MOLECULAR AND CELLULAR BIOLOGY, Apr. 1984, p. 813–816
0270-7306/84/040813-04$02.00/0

Vol. 4, No. 4

The Number of Ribosomes on Simian Virus 40 Late 16S mRNA Is Determined in Part by the Nucleotide Sequence of Its Leader

ALICE BARKAN† AND JANET E. MERTZ*

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

Received 3 October 1983/Accepted 3 January 1984

The size distributions of polyribosomes containing each of three simian virus 40 late 16S mRNA species that differ in nucleotide sequence only within their leaders were determined. The two 16S RNA species with shorter leaders were incorporated into polysomes that were both larger (on average) and narrower in size distribution than was the predominant wild-type 16S RNA. Therefore, the nucleotide sequence of the leader can influence the number of ribosomes present on the body of an mRNA molecule. We propose a model in which the excision from leaders of sizeable translatable regions permits more frequent utilization of internally located translation initiation signals, thereby enabling genes encoded within the bodies of polygenic mRNAs to be translated at higher rates. In addition, the data provide the first direct evidence that VP1 can, indeed, be synthesized in vivo from the species of 16S mRNA that also encodes the 61-amino acid leader protein.

Procaryotic mRNAs contain sequences upstream from translation initiation codons that facilitate ribosome recognition and binding (reviewed in reference 5). The translation of eucaryotic mRNAs, however, is not strictly dependent on AUG-proximal sequences: there is little conservation of sequences upstream of AUGs that function as initiation codons (8–10a), and changes in the sequence context of some AUG codons have little or no effect on their ability to function as initiation signals (14). Results presented below indicate that RNAs differing in nucleotide sequence only within their leaders are incorporated into polyribosomes that differ both in average size and in distribution width.

The RNAs analyzed in the experiment described here are products of the late region of simian virus 40 (SV40). The predominant late RNA encoded by wild-type (WT) virus (RNA species *a* in Fig. 1A) is 16S in size and is the product of a single splicing event that joins a leader segment to a body segment encoding the 364-amino acid virion protein VP1. Two minor species of 16S RNA (*b* and *c*) also seen in WT-infected cells are spliced a second time within their leaders. A fourth 16S RNA species (*d*), encoded by the viable mutant *dl-805* (1, 12), lacks sequences within its leader because of a deletion in its DNA template. Although these four RNA species differ in sequence within their leaders, their bodies are identical.

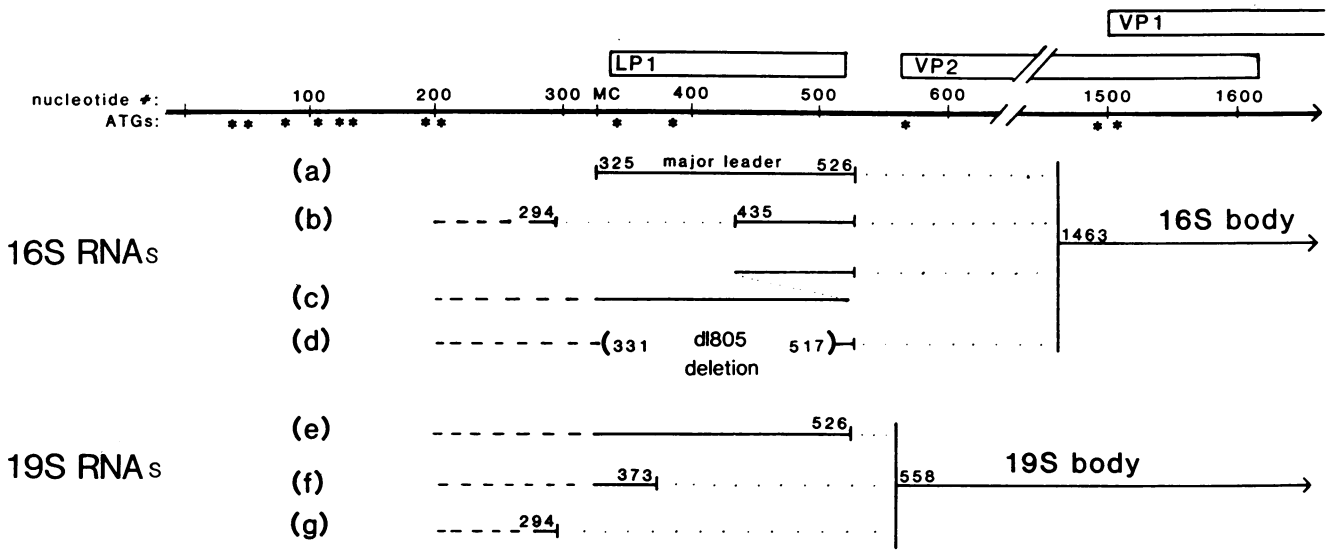
To determine whether sequences within the leader can affect translation of genes encoded within the body of an RNA, the distribution in polyribosomes of these 16S RNA species, as well as of the viral late 19S RNAs (species *e*, *f*, and *g* in Fig. 1A) which differ from the 16S RNAs in both their leaders and body, were compared. Ribonucleoprotein isolated from cells coinfected with WT800 and *dl*805 was fractionated in a sucrose gradient (Fig. 2). Fractions of the gradient were pooled as indicated in the figure. The viral RNA present in each sample was analyzed by S1 nuclease mapping (2) with a probe that enabled the 19S and three of the four 16S RNA species to be distinguished (Fig. 1B). The results (Fig. 3, lanes 1 though 10) indicate that the major 16S RNA species (*a*) was most abundant in pool 5, whereas the amounts of the other three RNA species (*b*, *d*, and *e* plus *f* plus *g*) each peaked in pool 3. Densitometric analysis of the DNA protected by unfractionated cytoplasmic RNA (cyto *b*) indicated that the doubly spliced RNA species *b* accounted for only 20 to 25% of the total WT 16S RNA present in these cells. Therefore, the finding that RNA species *b* was more abundant in pool 2 than was the predominant 16S RNA species is particularly striking.

Also noteworthy were the shapes of the polyribosome distributions: whereas species *a* mRNA molecules were found in sizeable numbers throughout the polysome profile, the mRNAs of each of the other species were found almost exclusively on large polysomes. In addition, species *a* may have exhibited three optima (i.e., in pools 3, 5, and 7). Although the latter of these somewhat unexpected observations still needs to be confirmed by additional experiments, a reasonable model for translation of eucaryotic mRNAs, outlined below, predicts just such results.

These data were not an artifact of the mapping procedure, since all hybridization reactions were performed with a molar excess of probe DNA (compare 2× cyto *b* with cyto *b*). Results similar to those shown in Fig. 2 were also obtained when the experiment was repeated with cells infected on another day (data not shown). Therefore, the sedimentation profiles of polyribosomes containing the doubly spliced WT800 16S RNA species *b*, *dl-805* 16S RNA, and 19S RNA were quite different from that of polyribosomes containing singly spliced WT 16S RNA (species *a*).

The number of ribosomes on an mRNA molecule is a function of (i) the frequency of translation initiation of each gene encoded within the RNA molecule; (ii) the rate of movement of the translating ribosomes; and (iii) the size of the translated portion of each gene. The singly spliced WT 16S RNA species *a* is an unusual eucaryotic mRNA in that its translation does not initiate solely from the AUG codon nearest its 5′ end. The initiation codon for VP1 is preceded by two AUGs, at least one of which is a functional initiation codon (see Fig. 1A) (6, 7, 13). However, LP1, the protein encoded within the leader of this RNA species, is only 61 amino acids in length. Therefore, translation of the leader segment of this RNA species can contribute, at most, three to five of the ribosomes present per RNA molecule found in

* Corresponding author.
† Present address: Department of Genetics, University of California, Berkeley, CA 94720.



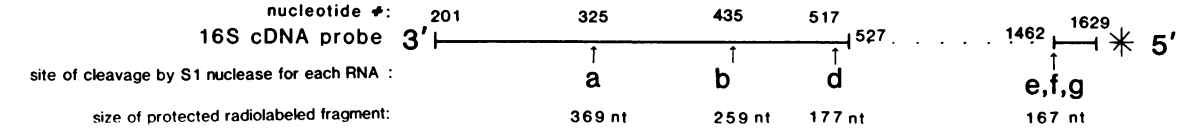


FIG. 1. (A) Structures of the SV40 late RNA species. Each mRNA consists of one of the indicated leader segments joined covalently to the corresponding body segment. MC indicates the major cap site, which encodes the 5' end of approximately 90% of the WT late-strand viral mRNA molecules (3). The remaining 5' ends, represented here by dashed lines, map at numerous sites between nucleotides 28 and 325. Dotted lines are drawn between sequences that are adjacent in the RNA but not in the viral genome. Asterisks indicate the locations of ATG triplets present in the late RNA. Eleven additional ATGs, including that encoding the initiation codon for the virion protein VP3, are present between nucleotides 650 and 1430, which are not shown in this drawing. RNA species *a*, *b*, and *c* make up approximately 80, 15, and 5%, respectively, of total WT 16S RNA (3). RNA species *d* accounts for greater than 90% of *dl-805* 16S RNA (4). RNA species *e*, *f*, and *g* comprise approximately 5, 50, and 40%, respectively, of total WT800 spliced 19S RNA (M. B. Somasekhar and J. E. Mertz, manuscript in preparation). Approximately 80% of *dl-805* cytoplasmic 19S RNA is species *g* (P. Ghosh, personal communication; Somasekhar and Mertz, in preparation). The molar ratios of 16S to 19S RNA are approximately 4 for WT800 and 1 for *dl-805* (A. Barkan, Ph.D. thesis, University of Wisconsin, Madison, 1983). The nucleotide numbering system is that given by Tooze (15). (B) Strategy for distinguishing among the various SV40 late RNA species by mapping with S1 nuclease. The probe for S1 nuclease mapping was derived from pSV1-16SIVS, a cloned mutant of SV40 that lacks nucleotides 527 to 1462 which encode the intervening sequence of the SV40 major late 16S RNA (generously supplied by J. Sklar and P. Berg). The 493-base pair *Sph*I-*Acc*I restriction fragment (SV40 nucleotides 201 to 1629) of pSV1-16SIVS was 5' end labeled with ^{32}P at the *Acc*I site, hybridized to RNA, and treated with S1 nuclease as described in the legend to Fig. 3. Arrows show the points nearest the 5' end of the probe at which S1 nuclease will cleave hybrids between the probe and each of the RNA species indicated in Fig 1A. RNA species *c*, containing both heterogeneity at its 5' end and a tandem duplication of nucleotides 435 to 526, may protect variable lengths of the probe from cleavage with S1 nuclease; fortunately, since it accounts for only a very minor fraction of the total viral RNA, its presence can be ignored in this study.

polysomes. Consequently, we can conclude from the finding that some of the RNA molecules of this species are present in fast-sedimenting polysomes (see lanes 1 to 3 of Fig. 3) that ribosomes must also be bound to the interior of these molecules, the region that encodes VP1. Therefore, these molecules can indeed function in the synthesis of VP1 as well as LP1. In addition, if we assume that the rate of movement of ribosomes along the sequences encoding VP1 is unaffected by sequences within the leader, these data indicate that the frequency of initiation of translation of VP1 is much lower on this predominant, singly spliced 16S RNA species than on either *dl-805* 16S RNA (*d*) or the doubly spliced WT 16S RNA species *b*.

These latter RNAs are similar in structure: whereas *dl-805* RNA lacks nucleotides 331 to 517, the doubly spliced WT 16S RNA species *b* lacks the overlapping nucleotides 295 to 435. Thus, the similar distribution on polysomes of these latter two RNA species may result from their common loss of an element mapping within the leader that affects the frequency of translation initiation of the sequences encoding VP1.

In a modified scanning hypothesis for translation start

codon selection (9, 10, 10a), Kozak has proposed that the 40S ribosomal subunit binds to the 5′ end of an mRNA molecule and migrates along it, forming an initiation complex at one of the AUG codons with an efficiency determined by sequences within its immediate vicinity. According to this model, the two AUGs present in the leader of the singly spliced WT 16S RNA are "leaky" and allow a portion of scanning 40S ribosomal subunits to reach the sequences within the body of the RNA that encode VP1. However, because some of the scanning ribosomes both initiate and terminate translation within the leader, fewer ribosomes ever reach the initiation codon involved in the synthesis of VP1. Therefore, VP1 is translated less efficiently from 16S RNA species *a* than from the 16S RNA species that lack AUGs within their leaders.

On the surface, our finding that the two 16S RNA species lacking AUG codons within their leaders were incorporated into faster sedimenting polyribosomes than was the predominant WT 16S RNA is consistent with the modified scanning hypothesis of Kozak: the absence of the two AUG codons found in the major 16S RNA leader of WT SV40 enables a higher percentage of the scanning 40S ribosomal subunits to reach the initiation codon for VP1. Therefore, initiation of translation of the sequences encoding this protein occurs more frequently, resulting in a larger average number of ribosomes being present on these RNA molecules.

However, the scanning hypothesis fails to predict the peculiar shape of the polysome distribution found for RNA species *a*. Although the presence of upstream weak initiation codons followed by in-phase termination codons would shift

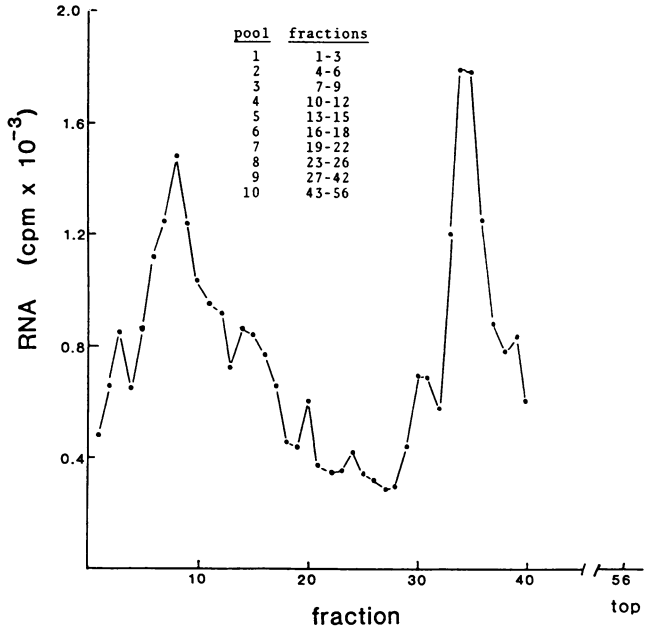


FIG. 2. Fractionation of ribonucleoprotein from SV40-infected cells. CV-1P cells were coinfected with WT800 and *dl-805* (approximately 20 PFU of each virus per cell). Cellular RNA was metabolically labeled with [^{3}H]uridine (50 μCi per ml of complete Dulbecco modified Eagle medium; 38 Ci/mmol) during the 16 h before cell lysis. Forty hours after infection, cells from two 100-mm dishes were harvested and lysed by incubation for 5 min at 0°C in 1 ml of TKM buffer (25 mM Tris-hydrochloride [pH 7.6], 40 mM KCl, 7.5 mM $MgCl_2$) containing 0.5% Nonidet P-40. The nuclei were pelleted by centrifugation at 1,500 rpm for 5 min and washed once. The resulting supernatants were pooled, made 10 μg/ml in heparin, and centrifuged in an SS-34 rotor at 10,000 rpm for 10 min. After removal of a portion for the analysis of unfractionated cytoplasmic RNA (cyto *b* of Fig. 3), the post-mitochondrial supernatant (1 ml) was made 0.5% in sodium deoxycholate, layered onto an 8-ml, 15 to 50% gradient of sucrose in TKM buffer containing a 2-ml pad of 56% sucrose, and centrifuged in an SW41 rotor for 3.25 h at 28,500 rpm and 4°C (11). Fractions (0.2 ml) were collected from the bottom of the gradient. The amount of cellular radiolabeled RNA in 1/30th of each fraction, quantified by a DE81 filter binding assay, is shown. Based upon this distribution of radiolabeled RNA, we assume that monosomes are contained within pool 9.

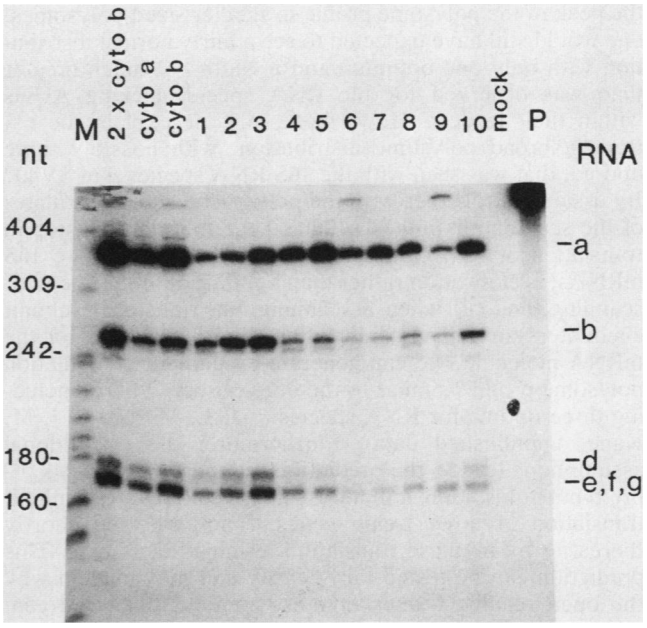


FIG. 3. S1 nuclease mapping of viral RNA species present in ribonucleoprotein fractions. Sodium dodecyl sulfate and EDTA were added to final concentrations of 1% and 7 mM, respectively, to each of the pools of fractions indicated in Fig. 2. RNA samples were extracted twice with phenol-chloroform, precipitated with ethanol, suspended in 10 mM Tris-hydrochloride (pH 7.6)–1 mM EDTA, and analyzed by S1 nuclease mapping as follows. An excess of radiolabeled probe (see Fig. 1B for structure) was hybridized to each RNA sample at 52°C for 16 h in 6 μl of 400 mM NaCl–40 mM PIPES (piperazine-*N*-*N*′-bis(2-ethanesulfonic acid; pH 6.4)–1 mM EDTA–80% formamide. Each sample was then diluted with 60 μl of 250 mM NaCl–30 mM sodium acetate (pH 5.5)–1 mM $ZnSO_4$–10 μg of denatured calf thymus DNA per ml containing S1 nuclease (generously supplied by J. Ross). After incubation for 1 h at 37°C, 20 μl of 2.5 M ammonium acetate (pH 7.5)–50 mM EDTA–400 μg of tRNA per ml was added. Nuclease-resistant material was precipitated with ethanol, denatured by incubation at 100°C for 1 min in 80% formamide–1 mM EDTA–10 mM NaOH, and electrophoresed in a 5% polyacrylamide gel (0.25% bisacrylamide) containing 7 M urea and 90 mM Tris-borate (pH 8.3)–2.5 mM EDTA. Radiolabeled DNA was detected by autoradiography at −70°C with Kodak XAR film. Lanes 1 through 10 of the autoradiogram show the DNAs protected by 6×10^4 cells' worth of RNA isolated from the corresponding pools of fractions shown in Fig. 2. S1 mapping was also performed with 6×10^3 (cyto *b*) and 1.2×10^4 (2× cyto *b*) cells' worth of RNA purified from the unfractionated cytoplasm of cells coinfected with WT800 and *dl-805*. Mock and cyto *a* show, respectively, the results obtained with no added monkey cell RNA and with RNA preparations purified from the cytoplasm of 6×10^3 WT800-infected cells and 6×10^3 *dl-805*-infected cells that were mixed together before mapping. This latter control shows that the ratios of the various SV40 late RNA species are similar in singly and mixedly infected cells. M, Molecular weight markers; P, full-length probe (one-half the amount used in each hybridization reaction).

the peak of the polysome profile to smaller sized polysomes, one would still have expected to see a fairly normal distribution with only one optimum and a width not much greater than was observed for the RNA species lacking AUGs within their leaders. How, then, can one explain the extremely broad polysome distribution, with possibly three optima, that was seen with the 16S RNA species *a* of SV40? By assuming, together with the previously stated postulates of the scanning hypothesis, that (i) the rate-limiting step in translation of some mRNAs, including the SV40 late 16S mRNAs, is elongation rather than hopping onto the 5′ end or scanning, and (ii) when a scanning 40S ribosomal subunit encounters an actively translating ribosome it falls off the mRNA molecule, we can generate by computer simulation polysome profiles similar to the ones observed here, including three optima for RNA species *a* (J. E. Mertz and J. M. Kane, unpublished data). Furthermore, these additional assumptions lead to the prediction that even fairly weak 5′-proximal translation initiation signals can severely inhibit translation of downstream genes if not followed shortly thereafter by in-phase translation termination codons. This prediction can be tested fairly easily and may explain why the open reading frames of most nonfunctional upstream AUGs are very short (10a).

We thank Jeff Sklar and Paul Berg for the generous gift of the SV40 16S cDNA clone, Marilyn Kozak for discussions, Jeff Ross and Howard Temin for critical review of this manuscript, and the anonymous referees for insightful comments concerning this study.

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 Public Health Service training grant T32 CA-09135 from the National Cancer Institute.

LITERATURE CITED

1. **Barkan, A., and J. E. Mertz.** 1981. 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. J. Virol. **37**:730–737.
2. **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**:721–732.
3. **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.
4. **Ghosh, P. K., P. Roy, A. Barkan, J. E. Mertz, S. M. Weissman, and P. Lebowitz.** 1981. Unspliced functional late 19S mRNAs containing intervening sequences are produced by a late leader mutant of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **78**:1386–1390.
5. **Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo.** 1981. Translational initiation in prokaryotes. Annu. Rev. Microbiol. **35**:365–403.
6. **Jackson, V., and R. Chalkley.** 1981. Use of whole-cell fixation to visualize replicating and maturing simian virus 40: identification of a new viral gene product. Proc. Natl. Acad. Sci. U.S.A. **78**:6081–6085.
7. **Jay, G., S. Nomura, C. W. Anderson, and G. Khoury.** 1981. Identification of the SV40 agnogene product: a DNA binding protein. Nature (London) **291**:346–349.
8. **Kozak, M.** 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? Cell **15**:1109–1123.
9. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
10. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.

10a. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.

11. **McMullen, M. D., P. H. Shaw, and T. E. Martin.** 1979. Characterization of poly(A$^+$) RNA in free messenger ribonucleoprotein and polysomes of mouse taper ascites cells. J. Mol. Biol. **132**:679–694.
12. **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.
13. **Mertz, J. E., A. Murphy, and A. Barkan.** 1983. Mutants deleted in the agnogene of simian virus 40 define a new complementation group. J. Virol. **45**:36–46.
14. **Sherman, F., J. W. Stewart, and A. M. Schweingruber.** 1980. Mutants of yeast initiating translation of iso-1-cytochrome c within a region spanning 37 nucleotides. Cell **20**:215–222.
15. **Tooze, J. (ed.).** 1981. DNA tumor viruses. Molecular biology of tumor viruses, 2nd ed., part 2, revised. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.