Efficient and Accurate In Vitro Processing of Simian Virus 40-Associated Small RNA

NISSIM HAY,¹ ORNA AMSTER-CHODER,¹ AND YOSEF ALONI^{1,2*}

Department of Genetics, Weizmann Institute of Science, Rehovot 76100, Israel,^{1} and Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544²*

Received 6 March 1985/Accepted 3 September 1985

Nuclei were isolated from simian virus 40 (SV40)-infected cells with a hypotonic, detergent-free buffer and incubated in vitro in a high-ionic-strength buffer containing [α -³²P]UTP. The labeled viral RNAs produced were analyzed by gel electrophoresis together with 3-h-labeled viral RNAs extracted from SV40-infected cells. The in vitro-synthesized RNA contained a major RNA species of 62 to 64 nucleotides that appeared on the gel at the same position as in vivo-synthesized SV40-associated small RNA (SAS-RNA). Analyses of the in vitro-synthesized 62- to 64-nucleotide RNA by hybridization to restriction fragments and by the use of an SAS-RNA deletion mutant clearly identified it as SAS-RNA. The intensity of the band of the in vitro-synthesized SAS-RNA increased with an increase in the labeling time or when a short pulse was followed by a chase. Moreover, the SAS-RNA band disappeared when ITP replaced GTP in the transcription reaction mixture. These results indicate that SAS-RNA is processed from a precursor molecule and that an RNA secondary structure could be an element recognized by the processing enzyme.

After simian virus 40 (SV40) infection, RNA polymerase II transcribes both SV40 mRNAs and 62- to 64-nucleotide (nt) RNA, known as SV40-associated small RNA (SAS-RNA) (6-8). It was suggested that SAS-RNA is specifically processed from the noncoding sequence in long RNA transcripts which are initiated at normal late promoters (6). The SAS-RNA-coding region is located about 170 nt's downstream from the polyadenylation site (0.17 map unit) of the late mRNAs (see Fig. 1A for its map location). Thus, the processing reactions which the SV40 primary late RNA transcripts undergo yield two RNA species: (i) mature poly(A)⁺ mRNAs and (ii) SAS-RNA. SAS-RNA has unique uncapped 5' and poly(A)⁻ 3' ends (6), so its processing reaction must therefore be highly specific.

To directly address the questions of whether SAS-RNA is a processing product and, if so, how the processing enzymes recognize the SV40 primary late RNA substrates and accomplish such precise maturation, an in vitro processing system must be developed.

For this purpose, we prepared an isolated-nucleus system from SV40-infected cells. In vitro incubation of the nuclei in a low-ionic-strength buffer leads to the premature termination of SV40 late transcripts and to the synthesis of a 94- to 98-nt attenuator RNA (see Fig. 1A for its map location) (1, 4, 5, 13-15, 20, 22, 23); incubation in a high-ionic-strength buffer leads to the synthesis of long viral RNA transcripts (5, 13-15, 20).

In the present communication, we show that when the nuclei are incubated in a high-ionic-strength buffer for a long time or when they are pulse-labeled with [α -³²P]UTP and the label is then chased with a high concentration of UTP, efficient and accurate processing of SAS-RNA can occur.

BSC-1 cells were infected with SV40 (30 to 50 PFU per cell) as previously described (18). At 42 to 48 h postinfection, the cells were washed and collected in cold hypotonic buffer (50 mM Tris hydrochloride [pH 7.9], 1.5 mM MgCl₂, 1 mM dithiothreitol) by centrifugation for 1 min at 1,000 $\times$ g. Nuclei were isolated by suspending 5 $\times$ 10⁷ cells in 10 ml of

hypotonic buffer (17) and pipetting them up and down 10 times with a Pasteur pipette, followed by centrifugation for 2 min at 1,000 $\times$ g. This step was repeated twice. The nuclei (approximately 3 $\times$ 10⁷) were suspended in a transcription reaction mixture containing 5 mM KCl, 1.5 mM MnCl₂, 1 mM dithiothreitol, 12.5% glycerol, 100 mM (NH₄)₂SO₄, 30 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES)-NaOH (pH 7.9), a 400 μ M concentration each of ATP, GTP, and CTP, 5 to 20 μ M UTP (as indicated in the figure legends), and 300 to 500 μ Ci of [α -³²P]UTP (400 mCi/mmol; Radiochemical Centre, Amersham, England) in a final volume of 0.6 ml. Transcription was carried out at 30°C for various times, as indicated in the figure legends, and stopped by the addition of 100 μ g of RNase-free DNase (Worthington Diagnostics, Freehold, N.J.) per ml for 1 min at 30°C, followed by extraction with phenol-chloroform (19). The aqueous fraction was passed through a Sephadex G-25 syringe column to remove free labeled nucleotides. After ethanol precipitation, the precipitate was suspended in TKM (25 mM KCl, 2.5 mM MgCl₂, 50 mM Tris hydrochloride [pH 6.7]), 200 μ g of RNase-free DNase per ml was added, and incubation was carried out for 1 h at 4°C. The DNase treatment was stopped by phenol-chloroform extraction, and the RNA was precipitated with ethanol. SV40-specific RNA was isolated by hybridization to and elution from SV40 DNA on nitrocellulose filters. Hybridization was carried out in 70% formamide—0.3 M NaCl—10 mM HEPES—NaOH (pH 7.5)—1 mM EDTA—0.1% sodium dodecyl sulfate at 37°C for 40 to 48 h. At the end of the incubation time, the filters were washed once with 0.5 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate), incubated in a solution of 35% formamide—0.01 M NaCl—1 mM EDTA—10 mM HEPES—NaOH (pH 7.5) at 37°C for 30 min and then washed three times with 0.5 $\times$ SSC. The bound RNA was eluted by incubating the filter in a solution of 90% formamide—10 mM HEPES—NaOH (pH 7.5) for 1 h at 37°C. For the preparation of in vivo-labeled RNA, infected cells were labeled for 3 h with ³²P_i (5 mCi/10⁷ cells) in phosphate-free medium. Nuclear and cytoplasmic fractions were prepared by using the hypotonic buffer procedure described above. ³²P-labeled

* Corresponding author.

FIG. 1. Genomic map of SV40 (A) and identification of SAS-RNA produced in isolated nuclei by size analysis (B). (A) The map indicates the positions of the primary late RNA transcript (we have suggested that it terminates after transcription around the genome at the attenuation site at 0.74 map unit [13]) and the spliced early and late mRNAs (inner arrows). The map position of attenuator RNA (Att. RNA) is indicated between 0.72 and 0.74 map unit. The map position of SAS-RNA is indicated between residues 2844 and 2905 (25) in the polarity of the late mRNAs. a through e, Restriction fragments obtained by cleaving the viral DNA with *Hpa*I, *Eco*RI, and *Bgl*I restriction enzymes (3). Ori, origin of replication. (B) For in vivo experiments, viral RNA was extracted and purified from the cytoplasmic fraction of infected cells labeled for 3 h with 32 P. For in vitro experiments, nuclei isolated from infected cells were incubated in a transcription reaction mixture containing 300 μ Ci of $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ (400 mCi/mmol) and 6 μ M UTP for 6 min. Unlabeled UTP was then added to a final concentration of 400 μ M, and incubation at 30°C was continued for an additional 30 min. The labeled RNA produced was purified. The in vivo- and in vitro-synthesized RNAs were analyzed by gel electrophoresis. The arrowhead points to the position of SAS-RNA. m, Size markers obtained by labeling *Hpa*II restriction fragments of pBR322 DNA.

FIG. 2. Size analysis of viral RNAs produced in isolated nuclei infected with either the wild-type (WT) or *dl2194* strain of SV40 (A) and hybridization of the 62- to 64-nt RNA with restriction fragments of SV40 (B). (A) Nuclei isolated from cells infected with either the WT or *dl2194* strain of SV40 were incubated in a transcription reaction mixture containing 500 μ Ci of $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ (400 mCi/mmol) and 20 μ M UTP for 20 min at 30°C. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The open arrowhead points to the position of attenuator RNA, and the filled arrowhead points to the position of SAS-RNA. m, Size markers as in Fig. 1B. (B) The RNA at the position of SAS-RNA was eluted from the gel and hybridized to a blot containing five restriction fragments (as indicated in Fig. 1A).

viral RNA was extracted from the cytoplasmic fraction (19), and viral RNA was purified by the hybridization-elution procedure described above for nuclear RNA.

For analysis of the RNA by gel electrophoresis, the RNA was denatured in 5 μ l of 90% formamide—10 mM HEPES—NaOH (pH 7.5) for 1 min at 90°C and subjected to electrophoresis on an acrylamide gel (bis/acrylamide ratio, 1:29) containing 7 M urea in TBE. Electrophoresis was carried out at a constant current of 20 mA for 4 to 5 h.

Figure 1B shows that the RNA synthesized *in vitro* contained a major RNA species that appeared on the gel at the same position as the SAS-RNA synthesized *in vivo* (arrowhead) (6), suggesting accurate *in vitro* synthesis of SAS-RNA. It is worth noting that in comparison with *in vivo* SAS-RNA synthesis, SAS-RNA was synthesized *in vitro* at a high level. We noted the appearance of a minor band above SAS-RNA. This may constitute SAS-RNA molecules longer by a few nucleotides or a different form of SAS-RNA species.

To determine whether the 62- to 64-nt RNA synthesized *in vitro* was indeed SAS-RNA, we prepared nuclei from cells infected with a viable deletion mutant of SV40 (*dl*2194), in

FIG. 3. Size analysis of viral RNAs produced in isolated nuclei under either pulse or pulse-chase conditions. Nuclei were incubated in a transcription reaction mixture containing 500 μ Ci of [α - 32 P]UTP (400 mCi/mmol) and 5 μ M unlabeled UTP for 5 min at 30°C (pulse). After the pulse, 400 μ M unlabeled UTP was added, and incubation was continued for 30 min (chase). The labeled RNAs produced were purified and analyzed by gel electrophoresis. p, Pulse; c, chase; m, size markers as in Fig. 1B. The open arrowhead points to the position of attenuator RNA, and the filled arrowhead points to the position of SAS-RNA.

FIG. 4. Size analysis of viral RNAs produced in isolated nuclei incubated in the presence of either GTP or ITP in the transcription reaction mixture. Lane GTP shows a standard transcription reaction mixture such as that described in the legend to Fig. 1B. In lane ITP, ITP (400 μ M) replaced GTP in the standard transcription reaction mixture. Transcription was allowed to proceed as described in the legend to Fig. 2. The labeled RNAs produced were purified and analyzed by gel electrophoresis. The arrowhead points to the position of the SAS-RNA. m, Size markers as in Fig. 1B.

which the coding region of SAS-RNA has been deleted (12), and compared the syntheses of SAS-RNA in mutant- and wild-type-infected cells. The absence of SAS-RNA in cells infected with *dl*2194 has been previously demonstrated (6).

Figure 2A, lane WT, shows the appearance of a major band of 62- to 64-nt RNA (filled arrowhead). Trace amounts of 94- to 98-nt attenuator RNA (open arrowhead) and two additional bands of 125- and 155-nt RNA are also recognizable, together with viral RNA of heterogeneous lengths. The appearance of the 125- and 155-nt RNA bands was not always reproducible. The 125-nt RNA was not SV40 specific; it was presumably trapped on the filter during hybridization (see also reference 6). Figure 2A, lane *dl*2194, shows that when nuclei from these cells were incubated *in vitro* under the same conditions as nuclei from wild-type-infected cells, no major band of 62- to 64-nt RNA appeared. Instead, viral RNA molecules which were longer than those in lane WT were observed. Attenuator RNA was not recognizable in the RNA preparation from *dl*2194-infected cells. However, because this observation was not reproducible and because the present conditions were not optimal for transcription termination at the attenuation site (13), it is

FIG. 5. Schematics of putative RNA secondary structures in the proximity of SAS-RNA of strains 776 (A) and VA4554 (B). Note that SAS-RNA forms bulges and that residues 11 to 20 of U₄ RNA (10) base pair with the 5' and 3' ends of SAS-RNA. Note also that the involvement of the U₄ ribonucleoprotein in processing-polyadenylation has been suggested (9). The sequences deleted in *d1263* and the alterations in sequences in VA4554 are indicated (8). Changes in free energy (ΔG) were calculated by the method of Tinoco et al. (24).

difficult to draw any conclusions concerning the role of SAS-RNA in the attenuation process. A hypothetical role for SAS-RNA in attenuation has previously been discussed (4). Figure 2B shows that when 62- to 64-nt RNA was eluted from the gel and hybridized to a blot of restriction fragments (as shown in Fig. 1A), it hybridized exclusively with fragment c, spanning 0.17 to 0.37 map unit. This fragment of the viral genome includes the SAS-RNA-coding region (6). These results strongly suggest that 62- to 64-nt RNA is the SV40-encoded SAS-RNA. Further evidence for the identification of 62- to 64-nt RNA as SAS-RNA came from an RNase T₁ fingerprint analysis, which yielded the expected pattern based on the known base sequence of the SAS-RNA-coding region (N. Hay, H. Skolnik-David, and Y. Aloni, unpublished results).

To determine whether SAS-RNA is produced *in vitro* by a specific cleavage mechanism or is rather an independent transcription product, the following experiments were performed. Isolated nuclei of SV40-infected cells were incubated *in vitro* in a high-ionic-strength buffer [100 mM (NH₄)₂SO₄] under either pulse or pulse-chase conditions. The labeled viral RNAs were then purified and analyzed by gel electrophoresis as in Fig. 1B. Figure 3, lane p, shows that during a 5-min pulse with [α -³²P]UTP, 94- to 98-nt attenuator RNA (open arrowhead) and long viral RNA molecules of heterogeneous lengths were synthesized. A faint band of SAS-RNA was also observed. The production of SAS-RNA

may indicate transcription from an independent promoter or processing of SAS-RNA from precursor nascent chains. Figure 3, lane c, shows that after a 30-min chase with unlabeled UTP, a different pattern of viral RNA species emerged. The 94- to 98-nt attenuator RNA almost disappeared, the average size of the viral RNA decreased, and a major band of 62- to 64-nt SAS-RNA appeared (filled arrowhead). Since the UMP residues of SAS-RNA were labeled only during the pulse and since the labeled SAS-RNA species accumulated after the chase, we conclude that SAS-RNA is produced *in vitro* from long precursor molecules by an efficient and accurate cleavage mechanism. The cleavage probably occurs when the precursor molecule is still nascent. In this case, the various small RNA molecules which appeared after the chase (see Fig. 3, lane c) may represent either (i) labeled viral RNA sequences with their 5' ends adjacent to the 3' end of SAS-RNA which were transcribed from a region of the viral genome located downstream from the SAS-RNA-coding region or (ii) degradation products. The observation that SAS-RNA accumulated after the chase is indicative of its remarkable stability in our isolated-nucleus system.

We have begun to use the above-described *in vitro* isolated-nucleus system for analyzing the elements which constitute the signal for the SAS-RNA-processing enzyme. Our first approach was to determine the importance of an RNA secondary structure. For this purpose, we replaced GTP

with its analog ITP in the standard transcription reaction mixture. In previous studies, ITP was shown to abolish transcription termination in bacteria and bacteriophages (2, 11) and at the SV40 attenuator (13) by destabilizing RNA stem-and-loop structures. Figure 4 shows the production of long viral RNA molecules in the two transcription reaction mixtures. However, SAS-RNA was produced in the transcription reaction mixture which contained GTP (arrowhead) but not in that which contained ITP. Based on these results, we suggest that an RNA secondary structure may constitute at least part of the signal for the endonucleolytic enzyme. Alternatively, the G residue could be an essential nucleotide in the sequence required for processing, or GTP may provide the energy source for the processing reaction.

Heuverswyn et al. (16), Alwine and Khoury (8), and Sadofsky and Alwine (21) have suggested that a putative secondary structure could exist in single-stranded DNA or in late mRNAs in the proximity of the SAS-RNA-coding region. The existence of this structure could explain the observations that strain *d1263*, which lacks sequences upstream of the SAS-RNA-coding region (see Fig. 5A for the sequences deleted), did not produce SAS-RNA, whereas strain VA4554, in which several sequences are inserted upstream of the SAS-RNA-coding region, leading to a less stable secondary structure, produced three to five times less SAS-RNA (8). We present here schematics of alternative secondary structures for the RNAs of the two wild-type strains 776 and VA4554 (Fig. 5A and B, respectively). In our proposed secondary structures, SAS-RNA forms bulges in stem-and-loop structures. In these structures, the cleavage sites seem to be in a position easily accessible to the processing enzymes. Furthermore, the conformational rearrangements which accompany the excision of SAS-RNA would result in a decrease in free energy from -14 kcal (1 cal = 4.184 J) to -21 kcal and from -5 kcal to -12 kcal for strains 776 and VA4554, respectively. The -7 -kcal decrease in free energy could provide all or a portion of the energy required for the endonucleolytic reaction.

We have noticed that residues 11 to 20 in U₄ RNA are complementary to sequences at the 5' and 3' ends of SAS-RNA and can therefore base pair with them (Fig. 5A and B). This base pairing can help stabilize the secondary structures (Fig. 5A and B), in which the 5' and 3' ends of SAS-RNA are held in close proximity. Cleavage could occur either by nuclease activity associated with the U₄ ribonucleoprotein or through an RNase directed to the proper form of the secondary structure.

It seems that a common feature of mRNA splicing, processing-polyadenylation, and processing of SAS-RNA is that the cleavage sites are primarily defined by one of a set of nucleotide sequences but that their actual function is modulated by the secondary and tertiary structures of the RNA and possibly by proteins associated with them. Our present success in reproducing the processing of SAS-RNA *in vitro* is a first step in providing information regarding the details of this mechanism and thereby of pre-mRNA processing.

We thank O. Resnekov for help in preparing cytoplasmic RNA, Ruchama Leiserowitz for technical assistance, and J. Feunteun for providing strain *d12194*. Yosef Aloni thanks A. J. Levine for his kind hospitality in allowing completion of this work in his laboratory.

This research was supported by U.S. Public Health Service research grant CA 14995 from the National Institutes of Health and in part by the MINERVA Foundation, Munich. Part of the work reported in this paper was undertaken during the tenure of an American Cancer Society-Eleanor Roosevelt-International Cancer

Fellowship awarded to Yosef Aloni by the International Union Against Cancer.

LITERATURE CITED

1. Abulafia, R., A. Ben-Ze'ev, N. Hay, and Y. Aloni. 1984. Control of late simian virus 40 transcription by the attenuation mechanism and transcriptionally active ternary complexes are associated with the nuclear matrix. *J. Mol. Biol.* **172**:467-487.
2. Adhya, S., and M. Gottesman. 1978. Control of transcription termination. *Annu. Rev. Biochem.* **47**:967-996.
3. 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. USA* **74**:3686-3690.
4. Aloni, Y., and N. Hay. 1983. Attenuation and modulation of mRNA secondary structure in a feedback control system regulating SV40 gene expression. *Mol. Biol. Rep.* **9**:91-100.
5. Aloni, Y., N. Hay, H. Skolnik-David, P. Pfeiffer, R. Abulafia, R. Pruzan, E. Ben-Asher, E. B. Jakobovits, O. Laub, and A. Ben-Ze'ev. 1983. Attenuation in the control of gene expression in animal viruses, p. 1-48. *In* A. Kohn and P. Fuchs (ed.), *Developments in molecular virology*, vol. 4. Martinus Nijhoff Publishers BV, Dordrecht, The Netherlands.
6. Alwine, J. C. 1982. Hybrid selection of small RNAs by using simian virus 40 DNA: evidence that the simian virus 40-associated small RNA is synthesized by specific cleavage from large viral transcripts. *J. Virol.* **43**:987-996.
7. Alwine, J. C., R. Dhar, and G. Khoury. 1980. A small RNA induced late in simian virus 40 lytic infection can associate with early viral mRNAs. *Proc. Natl. Acad. Sci. USA* **77**:1379-1383.
8. Alwine, J. C., and G. Khoury. 1980. Simian virus 40-associated small RNA: mapping on the simian virus 40 genome and characterization of its synthesis. *J. Virol.* **36**:701-708.
9. Berger, S. M. 1984. Are U4 small nuclear ribonucleoproteins involved in polyadenylation? *Nature (London)* **309**:179-182.
10. Busch, H., R. Reddy, L. Rothblum, and J. C. Choi. 1982. SnRNAs, SnRNPs and RNA processing. *Annu. Rev. Biochem.* **51**:617-654.
11. Farnham, P. J., and T. Platt. 1980. A model for transcription termination suggested by studies on the trp attenuator *in vitro* using base analogs. *Cell* **20**:739-748.
12. Feunteun, J., G. Carmichael, J. C. Nicolas, and M. Kress. 1981. Mutants carrying deletions in the two simian virus 40 early genes. *J. Virol.* **40**:625-634.
13. Hay, N., and Y. Aloni. 1984. Attenuation in SV40 as a mechanism of transcription termination by RNA polymerase B. *Nucleic Acids Res.* **12**:1401-1414.
14. Hay, N., and Y. Aloni. 1985. Attenuation of late simian virus 40 mRNA synthesis is enhanced by the agnoprotein and is temporally regulated in isolated nuclear systems. *Mol. Cell. Biol.* **5**:1327-1334.
15. Hay, N., H. Skolnik-David, and Y. Aloni. 1982. Attenuation in the control of SV40 gene expression. *Cell* **29**:183-193.
16. Heuverswyn, H. V., C. Cole, P. Berg, and W. Fiers. 1979. Nucleotide sequence analysis of two simian virus 40 mutants with deletions in the region coding for the carboxyl terminus of the T antigen. *J. Virol.* **30**:936-941.
17. Jakobovits, E. B., and Y. Aloni. 1980. Isolation and characterization of various forms of simian virus 40 DNA-protein complexes. *Virology* **102**:107-118.
18. Laub, O., and Y. Aloni. 1976. Transcription of simian virus 40. VI. SV40 DNA-RNA polymerase complex isolated from productively infected cells transcribed *in vitro*. *Virology* **75**:346-356.
19. Penman, S. 1966. RNA metabolism in the HeLa cell nucleus. *J. Mol. Biol.* **17**:117-130.
20. Pfeiffer, P., N. Hay, R. Pruzan, E. B. Jakobovits, and Y. Aloni. 1983. *In vitro* premature termination in SV40 late transcription. *EMBO J.* **2**:185-191.
21. Sadofsky, M., and J. C. Alwine. 1984. Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. *Mol. Cell. Biol.* **4**:1460-1468.

22. **Skolnik-David, H., and Y. Aloni.** 1983. Pausing of RNA polymerase molecules during *in vivo* transcription of the SV40 leader region. *EMBO J.* 2:179–184.
23. **Skolnik-David, H., N. Hay, and Y. Aloni.** 1982. Site of premature termination of late transcription of simian virus 40 DNA. Enhancement by 5,6-dichloro-1- β -ribofuranosylbenzimidazole. *Proc. Natl. Acad. Sci. USA* 79:2743–2747.
24. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. Uhlenbeck, D. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. *Nature (London) New Biol.* 246:40–41.
25. **Tooze, J.** 1981. Appendix A. The SV40 nucleotide sequence, p. 799–841. *In* J. Tooze (ed.), *DNA tumor viruses*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.