DNA Sequence Analysis of Simian Virus 40 Mutants with Deletions Mapping in the Leader Region of the Late Viral mRNA's: Mutants with Deletions Similar in Size and Position Exhibit Varied Phenotypes

ALICE BARKAN AND JANET E. MERTZ*

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706

The nucleotide sequences of 10 viable yet partially defective deletion mutants of simian virus 40 were determined. The deletions mapped within, and, in many cases, 5' to, the predominant leader sequence of the late viral mRNA's. They ranged from 74 to 187 nucleotide pairs in length. Six of the mutants had lost the sequence that corresponds to the "cap" site (5' terminus) of the most abundant class of 16S mRNA's. One of these mutants had a deletion that extended 103 nucleotide pairs into the region preceding this primary cap site and, therefore, was missing many secondary cap sites as well. A seventh mutant lacked the entire major 16S leader sequence except for the first six nucleotides at its 5' end and the last nine at its 3' end. Although these mutants differed in the size and position of their deletions, we were unable to discover any simple correlations between their growth characteristics and their DNA sequences. This finding indicates that the secondary structures of the RNA transcripts may play a more important role than the exact nucleotide sequence of the RNAs in determining how they function within the cell.

Lytic infection of monkey kidney cells by simian virus 40 (SV40) can be divided into early and late phases demarcated by the onset of viral DNA replication. Late in infection, two major species of viral mRNA, sedimenting at 16S and 19S, can be found in the cytoplasm of infected cells. Each species is the product of a "splicing" event which joins a 5'-terminal "leader" sequence to the "body" of the mRNA containing the coding regions for one or more of the viral capsid proteins (1, 2, 4, 6, 8, 11, 15, 20). The leaders fall into numerous classes that vary with respect to the position of both their 5' and 3' termini (6, 20). Although most, if not all, late viral mRNA's found in cells infected by SV40 and the related papovavirus, polyoma virus (17), contain a leader sequence, their function(s) and the mechanism(s) by which they are formed are presently unknown.

An approach that has classically been used to determine the function of a gene is to examine mutants defective in the gene of interest. A number of SV40 mutants with deletions mapping in the leader region have already been isolated (9, 14, 19, 27, 29). Most of them are viable but somewhat defective in their rate of infectious virion production as indicated by both the late appearance and small size of the plaques they produce (9, 19, 29) and by their low production of PFU per cell during a single cycle of

growth (27). Knowledge of the nucleotide sequences of such mutants should aid in defining both the function(s) of the leader sequences and the molecular signals involved in the biogenesis of mature mRNA's. The results of DNA sequence analysis of 10 mutants of SV40 containing deletions that map within the region coding for the late leader sequences are presented. An attempt has been made to correlate the growth characteristics of these mutants with their altered nucleotide sequences.

MATERIALS AND METHODS

Origin of mutants. The deletion mutants discussed in this paper, dl-801 through dl-810, are naturally arising mutants of SV40 strain WT800. Their selection, propagation, and growth characteristics have been reported elsewhere (19). In brief, these mutants were isolated from a virus stock that had been serially passaged four times at high multiplicities of infection in primary African green monkey kidney cells. They were selected from viral DNA that was resistant to cleavage by the restriction endonuclease *Hpa*II (an enzyme that cleaves the wild-type SV40 genome once at map position 0.725 [see Fig. 1A]), yet formed plaques on CV-1P cells in the absence of helper virus. Virus stocks were prepared in MA-134 cells and titrated on CV-1P cells, both established lines of African green monkey kidney cells. Viral DNAs were purified essentially as described previously (18).

Mapping the restriction endonuclease cleav-

age sites in the mutant genomes. Purified DNA of wild-type SV40 and each of the mutants was incubated with an excess of the indicated restriction enzyme at 37°C in a standard reaction buffer (10 mM Tris-hydrochloride, pH 7.6, 10 mM MgCl₂, 1 mM dithiothreitol, and 100 µg of gelatin per ml). The products of the reactions were separated by electrophoresis in 1.3% agarose gels in a Tris-acetate-EDTA buffer (0.16 M Tris-hydrochloride, 0.08 M NaOAc, 9 mM disodium EDTA, pH 8.3). Bands were visualized by UV fluorescence of ethidium bromide-stained DNA.

Nucleotide sequence analysis. The exact size and location of the deletion in each mutant genome were determined by the dideoxynucleotide sequencing method of Sanger et al. (23). Primers for the sequencing reactions were made by cleaving WT800 DNA with the appropriate restriction endonucleases at 37°C in a standard reaction buffer. The resulting DNA fragments were separated on a 6% polyacrylamide gel and visualized by UV fluorescence of ethidium bromide-stained DNA. Bands were cut out of the gel, and the DNAs were eluted by incubating the crushed gel slices overnight at 20°C in a solution of 0.5 M NH₄OAc, 0.01 M Mg(OAc)₂, 0.1% sodium dodecyl sulfate, and 0.1 mM EDTA. The DNAs were then purified by extraction with phenol, followed by extraction with ether. After ethanol precipitation, the DNAs were suspended in distilled water and stored at -20°C.

Templates were prepared by cleaving the DNA with an appropriate restriction enzyme to make it linear, followed by digestion with exonuclease III (5 U/µg, in standard reaction buffer at 37°C for 45 min) to yield a region of single-stranded DNA.

After inactivation of the exonuclease III by incubation at 70°C for 10 min and annealing of the primer to the template, polymerization reactions were catalyzed by reverse transcriptase at 42°C. The primer was then removed by cleavage with an appropriate restriction endonuclease. The polymerization products were denatured by boiling in 95% formamide and then fractionated on 0.35-mm thick 8 or 20% polyacrylamide gels containing 8 M urea (22).

Single-cycle growth curves. Freshly confluent monolayers of CV-1P cells were infected with WT800 or dl-810 virus at a multiplicity of infection of approximately 30 PFU per cell. After incubation for 2 h at 37°C, the cells were washed twice with prewarmed medium to remove unadsorbed virus. Dulbecco's modification of Eagle medium supplemented with 2% fetal bovine serum was added, and incubation was continued at 37°C.

Production of PFU. Infected cells plus the accompanying culture fluid were harvested at the indicated times after infection, frozen and thawed three times, extracted with 0.2 volume of chloroform and centrifuged to remove cellular debris. After addition of 0.05 volume of 1.0 M Tris-hydrochloride (pH 7.6) and fetal bovine serum to 2%, the concentration of PFU in the supernatant fluid was determined by plaque assays on monolayers of CV-1P cells as described previously (18).

Rate of viral DNA synthesis. At 8-h intervals, newly synthesized DNA was radiolabeled by replacing the culture fluid with 4 ml of medium containing 2% dialyzed bovine serum and 25 µCi of [³H]thymidine

(20 Ci/mmol). The cells were then incubated for an additional 7.7 ± 0.2 h at 37°C. Cells were lysed by the method of Hirt (10) at the times indicated. After extraction of the supernatant with phenol, the amount of radiolabeled viral DNA present was determined as follows: DNA was fractionated in 1.3% agarose gels, stained with ethidium bromide, and visualized by UV fluorescence. The band containing supercoiled SV40 DNA was cut out, the agarose was melted, and the radioactivity within the band was determined by liquid scintillation counting. The averages of duplicate assays are indicated.

Enzymes and biochemicals. Exonuclease III and the restriction endonucleases *HincII*, *KpnI*, and *PvuII* were purchased from New England Biolabs. *HpaI* was purified by a modification of the procedure of Sharp et al. (26). *BglII* was prepared by an unpublished procedure of R. Roberts. Avian myeloblastosis virus reverse transcriptase was obtained from Life Sciences, Inc. Deoxynucleotides and dideoxynucleotides were purchased from PL Biochemicals. [α-³²P]dATP and [α-³²P]dCTP, each at approximately 350 Ci/mmol, were purchased from Amersham Corp. and New England Nuclear Corp., respectively.

RESULTS

Ten viable SV40 mutants each lacked a segment of the region of the viral genome that encodes the leader sequences of the late viral mRNA's. Mutants dl-801 through dl-810 were selected for resistance to cleavage by the restriction endonuclease *HpaII* which cleaves the wild-type SV40 genome once at map position 0.725 (Fig. 1A). The altered electrophoretic mobility of their *HindIII* + *III* C fragment (map position 0.648 to 0.754) and electron microscopic heteroduplex analysis indicated that these mutants each have a small deletion in the region of the genome that includes the *HpaII* cleavage site (19). Because these deletions fall within the region coding for the leaders to the late viral mRNA's, the mutants must make mRNA's with altered leader sequences. They therefore might prove useful in elucidating the function(s) and mechanism(s) of formation of these leader sequences.

To localize the position of each deletion more precisely, we determined the patterns produced by cleavage of each mutant DNA with the restriction endonucleases *PvuII*, *KpnI*, and *HpaI* (Fig. 1A). The results of these experiments are summarized in Table 1.

DNA sequence determination. The exact nucleotide sequence around the region deleted in each mutant was determined by the dideoxynucleotide chain termination method of Sanger et al. (23). The restriction maps were used to aid in the choice of DNA primers.

The positions of the deletions in mutants dl-801, dl-802, dl-804, and dl-805 were determined by using the *BglII-KpnI* fragment to prime the

FIG. 1. (A) Map of the late region of the SV40 genome. Map coordinates are expressed as the fractional length of wild-type (WT) SV40 where 0.00 is defined as the *EcoRI* restriction endonuclease cleavage site. Nucleotide residue numbers refer to the wild-type SV40 sequence of Reddy et al. (21). *ori* is the approximate location of the origin of viral DNA replication. The 5' ends of the 19S and 16S leaders are depicted by dashed lines because of their heterogeneous nature (see the text). (B) Positions of the deletions in mutants dl-801 through dl-810 with respect to landmarks in the late leader region. The primers used in the sequencing reactions are represented by dashed lines. Deleted sequences are represented by solid lines. The position of the major cap and splice sites and the initiation codon for a putative leader protein are as indicated. An open reading frame is represented by the hatched box within the major leader.

TABLE 1. Position of deletion versus rate of plaque growth for leader region mutants

Mutant	Restriction enzyme site ^a					Size of deletion (np) ^b	Missing nucleotide residues ^c	Mean plaque diameter (mm) ^d at days postinfection:			
	<i>PvuII</i> (0.710)	<i>KpnI</i> (0.715)	<i>HpaII</i> ^d (0.725)	<i>HpaI</i> (0.754)	<i>HpaI</i> (0.758)			9.9	10.8	11.8	15.1
WT800	+	+	+	+	+			1.4	2.4	3.2	5.6
dl-801	+	+	—	+	+	74	249–322 ^e	0.9	1.8	2.6	4.4
dl-802	+	+	—	+	+	80	262–341	0.5	0.9	1.3	2.9
dl-803	+	—	—	+	+	75	203–277 ^e	0.3	0.9	1.8	3.8
dl-804	+	+	—	+	+	130	255–384 ^e	0.2	0.9	1.3	3.3
dl-805	+	+	—	—	+	187	249–435	0.3	1.0	1.3	3.1
dl-806	—	—	—	+	+	138	140–277 ^e	<0.1	<0.1	0.3	1.9
dl-807	+	+	—	+	+	169	224–392 ^e	0.3	1.2	2.4	3.5
dl-808	+	+	—	+	+	161	223–383	0.5	1.4	2.4	4.2
dl-809	+	—	—	+	+	170	217–386 ^e	0.5	1.2	2.0	3.7
dl-810	+	—	—	+	+	175	215–389 ^e	<0.1	0.4	1.2	3.4

^a Determined as described in the text. +, Presence of indicated restriction endonuclease cleavage site; —, absence of indicated endonuclease cleavage site. Map position (unit) is shown within parentheses.

^b np, Nucleotide pairs; calculated from the nucleotide sequences determined in this paper.

^c Determined as described in the text. Nucleotide residues are numbered by the wild-type SV40 DNA sequence of Reddy et al. (21).

^d Data from Mertz and Berg (19).

^e Where ambiguity exists as to the precise residue numbers missing due to a repeat in the sequence at the ends of the deletion, the deletion is assumed to begin as far distal to the origin of replication as possible.

sequencing reaction (Fig. 1B). Mutants dl-803 and dl-807 through dl-810 were sequenced by using the *BglII-PvuII* fragment, and dl-806 was sequenced by using the smallest *HpaI* fragment

(0.754 to 0.758 map units) as the primer. An autoradiogram of the gel from which the sequences of three of these mutants were determined is shown in Fig. 2. The nucleotide se-

FIG. 2. Autoradiogram of the sequencing gel from which the positions of the deletions in mutants dl-803 (a), dl-809 (b), and dl-810 (c) were determined. After performing the sequencing reactions, samples were denatured and loaded onto a 0.35-mm thick 20% polyacrylamide-8 M urea gel. The gel was run at 900 V until the bromophenol blue dye marker reached the bottom. Each lane is labeled with the letter corresponding to the dideoxynucleotide added to the polymerization reaction: T, thymidine; C, cytosine; G, guanosine; A, adenosine. The sequence, read from bottom to top, corresponds to the early strand of DNA read in a clockwise direction. The sites of the deletions (indicated by arrows) were deduced by comparing the nucleotide sequence of each mutant with that of wild-type SV40 (21).

quences of all 10 mutants are shown in Fig. 3 and summarized in Fig. 1B and Table 1.

Mutants with deletions similar in size and position exhibited considerable variation in their phenotypes. As previously noted (19), mutants dl-801 through dl-810 all reproducibly yield late-appearing plaques compared with wild-type SV40 when assayed on CV-1P cells (Table 1). However, the times of appearance of the resultant plaques vary considerably from mutant to mutant. Plaques formed by mutant

dl-801 appeared almost as soon as those of the parental strain WT800. On the other hand, mutants dl-806 and dl-810 formed plaques which were not visible until 10 to 11 days postinfection, whereas those formed by WT800 could readily be seen within 8 days.

An altered phenotype could be observed even during a single cycle of growth when dl-810 was compared with WT800. When CV-1P cells were infected with the wild-type or mutant virus at high multiplicities of infection, PFU first appeared at the same time, but accumulated at a considerably slower rate in the mutant infection (Fig. 4B). However, their onset and initial rate of viral DNA synthesis as determined by the incorporation of [3 H]thymidine during consecutive 8-h labeling periods were essentially identical (Fig. 4A). When this experiment was repeated with short (i.e., 1 h) labeling periods, the initial onset and rate of viral DNA synthesis was again very similar in cells infected by dl-806, dl-810, or WT800 virus (data not shown). Thus, as expected from the location of their deletions, these mutants appear to be defective in a function that occurs after the onset of viral DNA replication.

DISCUSSION

The 5'-terminal leader sequence of the SV40 late mRNA's is encoded by DNA sequences not adjacent to those coding for the known late proteins. Examples of such leader sequences have been found in other viral and cellular systems as well (3, 5). The functions of these sequences are, as yet, unknown. In this paper, we present the DNA sequences of 10 mutants of SV40 containing deletions mapping in the leader region of the late viral mRNA's. Correlation of the phenotypes of these mutants with their nucleotide sequences should be helpful in defining the functions of this particular leader region and the signals involved in the biogenesis of the late viral mRNA's.

The position of the deletion in each of these mutants with respect to various important landmarks in the leader region is diagrammed in Fig. 3. Included are sites where the late viral mRNA's are capped and spliced (6, 20) and where the potential initiation codon for a possible protein coding region is located (see below) (21). The deletions in dl-801, dl-802, dl-804, and dl-805 fall entirely within the region encoding the major leader. These mutants retained all of the sites that are capped in the wild-type mRNA's, the 16S mRNA splice joint, and two of the three 19S mRNA splice joints. dl-803 and dl-806 through dl-810 deleted the major 16S cap site. All 10 mutants have retained the 3' end of the major

FIG. 3. Nucleotide sequences of mutants dl-801 through dl-810 in the region coding for the leaders of the late mRNA's. The sequence shown is the SV40 DNA strand that corresponds to the late mRNA's. Numbers refer to the residue numbers of wild-type SV40 as determined by Reddy et al. (21). 5' termini and splice points are as determined by Ghosh et al. (6). The position of a possible initiation codon for a leader protein is indicated (positions 253 to 255). Deleted nucleotides are bracketed. Where the precise endpoints of a deletion are ambiguous due to a repeated sequence at its ends, the bracket is positioned as far to the right (i.e., from the origin of replication) as possible.

leader at residue 444, perhaps because its presence is essential for viability (see below).

It might be expected that the region of the viral genome preceding that coding for the 5' termini of the late mRNA's is involved in the initiation of viral RNA synthesis. However, we failed to find a direct correlation between the plaque morphologies of these mutants and the degree to which their deletions extend toward the origin of viral DNA replication (Table 1). dl-806, which along with dl-810 gives rise to the latest appearing plaques, had a deletion that extended 63 nucleotide pairs closer to the origin of replication than any of the other mutants. Part of the explanation for its poor growth may be that enzymes responsible for the initiation of transcription or processing of the 5' end of the viral RNA can recognize any of several sites, most of which have been deleted in dl-806. Preliminary experiments indicate that the amount of [³H]uridine incorporated into viral RNA during 5-min pulse-labeling periods is approximately the same in monkey cells infected with either dl-806 or WT800 virus (Barkan, unpublished data). Thus, the defect does not appear to be at the level of the efficiency of transcription initiation.

It is possible that deletions of the wild-type capping sequences could interfere with the processing of the 5' ends of mRNA precursors. Three predominant 19S 5' termini and one major 16S 5' terminus have been mapped onto the wild-type SV40 genome by Ghosh et al. (6) (Fig. 3). Loss of the major cap site by itself does not necessarily have a significant effect on viral replication since mutants that lack it range in phenotype from growing almost as well as wild-type SV40 (e.g., dl-808) to being quite defective (e.g., dl-810) (Table 1 and Fig. 4B). The nucleotide sequence of the 5' termini of the viral mRNA's found in cells infected by several of these mutants have been determined by P. Ghosh and collaborators (manuscript in preparation). It appears that mutant dl-805, although having a deletion that begins six nucleotides after the major cap site, synthesizes little, if any, mRNA that is capped at this site. Thus, this mutant either lacks the signal needed for capping at that site or the RNA is not detected due to instability. Further analysis of the viral mRNA's made after infection with this and other mutants should help to define the sequence information required for the production of stable, capped mRNA.

The role of RNA splicing in the biogenesis of

FIG. 4. Comparison of (A) the rate of viral DNA synthesis and (B) the production of PFU during single cycles of lytic growth in CV-1P cells infected with WT800 (●) or dl-810 (○) virus. DNA was radiolabeled for consecutive 8-h periods as described in the text. 1/40 of each Hirt supernatant was fractionated on an agarose gel, and the radioactivity in form I DNA was determined. This value was then multiplied by a factor of 40 so as to represent the total counts per minute (cpm) in form I DNA per dish of cells.

functional mRNA and the molecular signals involved in accurate splicing are unknown. Deletion of the intron that is spliced out of the SV40 16S mRNA precursor prevents the accumulation of stable 16S mRNA (7). This result suggests that splicing may be required for the biogenesis of stable 16S mRNA. The fact that none of these viable mutants has a deletion that extends beyond the 3' end of the leader region supports the idea that the presence of an intact splice joint is indeed necessary for viability. However, the nearness of a deletion to the major splice joint (residue 444) does not correlate with its plaque phenotype (Table 1). In fact, the

deletion in mutant dl-805 extends to within nine nucleotide pairs of the 16S splice junction, yet it is only mildly defective for growth (Table 1).

Although splicing may be required for the biogenesis of stable 16S mRNA, this does not appear to be a general phenomenon. When polysomal RNA from cells infected with mutant dl-805 is examined, the unspliced form is the most abundant and perhaps the only species of 19S mRNA found (P. Ghosh, P. Roy, A. Barkan, J. Mertz, S. Weissman, and P. Lebowitz, *Proc. Natl. Acad. Sci. U.S.A.*, in press). Nevertheless, this mutant is not nearly as defective as dl-806, a mutant which makes mostly normally spliced 19S mRNA. Although splicing appears to be affected by these deletions, deficiencies in splicing cannot entirely account for the altered phenotypes of the mutants.

The possibility that the leader sequence itself functions as an mRNA has not been ruled out. The leader region contains an open reading frame of 62 triplet codons (20) that might be used to synthesize a basic, histone-like protein. Although no such protein has yet been positively identified in SV40-infected cells, evidence is accumulating that it may actually be synthesized. First, these mutants define a new complementation group (Mertz, unpublished data). Second, the closely related human papovavirus BKV contains a corresponding sequence in which the potential amino acid sequence is more closely related to the putative SV40 leader protein than is the nucleotide sequence (31). Third, the fact that 12 of the 13 nucleotides preceding its initiation codon are identical to those preceding the known initiation codon for VP2 suggests that the gene may actually be translated *in vivo*. Finally, we and others (P. Southern and P. Berg; G. Khoury; personal communications) have tentatively identified a possible candidate for this polypeptide. A small (approximately 8,000-dalton) protein is found in extracts of cells infected with wild-type virus but is absent from cells infected with viable mutants containing deletions of the putative coding region (data not shown). However, as others have already pointed out (29), assuming this sequence is translated, the resulting protein must not be essential for growth of the virus in cultured cells since 9 of our 10 viable mutants lack its initiation codon and the remaining mutant (dl-802) is a frameshift mutant that has a UGA stop codon just 14 triplets after initiation. Clearly, the presence or absence of this small protein cannot account for the variety of phenotypes observed.

The leader region may be involved in the regulation of translation of the bodies of the late mRNA's. Legon (16) has found that 40S ribosomal subunits protect the leader of polyoma

virus late mRNA in *in vitro* binding assays. The 40S subunit bound to the leader may be the natural precursor to the 80S complex which forms at the initiation codon. The fact that mutants such as dl-805, which lack essentially the entire major leader sequence, can form visible plaques almost as quickly as wild-type SV40 suggests that the function of the leader does not depend on a specific sequence. It may be that the length of RNA between the cap and the initiation codon determines the translational efficiency. If the binding of the 40S subunit at the cap and the assembly of the 80S complex at the initiation codon are slow relative to the movement of the 40S subunit along the RNA, then translational efficiency would be improved by allowing a few 40S subunits to line up and feed into the initiation site. Examination of polysomal RNA in cells infected by these mutants will reveal whether a minimum length of RNA preceding the initiation codon is required for efficient translation.

The arrangement of nucleosomes on the SV40 minichromosome is not entirely random with respect to the viral DNA sequence. A small region on the late side of the origin of viral DNA replication (approximately 0.67 to 0.74 map units) is preferentially cleaved by endonucleases (25, 30) and often appears nucleosome free when inspected by electron microscopy (12, 24). The mutants dl-801 through dl-810 all contain deletions which fall within or overlap this nuclease-sensitive region. By interfering with the chromatin structure, these deletions may indirectly affect processes such as transcription or packaging of the minichromosome into virions.

The most striking finding of our work is that the deletions in mutants dl-809 and dl-810 map in almost identical positions (Fig. 3) yet their plaque morphologies are markedly different. Mutant dl-810 lacks only five nucleotide pairs more than the 170 deleted in dl-809; nevertheless, its plaques appear significantly later (Table 1). "S1 mapping" (28) failed to resolve any other differences between the two genomes (data not shown), although this does not rule out the existence of single base differences elsewhere in their genomes. Data of Ghosh et al. (manuscript in preparation) indicate that the structures of the leaders made by these two mutants are quite different with respect to the positions of both their caps and splices. The deletions in other less defective mutants (e.g., dl-803 and dl-805) overlap both ends of the deletion in dl-810 (Fig. 1B). Therefore, the loss of the specific five additional nucleotides in dl-810 is not responsible for its poor growth compared to dl-809. Consequently, we propose that the small change in

DNA sequence between these two mutants may lead to different preferred secondary structures in the RNAs resulting in different treatment by cellular enzymes somewhere along the pathway from primary transcription to translation of the mature mRNA's.

Khoury et al. (13) found an equally complex relationship between the size and position of some deletions in the early region of SV40 and the phenotypes of the mutant viruses. Proximity of these deletions to known splicing sites did not correlate with the amount of stable mRNA observed in the cytoplasm. These findings lead us to suggest that the secondary structure of the RNA in both the early and late regions plays a crucial role in the ultimate expression of the gene. Whether this control is at the level of transcription, processing, stability, transport of the RNA from the nucleus, or translation remains to be determined.

ACKNOWLEDGMENTS

We thank Gail Mazzara and Anne Johnson for teaching us DNA sequencing; Shi-Da Yu for technical assistance; Jerry Hertz, Paul Lebowitz, Jeff Ross, Bill Sugden, Howard Temin, and Sherman Weissman for helpful comments on the manuscript; and Joseph Beard for the gift of reverse transcriptase.

This work was supported by Public Health Service research grants CA-07175 and CA-22443 from the National Cancer Institute. A.B. was supported by a fellowship from the Wisconsin Alumni Research Foundation and Public Health Service training grant T32 CA-09135-04.

LITERATURE CITED

- Aloni, Y., R. Dhar, O. Laub, M. Horowitz, and G. Khoury. 1977. Novel mechanism for RNA maturation: the leader sequences of simian virus 40 mRNA are not transcribed adjacent to the coding sequences. *Proc. Natl. Acad. Sci. U.S.A.* 74:3686-3690.
- Berk, A. J., and P. A. Sharp. 1978. Spliced early mRNAs of simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 75:1274-1278.
- 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. U.S.A.* 75:4853-4857.
- Celma, M. L., R. Dhar, and S. M. Weissman. 1977. Comparison of the nucleotide sequence of the messenger RNA for the major structural proteins of SV40 with the DNA sequence encoding the amino acids of the protein. *Nucleic Acids Res.* 4:2549-2559.
- Chambon, P. 1977. Summary: the molecular biology of the eukaryotic genome is coming of age. *Cold Spring Harbor Symp. Quant. Biol.* 52:1209-1234.
- 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.
- Gruss, P., C.-J. 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.
- Haegeman, G., and W. Fiers. 1978. Evidence for "splicing" of SV40 16S mRNA. *Nature (London)* 273:70-73.
- Haegeman, G., H. van Heuverswyn, D. Gheysen, and W. Fiers. 1979. Heterogeneity of the 5' terminus of late

- mRNA induced by a viable simian virus 40 deletion mutant. *J. Virol.* **31**:484-493.
10. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365-369.
 11. 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.
 12. Jakobovits, E. B., S. Bratosin, and Y. Aloni. 1980. A nucleosome-free region in SV40 minichromosomes. *Nature (London)* **285**:263-265.
 13. Khoury, G., P. Gruss, R. Dhar, and C.-J. Lai. 1979. Processing and expression of early SV40 mRNA: a role for RNA conformation in splicing. *Cell* **18**:85-92.
 14. Lai, C.-J., 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.
 15. Lavi, S., and Y. Groner. 1977. 5'-terminal sequences and coding region of late simian virus 40 mRNAs are derived from non-contiguous segments of the viral genome. *Proc. Natl. Acad. Sci. U.S.A.* **74**:5323-5327.
 16. Legon, S. 1979. The binding of ribosomes to polyoma virus RNA. *J. Mol. Biol.* **134**:219-240.
 17. Legon, S., A. Flavell, A. Cowie, and R. Kamer. 1979. Amplification in the leader sequence of late polyoma virus mRNAs. *Cell* **16**:373-388.
 18. Mertz, J. E., and P. Berg. 1974. Defective simian virus 40 genomes: isolation and growth of individual clones. *Virology* **62**:112-124.
 19. Mertz, J. E., and P. Berg. 1974. Viable deletion mutants of simian virus 40: selective isolation by means of a restriction endonuclease from *Hemophilus parainfluenzae*. *Proc. Natl. Acad. Sci. U.S.A.* **71**:4879-4883.
 20. Reddy, V. B., P. K. Ghosh, P. Lebowitz, and S. Weissman. 1978. Gaps and duplicated sequences in the leaders of SV40 16S RNA. *Nucleic Acids Res.* **5**:4195-4213.
 21. Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, and S. M. Weissman. 1978. The genome of simian virus 40. *Science* **200**:494-502.
 22. Sanger, F., and A. Coulson. 1978. The use of thin acrylamide gels for DNA sequencing. *FEBS Lett.* **87**:107-110.
 23. Sanger, F., S. Nicklen, and A. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**:5463-5467.
 24. Saragosti, S., G. Moyne, and M. Yaniv. 1980. Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. *Cell* **20**:65-73.
 25. Scott, W. A., and D. J. Wigmore. 1978. Sites in simian virus 40 chromatin which are preferentially cleaved by endonucleases. *Cell* **15**:1511-1518.
 26. Sharp, P. A., B. Sugden, and J. Sambrook. 1973. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **12**:3055-3063.
 27. Shenk, T. E., J. Carbon, and P. Berg. 1976. Construction and analysis of viable deletion mutants of simian virus 40. *J. Virol.* **18**:664-671.
 28. 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 simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* **72**:989-993.
 29. 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.
 30. Varshavsky, A. J., O. Sundin, and M. Bohn. 1979. A stretch of "late" SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes. *Cell* **16**:453-466.
 31. Yang, R., and R. Wu. 1979. BK virus DNA sequence: extent of homology with simian virus 40 DNA. *Proc. Natl. Acad. Sci. U.S.A.* **76**:1179-1183.