Proc. Natl. Acad. Sci. USA
Vol. 76, No. 2, pp. 731–735, February 1979
Biochemistry

Rapid sequence determination of late simian virus 40 16S mRNA leader by using inhibitors of reverse transcriptase

(2′,3′-dideoxynucleoside triphosphates/splicing/complementary DNA/restriction fragment extension/ RNA-dependent DNA nucleotidyltransferase)

MINOU BINA-STEIN, MARILYN THOREN, NORMAN SALZMAN, AND JOHN A. THOMPSON

National Institute of Allergy and Infectious Diseases, Laboratory of Biology of Viruses, National Institutes of Health, Bethesda, Maryland 20014

Communicated by E. R. Stadtman, December 1, 1978

ABSTRACT **A method for the determination of the primary structure of spliced mRNA junction and leader sequences is described. By analogy to the DNA sequencing procedure of Sanger *et al.* [Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467], we use 2′,3′-dideoxynucleoside triphosphates as chain-terminating inhibitors of the reverse transcriptase (RNA-dependent DNA polymerase) reaction. By using specific DNA restriction fragments as primers in combination with this technique, we have determined the sequence of the spliced junction between the body and the leader sequence of the 16S late mRNA of simian virus 40. The method described should be of general utility in mapping spliced mRNA regions for which the corresponding protein sequence (if any) is unknown.**

Recent discoveries that many eukaryotic and viral mRNAs result from the excision of segments from primary transcripts and subsequent splicing (for reviews see ref. 1–3) suggest that a detailed knowledge of mRNA structure is necessary for an understanding of gene regulation. A number of investigators have analyzed the primary structure of mRNA by sequence determination of the complementary DNA (cDNA) (4–8). In those studies, the cDNA synthesis has been primed with synthetic oligonucleotides and the cDNAs have been sequenced either by Maxam and Gilbert's chemical method (7–9) or by Sanger and Coulson's plus and minus method (4–6, 10). We have approached the problem of sequencing desired regions of selected mRNAs by using given DNA restriction fragments, which anneal to the mRNA at specific regions, as primers for the reverse transcriptase (RNA-dependent DNA polymerase) reaction. Addition of dideoxynucleoside triphosphates as chain terminators (11) during the reverse transcriptase reaction allows rapid sequence determination of the cDNA. Using this method, we have studied the major late cytoplasmic 16S mRNA of simian virus 40 (SV40), the message that codes for the structural coat protein VP1 (12). It had been previously shown that VP1 mRNA is transcribed from two noncontiguous regions in the genome; the 5′ end of the mRNA is from position 0.72 to 0.76, while the body of the mRNA is transcribed from the region between 0.938 and 0.17 on the viral genome (8, 13–16) (Fig. 1*A*). We have confirmed the results showing that the 5′ leader sequence is about 200 nucleotides long (8). We find that the junction of the leader with the main body of the 16S mRNA precedes the initiation codon by 44 nucleotides, and we suggest that a subset of mRNAs with shorter leader sequences may exist.

MATERIALS AND METHODS

Preparation of SV40 DNA Fragments. Confluent monolayers of BSC-1 African green monkey cells were infected with SV40 (776) at a multiplicity of 0.1 plaque-forming unit per cell. SV40 DNA was isolated from virions (17) and further purified by isopycnic banding in cesium chloride and ethidium bromide (18). Restriction fragments m1, m2, m5, and T were obtained from the digestion of superhelical SV40 DNA with the restriction enzyme *Alu* I (New England BioLabs). The fragments were fractionated by RPC5 (gift from Bruce Roe) column chromatography (19–21). Digestion of the *Alu* I fragment P with *Hae* III followed by fractionation on RPC5 yielded the fragment designated PS (Fig. 1*B*).

The restriction fragments were dephosphorylated with alkaline phosphatase (Worthington) and 5′ end-labeled with [γ-^{32}P]ATP [specific activity, 3000 Ci/mmol (1 Ci = 3.7 × 10^{10} becquerels), New England Nuclear] and polynucleotide kinase (New England BioLabs) as described by Maxam and Gilbert (9).

Isolation of Late Cytoplasmic mRNA. Subconfluent (80–90%) monolayers of BSC-1 cells were infected with SV40 (5–10 plaque-forming units/cell) and labeled with [^{3}H]uridine. The cytoplasmic mRNA was isolated 48 hr after infection and extracted with phenol/chloroform (1:1, wt/wt) and 1% sodium dodecyl sulfate. DNA was removed by extraction with guanidine hydrochloride and the poly(A)-containing mRNAs were selected on poly(U)-Sepharose 4B (Pharmacia) (22).

Sequencing Procedure. Samples (0.03 μg) of m2 or PS DNA restriction fragments were hybridized to 100 μg of total poly(A)-containing late mRNA [in 0.5% sodium dodecyl sulfate/57% formamide (wt/vol)/0.4 M NaCl/0.02 M Tris·HCl, pH 7.5] (23) at 40°C for 24–48 hr and then precipitated with ethanol. The m2·RNA hybrids were dissolved in 40 μl of 0.01 M Tris·HCl, pH 7.5, with 0.25 mCi each of [α-^{32}P]dCTP and [α-^{32}P]dTTP (New England Nuclear, 300 Ci/mmol). The PS·RNA hybrids were dissolved in 0.01 M Tris·HCl, pH 7.5, with 0.25 mCi of [α-^{32}P]dCTP. The solutions were made 0.05 M in Tris·HCl, pH 8.3, 0.1 M in KCl, 0.01 M in $MgCl_2$, and 0.01 M in 2-mercaptoethanol. One hundred units of reverse transcriptase (J. W. Beard, Life Science, St. Petersburg, FL) was added and the reaction mixture was incubated for 30 sec at 40°C. Subsequently, equal aliquots were removed and added to tubes containing all four deoxynucleoside triphosphates (dNTPs) and one of the dideoxynucleoside triphosphates (ddNTPs) (Collaborative Research, Waltham, MA).

The concentrations of the deoxynucleotides in the final reaction mix (70 μl) were 28.5 μM dATP, 57 μM dGTP, 57 μM dCTP, and 57 μM dTTP. The final concentrations of the dideoxynucleotides varied, depending on the nucleotide: 43 μM ddATP, 14.2 μM ddGTP, 14.2 μM ddCTP, 14.2 μM ddTTP. At these nucleotide ratios (dNTP/ddNTP), the frequency with which the ddNTPs are incorporated is such that chain termination occurs at every point where the corresponding dNTP would have been incorporated; however, the ddNTP concentration is sufficiently low so that full length cDNA products are

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; dd-, 2′,3′-dideoxy-.

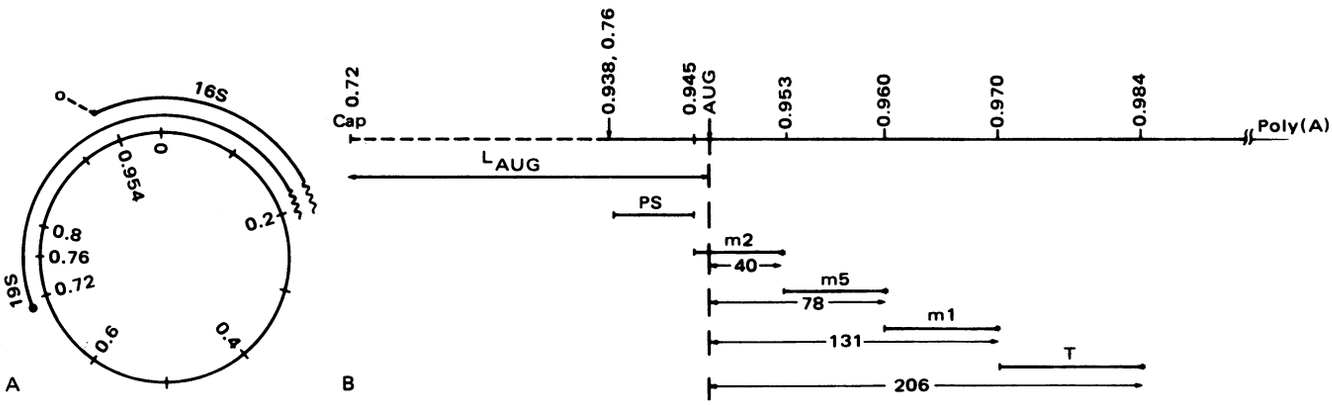

FIG. 1. (*A*) Map of the circular SV40 genome, indicating the segments of the genome that are templates for late 16S and 19S mRNAs. A leader sequence (0.72–0.76) is spliced to the main body of the late 16S mRNA (0.938–0.17). (*B*) Diagram of the 5′ end of the late 16S mRNA. The map positions of the SV40 DNA fragments used to prime the reverse transcriptase reaction are shown on the mRNA. The 5′ ends of m2, m5, m1, and T are 40, 78, 131, and 206 nucleotides, respectively, away from the VP1 initiation codon (AUG). L_{AUG} denotes the distance between the cap structure and the initiation codon.

also synthesized in the same reaction mixture. The reactions were incubated for 15 min at 40°C. Subsequently, 20 μl of a solution containing 250 μM of all four deoxynucleotides in the reaction buffer was added, and the incubation was continued for an additional 10 min. The reactions were terminated by addition of 10 μl of 1 M sodium hydroxide and then heated for 5 min at 100°C. The reaction mixtures were freed from unreacted radiolabeled nucleotides by fractionation on DEAE-cellulose (DE-52, Whatman) columns and were subsequently precipitated with ethanol. Samples were dissolved in 7 M urea and 0.25 mM EDTA and loaded onto standard DNA sequencing gels (9, 11).

RESULTS

Fig. 1*B* summarizes the principle of the approach. Four specific DNA restriction fragments—m2, m5, m1, and T—which occur successively from 0.945 to 0.984 on the SV40 genome and correspond to the main body of the 16S mRNA, are hybridized to late poly(A)-containing SV40 cytoplasmic mRNA. The m2 fragment contains the VP1 initiation codon. With reverse transcriptase we extend the primers and synthesize a cDNA copy of the 5′ end of 16S mRNA. From the length of the cDNA product, we estimate the distance between the 5′ end of the mRNA with respect to the initiation codon. The nucleotide sequence of the cDNA is obtained by analyzing the products when specific dideoxynucleoside triphosphates are added to the reverse transcriptase reactions. A comparison of the nucleotide sequence of the cDNA with that of the SV40 genome (24, 25) allows one to locate the spliced junction and to map the position of the leader sequence on the viral genome.

Determination of the Length of the 5′ End of mRNA with Respect to the Initiation Codon. The four ^{32}P end-labeled *Alu* I restriction fragments m2, m1, m5, and T were annealed with total poly(A)-containing cytoplasmic mRNA under conditions to maximize formation of RNA·DNA hybrids (23). The hybrids were fractionated by sedimentation in isokinetic sucrose gradients (26). Fig. 2 shows that the bulk of the ^{32}P-labeled fragments sediment at 16 S, the position of the major species of the late mRNAs. Some minor peaks sedimenting at 12 S, 19 S, and 22 S are also observed. The noncomplementary DNA strands as well as any unhybridized DNA sediment at 5 S (Fig. 2). The DNA fragments hybridized to the 16S mRNA were used to prime the reverse transcriptase. The autoradiograph in Fig. 3 shows that each specific DNA primer yields one major band and many minor bands. For each DNA primer (m2, m1, m5, and T), the major cDNA comprises 50% and the sum of minor species, each of which varies between 2% and 15%, comprises 41% of the total cDNA products (Table 1). The percentage and the size of the individual cDNA products were reproducible and did not depend on the mRNA preparations. The distance of the 5′ end of the mRNA from the initiation codon was calculated from the length of the major cDNAs. Fragments m2, m5, m1, and T (49, 38, 53, and 75 nucleotides) were extended by reverse transcriptase to form cDNA copies of lengths 275, 317, 383, and 446 nucleotides, respectively. Subtraction of the known length between the 5′ end of each cDNA and the initiation codon from the total length of that molecule (Fig. 1) yields 238, 242, 255, and 241 nucleotides for m2, m5, m1, and T, respectively, as the distance between the initiation codon and the 3′ end of the cDNA. Thus, the distance from the 5′ end of the 16S mRNA to the initiation codon is about 244 ± 6 nucleotides.

Determination of the Nucleotide Sequence of the Spliced Junction. Sanger *et al.* (11) have recently developed a rapid DNA sequencing technique that makes use of 2′,3′-dideoxynucleoside triphosphates as specific chain terminators for DNA

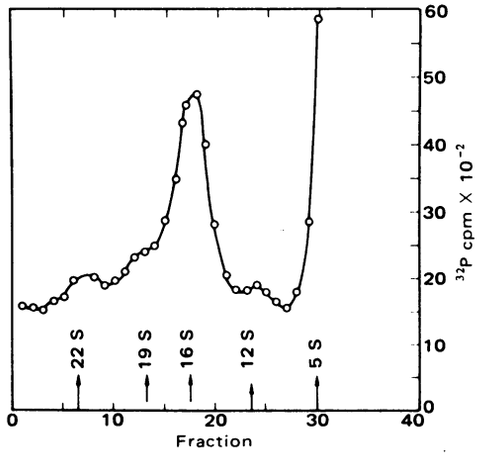

FIG. 2. Sucrose gradient fractionation of SV40 mRNA·DNA hybrids. Late poly(A)-containing cytoplasmic mRNA was hybridized to ^{32}P end-labeled restriction fragments. The hybridization mixture was layered onto isokinetic sucrose gradients, in 0.5% sodium dodecyl sulfate/0.1 M NaCl/1 mM EDTA/0.01 M Tris·HCl, pH 7.5. The sucrose concentrations in the mixing and reservoir chambers were 15% and 35.9%, respectively (26). Centrifugation was for 16 hr at 32,000 rpm at 20°C in a Spinco SW 41 rotor. Aliquots from each fraction were assayed for ^{32}P. Tritium-labeled 18S and 28S ribosomal RNAs were used as sedimentation markers in a parallel gradient. The arrows indicate the calculated S values.

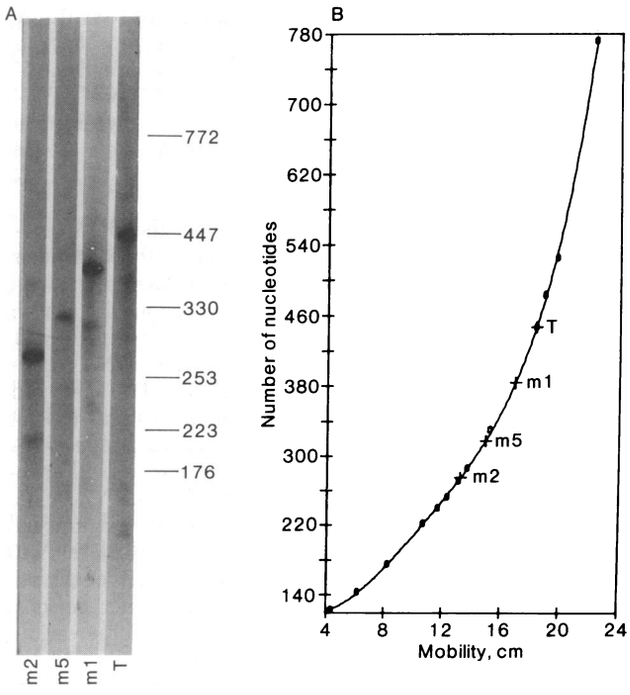

FIG. 3. (*A*) Autoradiograph of ^{32}P end-labeled cDNA electrophoresed on 3% polyacrylamide gel in 7 M urea. DNA restriction fragments m2, m5, m1, and T were hybridized to late cytoplasmic poly(A)-containing mRNA and fractionated on sucrose gradients (Fig. 2). The DNA hybridized to the 16S mRNA was used to prime the reverse transcriptase reaction. The specific fragments used as primers are indicated beneath the gels, which show the cDNA products that are formed. SV40 DNA *Alu* I and *Hin*dIII restriction fragments were used as size standards. The sizes of the markers were determined from the nucleotide sequence of the SV40 genone (24, 25). (*B*) cDNA was sized by computer analysis (MLAB–PDP 10 interactive system) (27–30). The mobilities of DNA markers (o) with respect to a fixed line at the bottom of the gel were fitted to a polynomial of the form $N(x) = \Sigma_{i=1}^{4} A_i x^i$, in which $N(x)$ denotes the length of DNA, N nucleotides long with mobility x, and A_i is an adjustable parameter. The mobility of the major cDNA product for each fragment, with respect to the same fixed line, is shown as + on the theoretical curve.

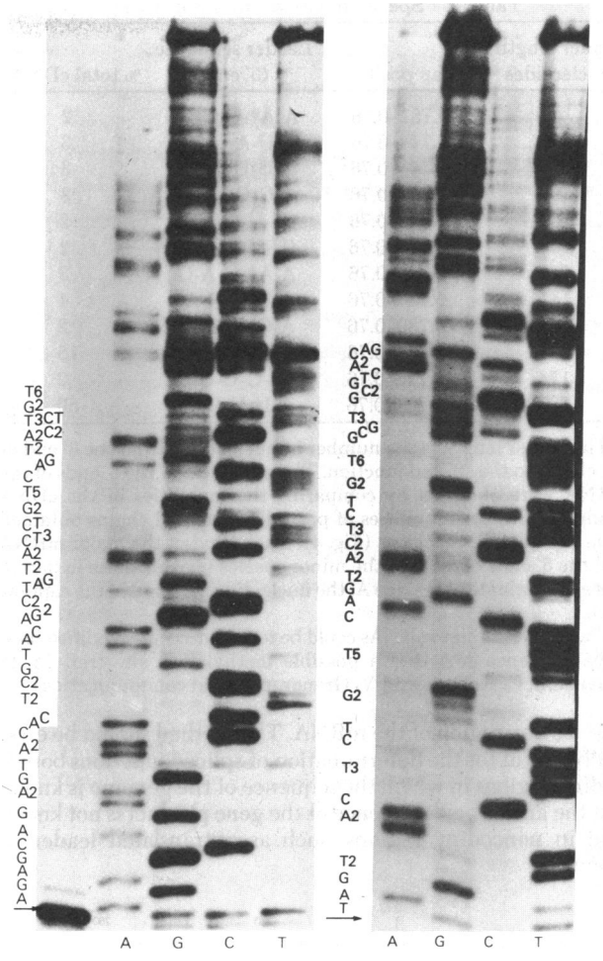

FIG. 4. Autoradiograph of the 15% polyacrylamide gels used for sequence determination of the cDNA with restriction fragments m2 (*A*) and PS (*B*) as primers. A, G, C, and T denote the specific ddNTP inhibitor used in the reaction.

polymerase. We have developed a variation of this method that allows equally rapid cDNA sequencing; thus, we use ddATP, ddGTP, ddCTP, and ddTTP to effect chain termination at A, G, C, and T residues during reverse transcription of an mRNA·restriction fragment primer hybrid. Two different DNA restriction fragments, m2 (residues 1413–1461) and PS (residues 1382–1412), were chosen for primers for the sequence analysis (Figs. 4 and 5).

Complementary DNA synthesis by reverse transcriptase is very sensitive to the concentrations of dNTPs. At limiting concentrations of the nucleotides, nonspecific chain termination occurs (33). The use of high concentrations of dNTP, however, lowers the specific activity of [α-^{32}P]dNTPs. To circumvent this problem, we have briefly pulsed the reverse transcriptase/DNA·RNA hybrid reaction with one or two specific α-^{32}P-labeled nucleoside triphosphates prior to the addition of the unlabeled deoxynucleoside triphosphates and a given chain terminator. The pulsing of m2·RNA hybrids with [α-^{32}P]dCTP and [α-^{32}P]dTTP results in the incorporation of five labeled nucleotides C-T-T-T-T (residues 1412–1407) at the 3′ terminus of the primer (Fig. 5). Therefore, the first nucleotide read from the autoradiogram (Fig. 4*A*) should be an adenine residue. The sequence obtained from this adenine (residue 1406) through residue 1379 corresponds to the expected contiguous region of the viral genome. Following residue 1379, the sequence corresponding to the leader appears (residue 443 uridine). The nucleotide sequence of the junction was further checked by using the PS fragment for primer. In this case, the first nucleotide detected on the autoradiograph (residue 1379) is identical to the nucleotide corresponding to the body of the 16S mRNA (Fig. 5); the subsequent nucleotides correspond to the leader sequence. Therefore, the nucleotide sequence of the spliced junction is (G-U-U-A-A-C-U|A-G-G-C-C-U-G). This result is in close agreement with the sequence (G-U-U-A-A-C-U|G-G-C-C-U-G) reported by Ghosh *et al.* (8).

DISCUSSION

The application of the dideoxynucleoside triphosphate method for rapid cDNA sequencing has two major advantages. First, because the DNA fragments hybridized to the mRNA are not ^{32}P end-labeled, unhybridized and noncomplementary DNA strands do not interfere with the sequence analysis. Second, because cDNA of very high specific activity is synthesized by incorporation of α-^{32}P-labeled nucleotides, only small amounts of the hybrid are needed.

A problem initially encountered with this technique was that low concentrations of deoxynucleoside triphosphates resulted in artifactual bands. We corrected this difficulty by increasing the concentrations of deoxynucleoside triphosphates and by the addition of a pulse step. With the dideoxy procedure, using DNA restriction fragments for primers, we can rapidly se-

Table 1. Species of late SV40 16S mRNA

Leader length, nucleotides	Map position	Leader sequence (5′ end)	% total cDNA
46	0.751–0.76	A-A*-C-G-C	2
61	0.748–0.76	A-U-A-C-U	.3
79	0.744–0.76	A-U-U-U-U	3
91	0.742–0.76	A-G-C-U-U	3
100	0.740–0.76	(U-U-G-U-U)	2
109	0.738–0.76	(U-U-U-U-U)	2
122	0.736–0.76	A-A*-C-A-G	3
141	0.733–0.76	A-U-G-G-A	4
154	0.730–0.76	A-G-G-U-U	2
166	0.728–0.76	A-G-G-C-C	15
178	0.725–0.76	(G-G-C-U-G)	2
201	0.721–0.76	A-U-U-U-C	50

The leader length is the number of nucleotides from the 3′ end of the cDNA to the spliced junction. The sequences of the 5′ ends of the mRNA were obtained by comparing the mobilities of the cDNA products with the mobilities of partially extended chains obtained from the ddNTP reactions (Fig. 4). The 5′ end of the major mRNA and the 5′ ends of most of the minor species (except those in parentheses) terminate at residue A, the nucleotide involved in the capping (31).

* The cDNA of these mRNAs could be terminating at a modified base; the sequence A-A-C is a possible methylation site (ref. 32; D. Canaani, C. Kahana and Y. Groner, personal communication).

quence any region of the mRNA. The method should be especially useful for the determination of spliced junctions both in coding regions in which the sequence of the genome is known but the amino acid sequence of the gene product is not known and in noncoding regions, such as untranslated leader sequences.

Comparison of the nucleotide sequence of SV40 16S mRNA with the sequence of the viral genome (24, 25) suggests that the nucleotide sequence around the 3′ end of the leader and the 5′ end of the body of the 16S mRNA are C-T-G-G and C-T-A-G-G, respectively (Fig. 5). The partial homology between these sequences precludes predicting which nucleotides are involved in the splicing reaction to generate the sequence C-U-A-G-G of the mRNA.

The length and the nucleotide sequence of the DNA complementary to the 5′ end of 16S mRNA indicate that the major leader is 200 nucleotides long and is joined to the body of the mRNA 44 nucleotides prior to the initiation codon. The length of the leader is in close agreement with the size obtained by S1 mapping technique (34, 35), suggesting that the cDNA represents a full copy of the 5′ end of the mRNA. The observed heterogeneity in the lengths of the cDNA products (Fig. 3) did not depend on either different RNA preparations or the DNA primers. Therefore, this heterogeneous population could represent either the existence of subspecies of 16S mRNA or the premature termination of cDNA synthesis due to unusual structures within the mRNA. Heterogeneity of the late SV40 mRNAs has been observed by direct 5′ end analysis of the purified mRNAs (D. Canaani, C. Kahana, and Y. Groner, personal communication). Comparison of the mobilities of the various cDNA products, synthesized with a given primer, with the mobilities of partially extended chains, obtained from the ddNTP reactions (data not shown), indicates that a number of the cDNA products terminated at a T residue (Table 1). Because adenine is the nucleotide involved in the cap structure (refs. 31 and 36; D. Canaani, C. Kahana and Y. Groner, personal communication), this suggests that some of the minor cDNA products could represent a subclass of 16S mRNAs with leaders shorter than the main leader.

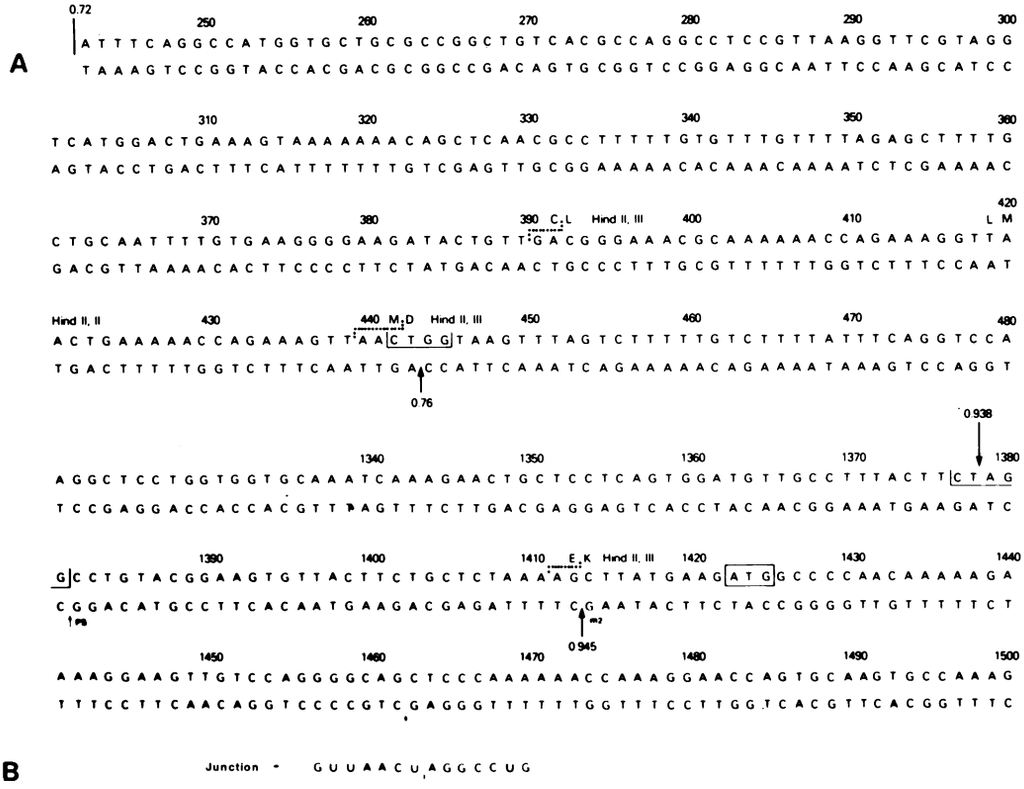

FIG. 5. (*A*) The nucleotide sequence of the major leader segment (nucleotides 243–443) and partial nucleotide sequence of the coding region of 16S mRNA (nucleotides 1378–1500) (24, 25). The length of the leader was determined by the total cDNA sizing technique (Fig. 3). The nucleotide sequence around the 3′ end of the leader and the 5′ end of the body of 16S mRNA is underlined. The splicing occurs 44 nucleotides from the VP1 initiation codon (box at 1423–1425). The numbering of the nucleotides is the same as in ref. 24. (*B*) The nucleotide sequence of the 16S mRNA at the junction of the splice. The sequence of the junction was determined by copying the mRNA into cDNA by reverse transcriptase in the presence of ddNTPs.

Even though the major leader of SV40 16S mRNA contains both initiation and termination codons, there is no evidence that this sequence is translated, and the function of the leader remains unknown. The conserved sequences of the 3′ end of 18S ribosomal RNAs, which have been implicated in the ribosome binding sites (37), are not complementary to any regions of the leader. However, the leader contains the hexanucleotide C-U-U-U-U-G also found in the 5′ noncoding region of eukaryotic mRNAs (38). Therefore, it seems plausible that the leader contains the information required for the initiation of protein synthesis.

Note Added in Proof. After the present manuscript was accepted, Reddy *et al.* (39) reported heterogeneity in the 16S mRNA leader. The 16S mRNA leaders 46, 61, 11, 100, and 201 nucleotides in length (Table 1) are colinear with SV40 DNA and correspond to the SV40 mRNAs designated 13, 12, 11, 10, and 6, respectively, by Reddy *et al.* The mRNAs with leaders 122 and 154 nucleotides in length (Table 1) are shown to be composed of leaders with gaps (nos. 9 and 7).

The application of the dideoxynucleotide method for RNA sequencing has been independently reported by Zimmern and Kaesberg (40), Hamlyn *et al.* (41), and McGeoch and Turnbull (42).

We greatfully acknowledge helpful discussions with Dr. R. Deely, Dr. M. Shani, and Mr. M. Radonovich. We thank Dr. A. Stein and R. T. Simpson for critical review of the manuscript.

1. Marx, J. L. (1977) *Science* **199,** 517–518.
2. Sambrook, J. (1977) *Nature (London)* **268,** 101–104.
3. Gilbert, W. (1978) *Nature (London)* **271,** 501.
4. Cheng, C. C., Brownlee, G. G., Carey, N. H., Dock, N. T., Gillam, S. & Smith, M. (1976) *J. Mol. Biol.* **107,** 522–547.
5. Baralle, F. E. (1977) *Cell* **10,** 549–558.
6. Baralle, F. E. (1977) *Nature (London)* **267,** 279–288.
7. Baralle, F. E. (1977) *Cell* **12,** 1085–1095.
8. Ghosh, P. K., Reddy, B., Swinscoe, J., Choudans, P. V., Lebowitz, P. & Weissman, S. W. (1978) *J. Biol. Chem.* **253,** 3643–3647.
9. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
10. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94,** 441–448.
11. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
12. Prives, C., Aviv, H., Gulboa, E., Revel, M. & Winocour, F. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **39,** 309–316.
13. Celma, M. L., Dahr, R., Pan, J. & Weissman, S. (1977) *Nucleic Acids Res.* **4,** 2549–2559.
14. Aloni, Y., Dahr, R., Laub, O., Horowitz, M. & Khoury, G. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 3686–3690.
15. Lavi, S. & Groner, Y. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5323–5327.
16. Haegman, G. & Fiers, W. (1978) *Nature (London)* **273,** 70–73.
17. Crawford, L. A. & Black, P. H. (1964) *Virology* **24,** 388–392.
18. Sebring, E. D., Kelly, T. J., Jr., Thoren, M. & Salzman, N. P. (1971) *J. Virol.* **8,** 487–490.
19. Hardis, S. C. & Wells, R. D. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 3117–3121.
20. Landy, A., Foeller, C., Rezelback, R. & Dudock, B. (1976) *Nucleic Acids Res.* **3,** 2575–2592.
21. Eshaghpour, H. & Crothers, D. M. (1978) *Nucleic Acids Res.* **5,** 13–21.
22. Deely, R. G., Gordon, J. I., Burns, A. T., Mullinex, K. P., Bina-Stein, M. & Goldberger, R. F. (1977) *J. Biol. Chem.* **252,** 8310–8319.
23. Casey, J. & Davidson, N. (1977) *Nucleic Acids Res.* **41,** 1539–1552.
24. Reddy, V. B., Thimmappaya, B., Dahr, R., Subramanian, K. N., Zain, S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978) *Science* **200,** 494–502.
25. Fiers, W., Contreras, R., Haegman, G., Rogier, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (London)* **273,** 113–120.
26. McCarthy, K. S., Jr., Vollmer, R. T. & McCarthy, K. S. (1974) *Anal. Biochem.* **61,** 165–183.
27. Knott, G. D. & Reece, D. K. (1972) in *Proceedings of the ON-LINE, 1972,* International Conference, Brunel University, England (On Line Computer System, Ltd.), Vol. 1, pp. 497–526.
28. Knott, G. D. & Schrager, R. I. (1972) in *Proceedings of the SIGGRAPH Computers in Medicine* (Association for Computing Machinery, New York), Vol. 6, pp. 138–151.
29. Bina-Stein, M. & Stein, A. (1976) *Biochemistry* **15,** 3912–3916.
30. Bina-Stein, M. (1978) *J. Biol. Chem.* **253,** 5213–5219.
31. Groner, Y., Carmi, P. & Aloni, Y. (1977) *Nucleic Acids Res.* **4,** 3959–3968.
32. Wei, C. M. & Moss, B. (1972) *Biochemistry* **16,** 1672–1676.
33. Efstratiades, A., Maniatis, T., Kafatos, F. C., Jeffrey, A. & Vournakis, J. N. (1975) *Cell* **4,** 367–378.
34. Berk, A. J. & Sharp, P. A. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 1274–1278.
35. Lai, C.-J., Dahr, R. & Khoury, G. (1978) *Cell* **14,** 971–982.
36. Haegman, G. & Fiers, W. (1978) *Nucleic Acids Res.* **5,** 2359–2371.
37. Hagenbuchle, O., Santer, M. & Steitz, J. A. (1977) *Cell* **13,** 551–563.
38. Baralle, F. E., Brownlee, G. G. (1978) *Nature (London)* **274,** 84–87.
39. Reddy, V. B., Ghosh, P. K., Lebowitz, P. & Weissman, S. M. (1978) *Nucleic Acids Res.* **5,** 4195–4213.
40. Zimmern, D. & Kaesberg, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4257–4261.
41. Hamlyn, P. H., Brownlee, G. C., Cheng, C.-C., Gait, M. J. & Milstein, C. (1978) *Cell* **15,** 1067–1075.
42. McGeoch, D. J. & Turnbull, N. T. (1978) *Nucleic Acids Res.* **5,** 4007–4024.