~300	—	—	—	none	none
*dl*2007	76	4663(498)	4588(573)	frameshift	15K	15K
Wild type	—	—	—	none	17K	17K

The nucleotide sequences of *dl*883, *dl*884, *dl*890, and *dl*891 have been determined by Thimmappaya and Shenk (1979) and Volckaert et al. (1978), and mutants *dl*885, *dl*886, *dl*888, and *dl*2007 by I. Seif and R. Dhar (unpubl.). The data have been previously published (Khoury et al. 1979).

[a]Inclusive sites of the deletions for which a sequence is known are given in the nucleotide sequencing numbers of both Reddy et al. (1978) and Fiers et al. (1978), in parentheses. The proximal site refers to the one closer to the 5′ end of RNA.

[b]Splice junction mutants remove the proximal small-T splice sequences (see Fig. 2). Those early viable deletion mutants that do not disturb this splice junction either preserve the small-T reading frame (in-frame) or alter it (frameshift).

[c]A size for small-T antigen could be predicted for those 54/59 deletion mutants for which the nucleotide sequence is known, considering whether the mutation was in phase or introduced a frameshift. Mutants that crossed splice junction (see Fig. 2) were predicted to make no small T antigen (see text for discussion of small T from *dl*884).

[d]Small T antigens based on autoradiographs from pulse-labeling experiments. Source of additional small T antigens seen at lower levels in cells infected with *dl*2000 (10K, 20K, 25K) and certain other mutants is under investigation.

Results of annealing RNA from *dl*1001-infected cells to the early strand of the wild-type SV40 *Hpa*II-*Bam*I A fragment (see Fig. 2), followed by nuclease-S1 digestion and alkaline agarose gel electrophoresis, are shown in Figure 3b. The 0.38-, 0.125-, and 0.07-map-unit-length bands, analogous to those seen in Figure 3a, represent the segments of early RNA molecules transcribed from the helper virus, SV40 *tsB*4. (The 0.07-map-unit-length fragment is partially obscured by the trailing of degraded oligonucleotides.) In addition, a new RNA segment of 0.17 map units is detected; further mapping studies using endo R · *Pst*I-cleaved SV40 DNA (data not presented) showed that it represents the portion of the body of mutant *dl*1001 RNA from the end of the deletion (0.33 SV40 map units) to the normal 3′ terminus for early RNA at 0.16 map units. The portion of *dl*1001 RNA between the common splice at 0.535 map units and the beginning of the deletion at 0.43 map units is 0.11 map units in length; it migrates just below the 0.125-map-unit leader and presumably is responsible for the broadness of this band (cf. Fig. 3a). Thus, infection with *dl*1001 leads to the production of a stable, shortened, polyadenylated early RNA. This observation supports the previous finding of a truncated T polypeptide in cells infected with *dl*1001 (Rundell et al. 1977).

Mutant *dl*1002 (Lai and Nathans 1974a) is missing essentially all of the *Hin*dII + III A fragment from 0.65 to 0.43 map units, a segment that includes the IVS and splice junctions for both large and small T antigens (see Fig. 2). In an experiment analogous to that described above for *dl*1001, monkey kidney cells were infected with *dl*1002 plus its helper virus *tsB*4, and the polyadenylated RNA was analyzed by annealing with the early *Hpa*II/*Bam*I wild-type SV40 DNA probe followed by S1 digestion and alkaline agarose gel electrophoresis. In repeated experiments (see Fig. 3c), we were able to detect only the helper virus RNA. If a stable early RNA were transcribed from *dl*1002, we would expect the S1 assay to demonstrate a fragment of approximately 0.27 map units in length (corresponding to the DNA segment mapping from the 3′ end of the early region at 0.16 map units). The absence of this 0.27-map-unit-length DNA suggests that no stable, early, SV40 cytoplasmic RNA is transcribed from mutant *dl*1002. This result is similar to our recent findings with a late deletion mutant (*dl*2301) from which the late SV40 splice junctions and intervening sequences (0.72–0.82 map units) have been deleted (Lai and Khoury 1979) (see Discussion).

Cytoplasmic RNA from Early Viable (54/59) SV40 Deletion Mutants

Cytoplasmic polyadenylated RNAs were purified from AGMK cells infected with various representatives of the early viable SV40 deletion mutants (Fig. 4; Table 1). These RNAs were annealed to the *Hpa*II/*Bam*I fragment A, early separated strand probes prepared separately from [^{32}P]DNA of each of the mutants; the products were first analyzed by treatment with exonuclease VII, which trims unpaired deoxyribonucleotides from the extremities of the hybrid molecules and thus positions the termini of RNA species, regardless of internal substitutions or

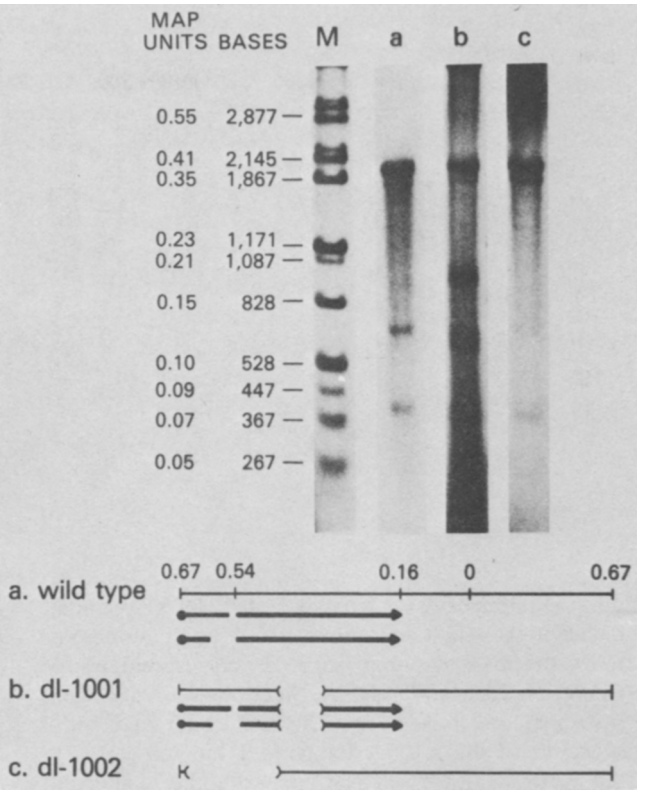

Figure 3. Mapping of early SV40 RNA transcribed from wild-type virus and the early nonviable deletion mutants. Cytoplasmic RNA was harvested 24 hr after infection of monkey kidney cells with SV40 or with either of the nonviable early deletion mutants *dl*1001 or *dl*1002 (plus its helper, *tsB*4). This RNA was annealed with an excess of ^{32}P-labeled probe as described in the text. RNA-DNA hybrids were treated with nuclease S1 and the products were analyzed by alkaline-agarose gel electrophoresis (see Materials and Methods). This size of marker (*M*) restriction enzyme fragments of SV40 DNA is given both in fractional lengths of the SV40 genome and in number of nucleotides. (*a*) Wild-type SV40 early RNA; (*b*) *dl*1001 plus *tsB*4 early RNA; (*c*) *dl*1002 plus *tsB*4 early RNA. A schematic diagram of the viral DNAs (thin lines) and RNAs (thick lines) is presented below the autoradiographs.

splices (Berk and Sharp 1978). Results not presented showed that DNA segments approximately 0.50 map units in length (depending on the size of the deletion) were protected from digestion. This indicates that the majority, if not all of the early mutant RNA molecules, have 5′ and 3′ ends analogous to those of the wild-type virus.

To identify the spliced RNA species of these early viable deletion mutants, similar annealing experiments were performed, and hybrid molecules were treated with endonuclease S1 and subjected to alkaline agarose gel electrophoresis. Results of this set of experiments are presented in Figure 5. In each case a prominent band of 0.38 SV40 fractional units was found, corresponding to the body of the early mRNAs (Fig. 2). The position of this body segment was confirmed in a separate experiment using an *Hpa*II/*Pst*I-cleaved SV40 DNA probe (the fragment map-

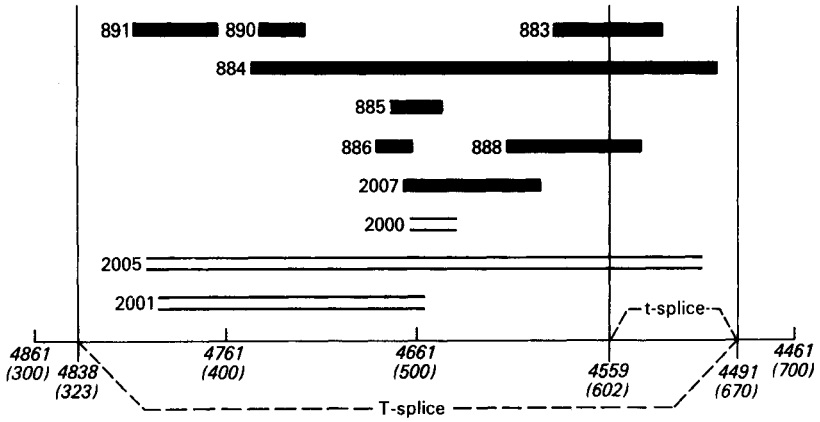

Figure 4. Location of the SV40 early viable deletion mutants. A diagram of the region of the SV40 genome between 0.60 and 0.53 map units is depicted (Fig. 1). The nucleotide sequence numbers of Fiers et al. (1978) are presented in parentheses (*300–700*) below those of Reddy et al. (1978b) (*4861–4461*). (-----) The sites of the splices for small- and large-T mRNAs; map positions are demarcated by vertical lines. Horizontal bars indicate the location of the early viable SV40 mutants. For some of the mutants, the nucleotide sequence of the deletions is known; these are represented by solid bars (▬); for the others, the approximate sizes and locations of the deletions based on cleavage patterns with appropriate restriction endonucleases are represented by open bars (═) (see Khoury et al. 1979).

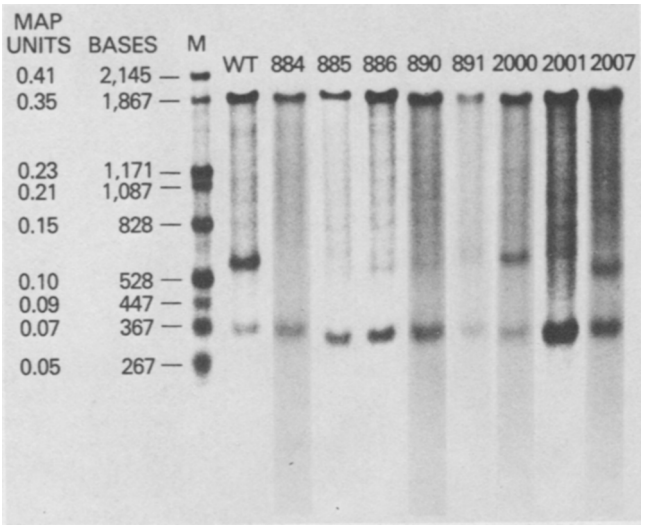

Figure 5. Mapping of early SV40 RNA transcribed from wild-type virus and early viable deletion mutants. Cytoplasmic RNA was harvested from AGMK cells 20 hr after infection in the presence of cytosine arabinoside by SV40 or by a number of early viable deletion mutants. Aliquots of the polyadenylated RNA fraction were annealed with an excess of the purified respective SS DNA probes, representing the early *Hpa*II-*Bam*I fragments of each virus. After annealing, DNA-RNA hybrids were analyzed by nuclease-S1 digestion and alkaline-agarose gel electrophoresis. The marker SV40 restriction enzyme fragments are the same as those of Fig. 3. The RNAs from early viable mutants *dl*883, *dl*888, and *dl*2005 were also analyzed, but they are not shown here (see Khoury et al. 1979).

ping counterclockwise from 0.72 to 0.28 map units; data not shown). Each mutant RNA population contains the 0.07-map-unit segment corresponding to the leader for the T-mRNA. Since the 0.07 and 0.38 fractional unit bands represent the leader and body segments of the mRNA for large T antigen (Fig. 2), the synthesis of this mRNA appears to be unaffected by the deletions in the 54/59 mutants (see below). It should be noted that none of the mutants for which we have a nucleotide sequence (*dl*883, *dl*884, *dl*890, *dl*891 [Thimmappaya and Shenk 1979; Volckaert et al. 1979]; *dl*885, *dl*886, *dl*888, and *dl*2007 [I. Seif and R. Dhar, unpubl.]) extends into the splice sites for large T antigen. Mutant *dl*891 maps 29 nucleotides from the large-T-antigen donor splice site and *dl*884 maps 11 nucleotides from the acceptor splice site. None of these deletions in the IVS region of the T-mRNA appear to have a significant effect, qualitatively or quantitatively, on the synthesis of T-mRNA.

Since all of these early viable mutants contain deletions that map in the segment unique to small T antigen (0.535–0.60 map units), it was expected that they would alter the size of the 0.125-map-unit leader RNA segment of t-mRNA (Fig. 5). It is remarkable, however, that most of these deletion mutants have a profound effect on the abundance of the t-mRNA as well as on the size of the t-mRNA leader. In the analysis of RNA from those mutants that delete the proximal splice junction for small T antigen (*dl*883, *dl*884, *dl*888, and presumably *dl*2005), there is no single discrete segment present in detectable amounts that would represent the shortened leader of an altered t-mRNA. This suggests that deletions that eliminate the proximal t-mRNA splice junction either eliminate or vastly diminish the production of an RNA for small T antigen (see May et al. 1978).

The deleted segments in early viable mutants *dl*2000 and *dl*2007 map within the T-mRNA IVS and do not involve any RNA splice junctions (Fig. 4). Nuclease-S1 analysis of the polyadenylated cytoplasmic RNA from AGMK cells infected by these mutants revealed three prominent bands, the body segment (0.38 map units), the T-mRNA leader (0.07 map units), and a shortened leader for t-mRNA (0.10 map units for *dl*2007 and 0.12 map units for *dl*2000). The decreased size of the leader in t-mRNA corresponds to the size of the wild-type t-mRNA leader minus the deletion. This suggests that the deleted nucleotides in *dl*2000 and *dl*2007 have little or no effect on the site or frequency of the formation of the t-mRNA splice.

The deleted regions of another group of early viable deletion mutants (including *dl*885, *dl*886, *dl*890, *dl*891, and *dl*2001), like *dl*2000 and *dl*2007, do not cross the t-mRNA splice junction (Fig. 4). Nuclease-S1 analysis of the early RNA of each of these mutants disclosed altered t-mRNA leaders (Fig. 5). The size of the shortened leader in most cases corresponded to that predicted from the size of the deletion. In contrast to the t-RNA leader segment generated by infection with wild-type SV40 or mutants *dl*2000 and *dl*2007, however, these mutants produced markedly reduced quantities of discrete t-mRNAs. This observation suggests that deletions that do not involve a splice junction can, nevertheless, significantly affect the amount of a stable, spliced RNA species that is generated.

Analysis of T Antigens Induced by the Early Viable Deletion Mutants

The analyses described so far show that the quantity of T-mRNA transcribed from the early viable deletion mutants is comparable to that transcribed from the wild-type virus, whereas the abundance of t-mRNA varies, depending on the mutant site. To correlate the levels of mRNA with the large and small T antigens produced by these mutants, we performed [^{35}S]methionine pulse-labeling of AGMK cultures infected with the deletion mutants or wild-type SV40, according to the protocol detailed in Materials and

Methods. Proteins were extracted and immunoprecipitated with an anti-SV40 small and large T serum and examined by SDS-polyacrylamide gel electrophoresis. The results of this analysis are shown in Figure 6 and Table 1. All the mutant-infected cells contained an abundant quantity of immunoprecipitated protein comigrating with the wild-type large T protein at 90,000 daltons. (Variable quantities of 88K and 85K T polypeptides, as seen in Fig. 6, have been described by others [Tegtmeyer et al. 1977]).

With one exception (*dl*884; see below), no small T protein was detected in cells infected by those mutants that cross the small-T-proximal splice junction (*dl*883, *dl*888, and *dl*2005). A small T protein, in amounts similar to that of the 17K small T polypeptide found in wild-type-infected cultures, was detected in cells inoculated with *dl*2000 and *dl*2007. As demonstrated above, these mutant-infected cultures also contained abundant t-mRNA. Deletion mutants 885, 886, 890, 891, and 2001 produced intermediate quantities of small T peptides, which appeared to correlate approximately to the amounts of truncated t-mRNAs that were detected in the preceding analysis (Fig. 5). In addition, the size of the shortened small T peptides directly corresponded to the size and position of the deletion in each mutant (see Table 1). For example, *dl*890 is an "in-frame" mutant with a deletion of 27 nucleotides (Thimmappaya and Shenk 1979; Volckaert et al. 1979); it shortens the small T antigen by the size of the deletion (from 17K to 16K). In contrast, *dl*891 is a "frameshift" deletion (Thimmappaya and Shenk 1979; Volckaert et al. 1979) that leads to a premature termination. The size of the deletion and the position of the new termination codon effectively remove 192 nucleotides of coding sequence from the t-mRNA. The new small T polypeptide has the predicted size (11K). The presence of a 12K small T polypeptide in cells infected by *dl*884 (Sleigh et al. 1978) (Fig. 6) is more difficult to explain. The size of this small T antigen indicates that it does not come from a contamination of the mutant stock by wild-type SV40; its detection in cells infected by each of three plaque isolates of the *dl*884 stock (unpubl.) makes contamination by another deletion mutant population seem unlikely. Yet the nucleotide sequence of *dl*884 (Thimmappaya and Shenk 1978; Volckaert et al. 1979) shows that the proximal t-mRNA splice junction is deleted. We have therefore considered the possibility that the small quantities of a 12K polypeptide in *dl*884-infected cells arise from transcripts that are unspliced and/or from transcripts that have a new splice (see Discussion).

Analysis of mRNA from Early SV40 Temperature-sensitive Mutants

In the preceding section we presented evidence that deletions remote from a potential splice site (i.e., in the 5′ flanking sequences for the t-mRNA splice) could nevertheless affect the abundance of the cytoplasmic t-mRNA, presumably by regulating the frequency with which a particular splicing reaction occurs. One interpretation of these data is that a critical factor in splicing is the conformation of the precursor RNA. This putative structure may depend in part on RNA-RNA, RNA-DNA, or RNA-protein interactions. Additional support for a splicing requirement greater than that of the primary nucleotide sequence at donor and acceptor sites comes from an analysis of the transcripts synthesized by early SV40 *tsA* mutants. At elevated temperatures (39–41°C) these mutants produce a defective large T antigen and overproduce early SV40 mRNA (Reed et al. 1976; Khoury and May 1977; Alwine et al. 1977). These observations plus the finding that T antigen normally binds to the origin of

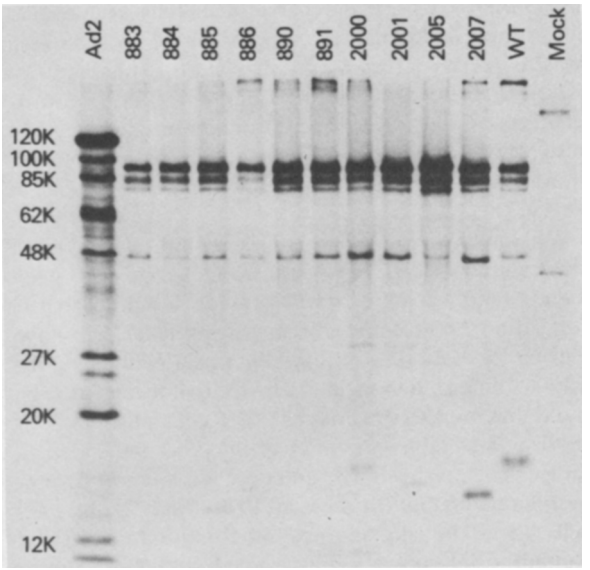

Figure 6. T antigens in cells infected by SV40 early viable deletion mutants. Confluent cultures of AGMK cells in 60-mm plates were infected by wild-type SV40 and several SV40 early viable deletion mutants (moi 20–50). At 40 hr postinfection, cells were starved for 1 hr in methionine-free medium and labeled for 1 hr in methionine-free medium with 100 μCi/ml [^{35}S]methionine (1295 Ci/mmole). Proteins were extracted as described previously (Simmons and Martin 1978), immunoprecipitated with hamster anti-SV40 T serum, and then electrohoresed in a 20% polyacrylamide gel at 75 V for 9 hr. The T-reactive proteins were analyzed by fluorography and their sizes estimated relative to Ad2 protein markers. (*Mock*) Uninfected cells. *dl*888 was also studied but is not shown here (see Khoury et al. 1979).

viral DNA replication (Reed et al. 1975; Alwine et al. 1975; Tenen et al. 1975; Tjian 1978), also the region of the proposed mRNA 5′ ends, led to the theory that T antigen acts as a repressorlike molecule. In this regard a substantial concentration of normal T antigen, in binding to SV40 DNA, is believed to reduce the synthesis of its own mRNA. Previous experiments demonstrating the overproduction of *tsA* early RNA were unable to distinguish whether this stimulation occurs at the level of transcriptional initiation, processing of the early mRNA, or both.

Since a single primary transcript of the SV40 early coding region appears to be the precursor for the mRNA encoding both small and large T antigens, depending on the location of the splice (Crawford et al. 1978; Berk and Sharp 1978), then the selection of the splicing event provides a potential regulatory step in early SV40 gene expression. Factors influencing the relative frequency at which either splice is used will control the relative abundance of the two early SV40 proteins, large and small T antigen.

Experiments were undertaken to determine if large T antigen regulates not only the transcription of the early region (the rate of synthesis of the primary early transcript), but also the relative amounts of large and small-T-antigen mRNAs. The data presented in Table 2 suggest that both early mRNAs are overproduced in the absence of a functional large T antigen. At elevated temperatures, however, the amount of large-T-antigen mRNA relative to small-T-antigen mRNA is enhanced both in cells infected by wild-type SV40 and the *tsA* mutant (J. Alwine and G. Khoury, in prep.). Thus, thermodynamic factors appear to influence the splicing of these RNAs. It should be noted in this set of experiments that the relative ratios of mRNA species are directly comparable to the ratios of large and small T antigens (Table 2). Thus, we have no evidence for further regulation of T antigens at a translational level. However, there is evidence for control of the abundance of a specific T antigen by selective splicing, which may be determined by the cell's thermodynamic environment.

Table 2. Ratio of the mRNAs or Proteins for Large and Small T Antigens in Cells Infected by Wild-type SV40 or *tsA*58 Mutants under Various Temperature Conditions

	AGMK cells			CV-1 cells		
	32°C[a]	shift[b]	41°C[c]	32°C[a]	shift[b]	41°C[c]
RNA						
Wild type	2.1	5.7	8.8	2.4	5.2	n.d.
*tsA*58	2.7	6.6	10.7	2.8	6.3	n.d.
Protein						
Wild type	2.48	4.61	5.50	1.8	3.5	4.9
*tsA*58	2.10	8.58	10.60	1.8	4.7	6.0

RNA: Early viral RNA species (cytoplasmic polyadenylated) were analyzed by the nuclease-S1 procedure of Berk and Sharp (1977). Ratios were determined from microdensitometer tracings of specific autoradiographic bands corresponding to large-T and small-T mRNAs. Protein: Pulse-labeled early viral proteins were quantitatively immunoprecipitated and electrophoresed in SDS-polyacrylamide gels. The ratios were determined by counting specific gel fragments containing the large and small T antigens.

[a]32°C samples: RNA was harvested or the proteins pulse-labeled 40 hr after infection at 32°C.

[b]Shifted samples: Cells were infected at 32°C for 40 hr, then shifted to 41°C for 5 hr, at which time the RNA was harvested or the proteins pulse-labeled.

[c]41°C samples: RNA was harvested or proteins pulse-labeled 24 hr after infection at 41°C.

Effect of the Host Cell on Processing of Viral mRNAs

In the absence of data on the splicing enzyme complex, it is difficult to know how many of these putative enzymes exist. It is also unclear whether they vary among species or among cell types, or if some enzymes are universally present in eukaryotic cells. Indirect support for the latter possibility, however, comes from a number of experiments: (1) BK virus, which grows in both AGMK cells and human embryonic kidney (HEK) cells induces the same spliced early and late viral mRNAs after infection of either cell type (Manaker et al. 1979). The same is true for adenovirus (T. Broker and L. Chow, pers. comm.). (2) SV40 or polyoma virus infection of a wide variety of cells generates the same spliced early viral mRNAs. (3) The mouse α- and β-globin RNA splice sequences are recognized in mouse cells into which SV40–mouse-globin recombinant DNA molecules have been inoculated (D. Hamer et al., pers. comm.). If there were a cell incapable of recognizing one or more sets of splice sites, it would not express a gene product from an mRNA requiring that splice. The fact that virtually all cells express SV40 T antigen indicates that all of them can appropriately splice the early SV40 mRNA. An exception is a series of mouse embryonal carcinoma cells that are unable to express SV40 T antigen. These cells, which are similar to cells of the early embryo, are the stem cells or precursors of teratocarcinomas. They are highly undifferentiated cells that rapidly cause tumors in susceptible hosts. Several groups were able to show that, although SV40 infection of embryonal carcinoma cells resulted in absorption and penetration of the virus, no T antigen was expressed (Swartzendruber and Lehman 1975; Kelly and Boccara 1976; Topp et al. 1977; Swartzendruber et al. 1977). These undifferentiated cells are also entirely refractory to infection with ecotropic murine type-C viruses (see Teich et al. 1977; Périés et al. 1977).

We have been involved in recent studies directed at determining the level of this block to viral gene expression; we hope it will provide insight into the regulation of host-cell genes. In particular, we have chosen to study the cell line F9, which is highly resistant to differentiation in vitro. Cells infected at high multiplicities (50–200 pfu/cell) of SV40 failed to synthesize detectable amounts of the SV40-encoded T antigens as assayed by immunofluorescence or immunoprecipitation (Segal et al. 1979) (Fig. 7). These F9 cells (even in the absence of SV40 infection) did contain a 55K protein that reacted with SV40 tumor

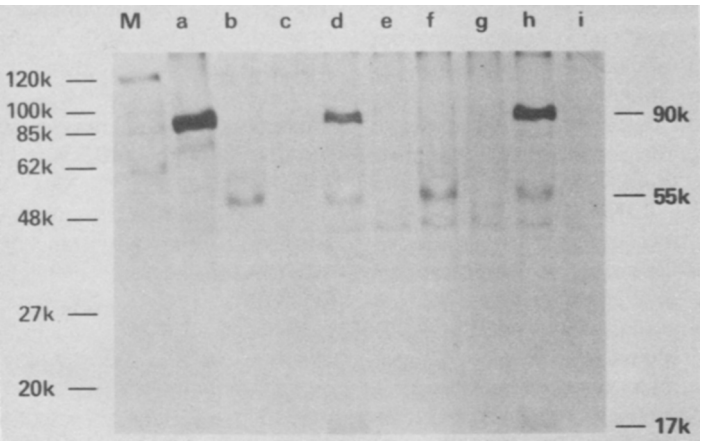

Figure 7. Analysis of proteins in SV40-infected undifferentiated and differentiated F9 cells. Undifferentiated and differentiated F9 cultures (1×10^7 cells/culture), uninfected and infected with SV40 (200 pfu/cell) for 18 hr, were labeled for 2 hr with [^{35}S]methionine (100 μCi/ml) in methionine-free medium. Proteins were extracted, purified, and immunoprecipitated with hamster anti-SV40 T serum as described by Segal et al. (1979). Immunoprecipitates were analyzed on 20% polyacrylamide slab gels (20% polyacrylamide–0.1% bisacrylamide, 0.2% SDS) for 10 hr at 75 V. The autoradiograph shows Ad2 marker proteins (*M*); control proteins from SV40-infected AGMK cells immunoprecipitated with anti-SV40 T serum (*a*); proteins extracted from SV40-infected undifferentiated F9 cells immunoprecipitated with either anti-SV40 T serum (*b*) or normal hamster serum (*c*); proteins from infected differentiated F9 cells immunoprecipitated with anti-SV40 T serum (*d*) or with normal hamster serum (*e*); proteins from uninfected differentiated F9 cells immunoprecipitated with anti-SV40 T serum (*f*) or with normal hamster serum (*g*). Protein extracts are from a spontaneously differentiating embryonal carcinoma line (PCC3) infected with SV40 and immunoprecipitated with anti-SV40 T serum (*h*) or with normal hamster serum (*i*). The bars at the right represent the positions of migration for large T antigen (90K), middle T antigen (55K), and small T antigen (17K).

sera (Fig. 7), which is thought to be a virus-induced protein of host origin (see Jay et al.; Linzer et al.; both this volume).

Since the block to early SV40 gene expression in embryonal carcinoma cells occurs at a step subsequent to adsorption and penetration, we turned to an examination of SV40 RNA in the infected cells using both C_rt analysis (Khoury et al. 1975) with SS [^{32}P]DNA probes and extension of nascent transcripts from viral transcriptional complexes (VTCs) (Ferdinand et al. 1977). Both methods indicated the presence of SV40 RNA in the infected cells; they also suggested that a portion of these RNA molecules had been initiated in the late as well as the early viral DNA strand (Segal et al. 1979). We then examined the RNA from infected F9 cells using the technique of Berk and Sharp (1978), which maps discrete transcripts. In contrast to the C_rt and VTC techniques, which measure only hybridizable RNA, no discrete spliced early SV40 transcripts were found. Yet at very low levels we did detect colinear viral RNA molecules, which appeared to be primary transcripts from the early region (Segal et al. 1979). From these data we concluded that the absence of SV40 T-antigen expression in teratocarcinoma cells resulted from an inability to generate stable early SV40 transcripts. This appeared to be in part related to the absence of splicing in early SV40 mRNA.

Recently, Strickland and Mahdavi (1978) have found that F9 cells can be induced to differentiate in tissue culture in the presence of vitamin-A derivatives, such as retinoic acid. This differentiation is accompanied by profound morphological changes as well as the acquisition of new antigenic determinants. For example, the SSEA-1 surface antigen disappears and the H-2 histocompatability antigen (H-2b) appears as the cells differentiate in tissue culture (Solter et al. 1979). When these differentiated cells are infected with SV40, one can easily demonstrate the synthesis of the two major spliced early RNAs and the presence of both large and small T antigens (S. Segal and G. Khoury, in prep.) (Fig. 7). Thus, the F9 cell line has a genetic composition that is capable of expressing the early SV40 genes. Yet, prior to differentiation this expression is blocked, at least in part, at a posttranscriptional level such that processed (spliced), mature viral transcripts do not accumulate, and T antigens are therefore not synthesized. Whether there is also an increased level of transcription of SV40 early RNA in the differentiated F9 cells is presently under investigation.

DISCUSSION

We have undertaken a number of experiments to determine what signals are involved in the splicing reaction for mRNAs and what possible significance splicing may have for gene regulation. For most of these studies we have used SV40, assuming that it will behave as the analog of a set of eukaryotic genes. In general, this would appear to be a sound assumption. A possible exception to the parallel between papovavirus and eukaryotic genes is that the viral genes are physically overlapping. Multiple splicing patterns within a

genetic unit (set of nucleotides), which allow for the expression of a group of nested polypeptides, have not yet been demonstrated for eukaryotic genes.

It is clear from an analysis of the nucleotide sequences at splice sites that preferred donors and acceptors exist in those regions destined to become mRNA splice junctions. In addition to the 5′ GT and 3′ AG at the termini of IVS (Breathnach et al. 1978), the extension of these rules outlined in Results has proved to be valuable in allowing us to position splices in mRNAs (e.g., BKV and polyoma virus mRNA) when we know their approximate location and the DNA sequence of their template (Seif et al. 1979).

Although there appears to be a nucleotide preference for the putative splicing enzyme, it seems justified at present to predict that RNA conformation also plays a central role in determining which splice junctions will be recognized in a specific splicing event. This prediction is based on a number of observations:

1. Deletion mutants in the early SV40 region flanking the small-T splice sites can significantly reduce the frequency at which this splice is made (Khoury et al. 1979) (see Table 1 and Figs. 4–6).
2. Similar effects have been observed with late SV40 deletion mutants, located within the 16S IVS (P. Gruss et al., in prep.).
3. The primary early SV40 transcript can be spliced in one of two ways, producing either small-T or large-T mRNA. The efficiency of this latter splice can be enhanced relative to the former if the incubation temperature is increased (Table 2). This suggests a differential thermodynamic property of the splicing intermediate (J. Alwine and G. Khoury, in prep.), which may indicate an alteration in the interaction between the template and the enzyme complex.

If the RNA conformation is important in regulating splicing efficiency, then the factors that determine the conformation remain to be elucidated. Even though secondary structure may be an inherent property of the RNA sequence itself, one must consider the possibility that the conformation of prespliced mRNAs is determined by their interaction with proteins and/or other nucleic acids.

Although the precise role for splicing of eukaryotic genes remains unclear, the functional significance of this processing event for the early and late SV40 genes is apparent. For the late SV40 transcripts, splicing enables the 5′ cap-containing leader sequence to move close to the coding functions for a specific mRNA (VP1 in 16S RNA and VP2 in 19S RNA). Since the cap has been shown to augment translation of eukaryotic transcripts and since the proximal AUG is generally recognized as the initiator codon, splicing permits translation to select a start signal that may have been internal in the primary transcript. In the early region of SV40, splicing of large-T-antigen mRNA removes a part of the small-T-antigen coding sequences as well as termination codons, thus permitting translation to continue in an open reading frame toward the 3′ end of the transcript. Since the small-T-antigen mRNA is spliced beyond the termination codon in its reading frame, the role of this splice is unclear. One possibility is that it prevents further processing of this mRNA to the large-T-antigen message (see Seif et al. 1979). Another hypothesis, not inconsistent with the first, is that many mRNAs require splicing for stabilization. A recent set of experiments bearing on this concept has involved an analysis of a class of late SV40 mutants, the so-called polar mutants (Cole et al. 1977), in which an upstream deletion in the late SV40 gene region prevents expression of the distal (VP1) function. Analysis of one of these mutants, *dl*2301 (Lai and Khoury 1979), has suggested that the defect resides in the inability to generate stable late mRNAs. In an extension of these experiments, Gruss et al. (1979) showed that a DNA that lacked precisely the late 16S RNA IVS could not express the 16S RNA gene function. This mutant, *dl*2350, which was generated from a reverse transcript of the SV40 late 16S mRNA, should be able to generate VP1 with the primary late transcript, bypassing the need for splicing. Its inability to do so has also been traced to a defect in stabilization of late cytoplasmic viral mRNA in the absence of splicing. In analogous studies, B. Howard and P. Berg (pers. comm.) have reached a similar conclusion.

With this result in mind, we are perhaps better able to approach the question of SV40 gene expression in the restrictive mouse teratocarcinoma cells. An analysis of the block to SV40 expression in the F9 teratocarcinoma cells was undertaken in anticipation that these cells would suppress the viral genes in the same way they prevent expression of the differentiated cellular functions. At least one problem uncovered was the inability of the undifferentiated cells to generate a stable spliced SV40 mRNA after viral infection. Of course, we cannot exclude the possibility that other regulatory mechanisms, such as deficient promoter activity, are also reponsible for the refractory nature of F9 cells to certain viruses. Nevertheless, retinoic acid treatment (Strickland and Mohdavi 1978) can induce differentiation in these cells to a state in which stable spliced SV40 early mRNAs are synthesized and T antigens are made (S. Segal and G. Khoury, in prep.). This allows us to exclude the possibility that the F9 line represents regressive "dead-end" cells that have lost the ability to express genes like SV40. Future experiments will be directed at further elucidation of the molecular mechanism of gene expression in embryonal carcinoma cells as they undergo changes associated with differentiation.

Acknowledgments

We thank N. Goldman, G. Jay, R. Manaker, E. May, and C. Prives for helpful suggestions during the course of this work.

REFERENCES

AKUSJÄRVI, G. and U. PETTERSSON. 1978. Nucleotide sequence of the junction between the coding region of the adenovirus 2 hexon messenger RNA and its leader sequence. *Proc. Natl. Acad. Sci.* **75:** 5822.

ALWINE, J., S. I. REED, and G. R. STARK, 1977. Characterization of the autoregulation of simian virus 40 gene A. *J. Virol.* **24:** 22.

ALWINE, J., S. I. REED, J. FERGUSON, and G. R. STARK. 1975. Properties of T antigens induced by wild-type SV40 and tsA mutants in lytic infection. *Cell* **6:** 529.

BERK, A. J. and P. A. SHARP. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12:** 721.

———. 1978. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci.* **75:** 1274.

BERNARD, O., N. HOZUMI, and S. TONEGAWA. 1978. Sequences of mouse immunoglobin light chain genes before and after somatic changes. *Cell* **15:** 1133.

BREATHNACH, R., C. BENOIST, K. O'HARE, F. GANNON, and P. CHAMBON. 1978. Ovalbumin gene: Evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl. Acad. Sci.* **75:** 4853.

CASEY, J. and N. DAVIDSON. 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4:** 1539.

CATTERALL, J. F., B. W. O'MALLEY, M. A. ROBERTSON, R. STADEN, Y. TANAKA, and G. G. BROWNLEE. 1978. Nucleotide sequence homology at 12 intron-exon junctions in the chick ovalbumin gene. *Nature* **275:** 510.

CHASE, J. W. and C. C. RICHARDSON. 1974. Exonuclease VII of *Escherichia coli*. *J. Biol. Chem.* **249:** 4553.

COLE, C. N., T. LANDERS, S. P. GOFF, S. MARTEUIL-BRUTTAG, and P. BERG. 1977. Physical and genetic characterization of deletion mutants of simian virus 40 constructed *in vitro*. *J. Virol.* **24:** 277.

CRAWFORD, L. V., C. N. COLE, A. E. SMITH, E. PAUCHA, P. TEGTMEYER, K. RUNDELL, and P. BERG. 1978. Organization and expression of early genes of simian virus 40. *Proc. Natl. Acad. Sci.* **75:** 117.

CRICK, F. 1979. Split genes and RNA splicing. *Science* **204:** 264.

DARNELL, J. E., JR. 1978. Implications of RNA·RNA splicing in evolution of eukaryotic cells. *Science* **202:** 1257.

DHAR, R., C.-J. LAI, and G. KHOURY. 1978. Nucleotide sequence of the DNA replication origin for human papovavirus BKV: Sequence and structural homology with SV40. *Cell* **13:** 345.

EFSTRADIATIS, A., F. C. KAFATOS, and T. MANIATIS. 1977. The primary structure of rabbit β globin mRNA as determined from cloned DNA. *Cell* **10:** 571.

FERDINAND, F.-J., M. BROWN, and G. KHOURY. 1977. Synthesis and characterization of late lytic simian virus 40 RNA from transcriptional complexes. *Virology* **78:** 150.

FEUNTEUN, J., M. KRESS, M. GARDES, and R. MONIER. 1978. Viable deletion mutants in the simian virus 40 early region. *Proc. Natl. Acad. Sci.* **75:** 4455.

FIERS, W., R. CONTRERAS, G. HAEGEMAN, R. ROGIERS, A. VAN DE VOORDE, H. VAN HEUVERSWYN, J. V. HERREWEGHE, G. VOLCKAERT, and M. YSEBAERT. 1978. Complete nucleotide sequence of SV40 DNA. *Nature* **273:** 113.

GANNON, F., K. O'HARE, F. PERRIN, J. P. LE PENNEC, C. BENOIST, M. COCHET, R. BREATHNACH, A. ROYAL, A. GARAPIN, B. CAMI, and P. CHAMBON. 1979. Organisation and sequences at the 5′ end of a cloned complete ovalbumin gene. *Nature* **278:** 428.

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 SV40. *J. Mol. Biol.* **126:** 813.

GILBERT, W. 1978. Why genes in pieces? *Nature* **271:** 501.

GRUSS, P., C.-J. LAI, R. DHAR, and G. KHOURY. 1979. Splicing as a requirement for biogenesis of functional SV40 mRNA. *Proc. Natl. Acad. Sci.* **76:** 4317.

HATTORI, J., G. G. CARMICHAEL, and T. L. BENJAMIN. 1979. DNA sequence alterations in Hr-t deletion mutants of polyoma virus. *Cell* **16:** 505.

HAYWARD, G. 1972. Gel electrophoretic separation of the complementary strands of bacteriophage DNA. *Virology* **49:** 342.

HIRT, B. 1968. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26:** 365.

KELLY, F. and M. BOCCARA. 1976. Susceptibility of teratocarcinoma cells to adenovirus 2. *Nature* **262:** 409.

KHOURY, G. and E. MAY. 1977. Regulation of early and late simian virus 40 transcription: Overproduction of early viral RNA in the absence of a functional T-antigen. *J. Virol.* **23:** 167.

KHOURY, G., P. GRUSS, R. DHAR, and C.-J. LAI. 1979. Processing and expression of SV40 mRNA: The effect of early deletions. *Cell* **18:** 85.

KHOURY, G., P. HOWLEY, M. A. BROWN, and M. A. MARTIN. 1975. The detection and quantitation of SV40 nucleic acid sequences using single-stranded SV40-specific RNA. *Cold Spring Harbor Symp. Quant. Biol.* **39:** 147.

KONKEL, D. A., S. M. TILGHMAN, and P. LEDER. 1978. The sequence of the chromosomal mouse β globin major gene: Homologies in capping, splicing and poly(A) sites. *Cell* **15:** 1125.

LAI, C.-J. and G. KHOURY. 1979. Deletion mutants of simian virus 40 defective in biosythesis of late viral mRNA. *Proc. Natl. Acad. Sci.* **76:** 71.

LAI, C.-J. and D. NATHANS. 1974a. Deletion mutants of simian virus 40 generated by enzymatic excision of DNA segments from the viral genome. *J. Mol. Biol.* **89:** 170.

———. 1974b. Mapping temperature-sensitive mutants of simian virus 40: Rescue of mutants by fragments of viral DNA. *Virology* **60:** 466.

LAI, C.-J., R. DHAR, and G. KHOURY. 1978. Mapping the spliced and unspliced late lytic SV40 RNAs. *Cell* **14:** 971.

LASKEY, R. A. and A. D. MILLS. 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* **56:** 335.

LEDER. A., H. I. MILLER, D. H. HAMER, J. G. SEIDMAN, B. NORMAN, M. SULLIVAN, and P. LEDER. 1978. Comparison of cloned mouse α and β globin genes: Conservation of intervening sequence locations and extragenic homology. *Proc. Natl. Acad. Sci.* **75:** 6187.

LEGON, S., A. J. FLAVELL, A. COWIE, and R. KAMEN. 1979. Amplification in the leader sequence of late polyoma virus mRNAs. *Cell* **16:** 373.

MANAKER, R., G. KHOURY, and C.-J. LAI. 1979. The spliced structure of BK virus mRNAs in lytically infected and transformed cells. *Virology* **97:** 112.

MAY, E., M. KRESS, and P. MAY. 1978. Characterization of two SV40 early mRNAs and evidence for a nuclear "prespliced" RNA species. *Nucleic Acids Res.* **5:** 3083.

MCDONELL, M. W., M. N. SIMON, and F. W. STUDIER. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* **110:** 119.

PAUCHA, E. and A. SMITH. 1978. The sequences between 0.59 and 0.54 map units on SV40 DNA code for the unique region of small t antigen. *Cell* **15:** 1011.

PÉRIÉS, J., E. ALVES-CARDOSO, M. CANIVET, M. C. DEBONS-GUILLEMIN, and J. LASNERET. 1977. Lack of multiplication on ecotropic murine C-type viruses in mouse teratocarcinoma primitive cells. *J. Nat. Cancer Inst.* **59:** 463.

PRIVES, C., E. GILBOA, M. REVEL, and E. WINOCOUR. 1977.

Cell-free translation of simian virus 40 early messenger RNA coding for viral T-antigen. *Proc. Natl. Acad. Sci.* **74:**457.

REDDY, V. B., P. K. GHOSH, P. LEBOWITZ, and S. M. WEISSMAN. 1978a. Gaps and duplicated sequences in the leaders of SV40 16S RNA. *Nucleic Acids Res.* **5:**4195.

REDDY, V. B., P. K. GHOSH, P. LEBOWITZ, M. PIATAK, and S. M. WEISSMAN. 1979. SV40 early mRNAs. I. Genomic localization 3′ and 5′ termini and two major splices in mRNA from transformed and lytically infected cells. *J. Virol.* **30:**279.

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. 1978b. The genome of simian virus 40. *Science* **200:**494.

REED, S. I., G. R. STARK, and J. C. ALWINE. 1976. Autoregulation of SV40 gene A by T-antigen. *Proc. Natl. Acad. Sci.* **73:**3083.

REED, S. I., J. FERGUSON, R. DAVIS, and G. R. STARK. 1975. T-antigen binds to simian virus 40 DNA at the origin of replication. *Proc. Natl. Acad. Sci.* **72:**1605.

RUNDELL, K., J. K. COLLINS, P. TEGTMEYER, H. L. OZER, C.-J. LAI, and D. NATHANS. 1977. Identification of simian virus 40 protein A. *J. Virol.* **21:**636.

SAKANO, H., J. H. ROGERS, K. HUPPI, C. BRACK, A. TRAUNECKER, R. MAKI, R. WALL, and S. TONEGAWA. 1979. Domains and the hinge region of an immunoglobin heavy chain are encoded in separate DNA segments. *Nature* **277:**627.

SEGAL, S., A. LEVINE, and G. KHOURY. 1979. Evidence for non-spliced SV40 RNA in undifferentiated murine teratocarcinoma cells. *Nature* **280:**335.

SEIF, I., G. KHOURY, and R. DHAR. 1979. BKV splice sequences based on analysis of preferred donor and acceptor sites. *Nucleic Acids Res.* **6:**3387.

SHENK, T., J. CARBON, and P. BERG. 1976. Construction and analysis of viable deletion mutants of simian virus 40. *J. Virol.* **18:**664.

SIMMONS, D. T. and M. A. MARTIN. 1978. Common methionine-tryptic peptides near the amino-terminal end of primate papovavirus tumor antigens. *Proc. Natl. Acad. Sci.* **75:**1131.

SLEIGH, M. J., W. C. TOPP, R. HANICH, and J. P. SAMBROOK. 1978. Mutants of SV40 with an altered small t protein are reduced in their ability to transform cells. *Cell* **14:**79.

SOEDA, E. and B. E. GRIFFIN. 1978. Sequences from the genome of a non-transforming mutant of polyoma virus. *Nature* **276:**294.

SOLTER, D., L. SHEVINSKY, B. B. KNOWLES, and S. STRICKLAND. 1979. The induction of antigenic changes in a teratocarcinoma stem cell line (F9) by retinoic acid. *Dev. Biol.* **70:**515.

STRICKLAND, S. and V. MAHDAVI. 1978. The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell* **15:**393.

SWARTZENDRUBER, D. E. and J. M. LEHMAN. 1975. Neoplastic differentiation: Interaction of SV40 and polyoma virus with murine teratocarcinoma cells in vitro. *J. Cell Physiol.* **85:**179.

SWARTZENDRUBER, D. E., T. D. FRIEDRICH, and J. M. LEHMAN. 1977. Resistance of teratocarcinoma stem cells to infection with SV40: Early events. *J. Cell Physiol.* **93:**25.

TEGTMEYER, P., K. RUNDELL, and J. K. COLLINS. 1977. Modification of simian virus 40 protein A. *J. Virol.* **21:**647.

TEICH, N. M., R. WEISS, G. R. MARTIN, and D. R. LOWY. 1977. Virus infection of murine teratocarcinoma stem cell lines. *Cell* **12:**973.

TENEN, D. G., P. BAYGEL, and J. LIVINGSTON. 1975. Thermolabile T (tumor) antigen from cells transformed by a temperature-sensitive mutant of simian virus 40. *Proc. Natl. Acad. Sci.* **72:**4351.

THIMMAPPAYA, B. and T. SHENK. 1979. Nucleotide sequence analysis of viable deletion mutants lacking segments of the SV40 genome coding for small t antigen. *J. Virol.* **30:**668.

TJIAN, R. 1978. The binding site on SV40 DNA for a T antigen-related protein. *Cell* **13:**165.

TONEGAWA, S., A. M. MAXAM, R. TIZARD, O. BERNARD, and W. GILBERT. 1978. Sequence of a mouse germ-line gene for a variable region of an immunoglobin light chain. *Proc. Natl. Acad. Sci.* **75:**1485.

TOPP, W., J. D. HALL, D. RIFKIN, A. J. LEVINE, and R. POLLOCK. 1977. The characterization of SV40-transformed cell lines derived from mouse teratocarcinoma: Growth properties and differentiated characteristics. *J. Cell Physiol.* **93:**269.

VAN DEN BERG, J., H. VAN OOYEN, N. MANTEL, A. SCHAMBÖCK, G. GROSVELD, R. FLAVELL, and C. WEISSMAN. 1978. Comparison of rabbit and mouse β globin genes showing strong evolutionary divergence of two homologous pairs of introns. *Nature* **276:**37.

VOLCKAERT, G., J. FEUNTEUN, L. V. CRAWFORD, P. BERG, and W. FIERS. 1979. Nucleotide sequence deletions within the coding region for small-t antigen of simian virus 40. *J. Virol.* **30:**674.

YANG, R. C. A. and R. WU. 1979. BK virus DNA sequence: Extent of homology with simian virus 40 DNA. *Proc. Natl. Acad. Sci.* **76:**1179.

ZAIN, B. S., J. SAMBROOK, R. J. ROBERTS, W. KELLER, M. FRIED, and A. R. DUNN. 1979. Nucleotide sequence analysis of the leader segments in a cloned copy of adenovirus 2 fiber mRNA. *Cell* **16:**851.

Cold Spring Harbor Symposia on Quantitative Biology

Regulation of SV40 Gene Expression through RNA Splicing

G. Khoury, J. C. Alwine, R. Dhar, et al.

Cold Spring Harb Symp Quant Biol 1980 44: 41-54
Access the most recent version at doi:10.1101/SQB.1980.044.01.007