Polyoma infected cells contain at least three spliced late RNA s

Mia Horowitz, Susan Bratosin and Yosef Aloni

Department of Genetics, The Weizmann Institute of Science, Rehovot, Israel

Received 8 September 1978

ABSTRACT

Poly(A)-containing polyoma cytoplasmic RNA was hybridized with linear double-stranded polyoma DNA and RNA displacement loops (R-loops) were formed. The structures visualized in the electron microscope are consistent with the conclusion that there are at least three late polyoma specific RNAs and that the leader sequences at the 5' ends of these viral RNAs are not coded immediately adjacent to the bodies of the RNAs. Measurements carried out on the R-loop structures have provided the locations on the physical map of polyoma DNA, for the bodies and leaders of the RNAs and the length of the bodies, leaders and the corresponding intervening DNA sequences.

INTRODUCTION

The vast amount of work on the transcription of SV40 and polyoma DNAs in lytically infected cells has pointed to remarkable similarities between the two viruses. In both viruses the stable late cytoplasmic mRNA is usually detected as two distinct species with sedimentation coefficients of 16S and 19S (1,2,3).

A technique has been developed by Thomas et al. (4), in which, under appropriate conditions of hybridization, complementary RNA can hybridize to double-stranded DNA by displacing the part of the DNA strand identical to this RNA. These hybrid molecules (R-loops) can be visualized in the electron microscope.

Using this technique, it has recently been shown that in adeno and SV40 mRNAs, sequences close to the 5' end of the mRNA (called "leader sequences") are coded for by DNA which is not contiguous with the main coding sequences (6, 22, 7, 8, 5).

In the present study we have used the same technique in analyzing hybrids formed between linear polyoma DNA and polyoma

late poly(A)-containing cytoplasmic RNA. It is shown that polyoma virus has at least three spliced late RNAs. The map positions of the bodies and the leaders of the three RNAs have been determined.

MATERIALS AND METHODS

Virus and cells: A-31 mouse cells (obtained from Dr. A. Hakina, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan) were infected with large-plaque polyoma virus 78V (9). The cells were labeled at 30 hrs post-infection for times as indicated with 5,6[³H]-uridine (200 μ ci/6 ml of medium per culture, 40 ci/mmol: The Radiochemical Center, Amersham, England).

Preparation of cytoplasmic RNA: At the end of the labeling period the cultures were washed and cytoplasmic fraction was prepared by lysing the cells with phosphate-buffered saline containing 0.5% Nonidet-P40 (10). RNA was extracted with phenol-chloroform-isoamyl alcohol (10) at room temperature and collected by ethanol precipitation and centrifugation, treated with DNase, phenol extracted and precipitated with ethanol (11). The precipitate was resuspended in 0.5M NaCl, 0.01 M Tris-HCl(pH 7.4), 0.1% SDS and chromatographed on an oligo-(dT)-cellulose column (13). The poly(A)-containing RNA was collected and precipitated with ethanol.

Preparation of polyoma DNA: Polyoma DNA was extracted by the Hirt procedure (14) as previously described (11). Form I DNA was further purified by equilibrium centrifugation in a CsCl-EtBr density gradient (15). To obtain linear DNA, Form I DNA was cleaved with HaeII restriction endonuclease in 60 mM NaCl, 6 mM MgCl₂, 6 mM mercaptoethanol, 6 mM DTT at 37°C for 3 hr. The enzyme was extracted with phenol and the DNA was precipitated with ethanol.

Hybridization with double-stranded linear polyoma DNA: Linear polyoma DNA (2 μ g) was precipitated with the viral RNA in ethanol and the pellet was collected by centrifugation and then resuspended in 90% (V/V) formamide, 0.01 M Tris-HCl (pH 7.4), 0.001 M EDTA, 0.1% SDS. The mixture was incubated at 37°C for 15 min, to denature the DNA. At the end of the incubation period the

formamide was diluted to 75% with 0.01 M Tris·HCl (pH 7.4), 0.001 M EDTA and the mixture was brought to 0.3 M NaCl. The hybridization mixture was incubated at 51°C for 18 hr. The RNA-DNA hybrids were collected by ethanol precipitation and centrifugation. The pellet was resuspended in 66% (V/V) formamide, 3M-urea, 20 mM EDTA, 0.2 M Tricine-NaOH (pH 8.1) and was visualized by the electron microscope.

Electron microscopy: Aliquots in a volume of 15 μ l were diluted into 70% (V/V) formamide, 2.8 M-urea, 0.02 M EDTA, 0.2 M Tricine-NaOH (pH 8.1) to bring the concentration of the DNA to 0.5 μ g/ml. Before spreading, samples were heated at 43°C for 40 sec, cooled at room temperature for 1 min (5) and cytochrome c added immediately before spreading to a final concentration of 100 μ g/ml. Aliquots were spread onto a distilled water hypophase at room temperature (5). The spreading was visualized by sprinkling graphite powder onto the hypophase. Molecules were picked-up on parlodion coated grids, stained with uranyl-acetate and rotary-shadowed with platinum-palladium. Photographs were taken in a Phillips 300 electron microscope at a magnification of 40,000 x and were projected at final magnification of 200,000 x. The molecules were traced and measured with a map measurer and for short distances with a graphic instrument.

RESULTS

Sedimentation analysis of cytoplasmic viral RNA

Polyoma late cytoplasmic RNAs can be partially fractionated by sucrose gradient sedimentation (Fig. 1). Two or more species of RNA sedimenting between 15S and 20S could be resolved. The sedimentation profile of total cytoplasmic viral RNA (Fig. 1A) is similar to that of poly(A)-containing viral RNA (Fig. 1B) indicating that the viral RNAs are polyadenylated. The nature of the viral RNA sedimenting faster than 19S is not clear at the present time. Because of the small amount of the faster sedimenting RNA a more sensitive analysis should be employed to verify its nature.

In order to better quantitate and analyze the cytoplasmic viral RNAs, fractions containing poly(A)⁺ viral RNAs were pooled, as indicated in Fig. 1B, collected by ethanol precipit-

Fig. 1. Sedimentation analysis of cytoplasmic viral RNA

Thirty hours after infection, cells were labeled for 3 hrs and labeled cytoplasmic RNA extracted (12). In (A), sedimentation analysis of the labeled RNA was performed in 15-30% (W/W) sucrose gradient in SDS buffer. Centrifugation was for 18 hrs at 25,000 rpm at 20°C in a Spinco SW27 rotor. At the end of the run, fractions were collected and the radioactivity in 10 μ l of each fraction were counted in a Triton-based scintillation fluid (o—o). The two peaks are of the 28S and 18S ribosomal RNAs. Fifty- μ l aliquots from each fraction were hybridized to 0.7 μ g polyoma DNA immobilized on nitrocellulose filters (11) (●—●). In (B) poly(A)⁺ RNA was first selected by chromatography on oligo-(dT)-cellulose column (13) and then analyzed as in A.

ation and centrifugation, and used for R-loop analysis as described below.

Analysis of R-loops structures

R-loop structures were obtained by annealing double-stranded linear polyoma DNA (generated by cleavage of polyoma Form I DNA with HaeII restriction endonuclease which cleaves at position 0.72 (3)) with poly(A)-containing polyoma RNA purified as

described in Fig. 1B.

Figure 2A, B, and C shows representatives of three types of R-loop structures with their schematic tracings. In all of them the R-loops are located between short and long segments of duplex DNA and free tails are recognizable at each of the forks. The locations of the termini of the R-loops were computed with reference to the HaeII cleavage site (at 0.72 map units (3)) and the results are shown in Fig. 3. The histogram shows the existence of three distinct groups of R-loops and that the abundance of the middle sized R-loop (18S) was about twice that of the largest R-loop (19S). We could not determine from the present experiment the relative abundance of the smallest R-loops (16S) because a major fraction of them were not collected in the sucrose gradient (see Fig. 1B). The relative high abundance of R-loops pertaining to the 18S RNA exclude the possibility that this viral RNA is transcribed from the "early" strand since the early transcripts account for less than 5% of total cytoplasmic RNA (12, 16, 17).

Fig. 2. Electron micrographs and their tracings of R-loops composed of linear polyoma DNA (Form I cleaved at position 0.72 map units) and cytoplasmic poly(A)⁺ RNA. Hybrids formed with 19S RNA in (A) with 18S RNA in (B) and with 16S RNA in (C). In the tracings: a and a' = double-stranded DNA; b = free tail of poly(A); c = a heteroduplex DNA-RNA which includes the body of the RNA; d = displaced single-stranded DNA; e = free leader.

Fig. 3. Histogram of the locations on the physical map of polyoma DNA of the 3' and 5' ends of the bodies of polyoma late RNAs. The locations of the 5' and 3' ends of the bodies were determined from measurements of R-loops such as shown in Fig. 2. (The map positions with the standard deviations which appear in the text were determined from the measurements of the molecules which are present in each peak.)

The largest R-loop spans, in a counter-clockwise direction, between 0.654 ± 0.013 and 0.260 ± 0.029 , the middle-size R-loop spans in a counter-clockwise direction between 0.582 ± 0.013 and 0.256 ± 0.024 and the smallest R-loop spans in a counter-clockwise direction between 0.477 ± 0.020 and 0.252 ± 0.021 . Based on these and the results of Siddell and Smith (18) we suggest that the three types of R-loops are formed by 19S, 18S and 16S late mRNAs, respectively. The existence of three late polyoma RNAs was shown also by R. Kamen using the S1 method of Berk and Sharp (19) (personal communication).

The forks proximal to the short and long segments of DNA represent the map locations of the 5' and 3' ends of the bodies of these viral RNA, respectively. We suggest that the tails at the 3' ends represent non hybridized poly(A) sequences and the tails at the 5' ends represent the leader sequences which are not coded immediately adjacent to the bodies of the RNAs, and which

are transcribed from a not contiguous segment of the viral DNA. Similar structures were observed for the R-loops of 16S and 19S SV40 late mRNAs (5). The frequencies of recognizable tails at the 5' and 3' ends were similar, which indicates that all or almost all of the poly(A)-containing 19S, 18S and 16S late RNAs have leader sequences at their 5' ends. We measured the lengths of the leaders in about 100 molecules of each of the 19S, 18S and 16S RNAs and found them to be of 0.029 ± 0.01 , 0.036 ± 0.01 and 0.036 ± 0.01 map units, respectively.

If the tails at the 5' ends represent the leader, then in some molecules, the sequences in the leader should hybridize to DNA sequences not contiguous to the locations of the templates of the bodies of the viral RNAs and the intervening DNA segment should loop out. Figures 4A, B and C show examples of these structures, with their schematic tracings for 19S, 18S and 16S late RNAs, respectively. Similar structures were observed for 16S and 19S SV40 late RNAs (5). Molecules as shown in Fig. 4 were

Fig. 4. Electronmicrographs of R-loop structures as in Fig. 2 with the difference that the leader is partially hybridized with its template. Hybrids formed with 19S RNA in (A) with 18S RNA in (B) and with 16S RNA in (C). In the tracings: a and a' = double-stranded DNA; b = free tail of poly(A); c = a heteroduplex DNA-RNA which includes the body of the RNA; d = displaced single-stranded DNA; e = rehybridized leader; f = intervening DNA sequences. Bar represents 0.5 μ .

found in about 30% of the R-loop structures. We have not observed structures with double R-loops (20) and as a result we could not determine the map location of the 5' ends of the leaders. The intervening DNA in the structures pertaining to 19S late RNA (Fig. 4A) was only barely recognizable in the EM and it was shorter than that in 19S SV40 mRNA.

The histogram in Fig. 5 indicates the calculated map positions of the 3' and 5' ends of the bodies, and that of the 3' ends of

Fig. 5. Histogram of the locations on the physical map of polyoma DNA of the 3' and 5' ends of the bodies of the 16S and 18S RNAs and the 3' ends of their leaders. The locations of the 5' and 3' ends of the bodies and the location of the 3' end of the leaders were determined from molecules such as in Fig. 4. The length of the intervening DNA is the distance between the 5' end of the body and the 3' end of the leader. The map positions of the 5' end of the body and the 3' end of the leader of the 19S RNA are not resolved.

the leaders for the 16S and 18S viral RNAs. The histogram shows also the boundaries of the intervening DNA segments, for each of the two viral RNAs.

Fig. 6 shows a comparison between R-loop structures with rehybridized leaders of polyoma and SV40 late RNAs. It could be seen that double-R-loops were recognizable only for SV40 mRNAs and R-loop structures corresponding to 18S RNA were recognizable only for polyoma late RNAs. We cannot determine at the present time whether the failure to detect R-loops of 18S RNA in SV40 late RNAs indicates their absence in the SV40 infected cells or

Fig. 6. A comparison between R-loop structures with rehybridized leaders of polyoma and SV40 late RNAs. For details see Fig. 4 reference 5, and in the text. Double R-loops (20) were recognizable only with SV40 late mRNAs (B), and the 18S RNA was found only in polyoma infected cells. Bar represents 0.5 μ m.

whether they are present there in a much lower frequency than in polyoma infected cells.

We measured the distances between the HaeII cleavage site at 0.72 map units and the attachment points of the leaders to the DNA (Fig. 4). Based on these measurements we calculated this position to be at 0.655 ± 0.01 map units for both 18S and 16S late RNAs. We suggest that this point represents the template location for the 3' end of the leader.

We were unable to differentiate between the 5' ends of the body of the 19S RNA and the 3' end of its leader presumably because the intervening DNA is less than 50 nucleotides long. It appears however that the 3' end of the leader of 19S RNA would also map at the same position (0.655 ± 0.01 map units). If the length of the template is the same as that of the leader then the 5' end of the leader should map at a location which is 150-200 nucleotides downstream from 0.655 ± 0.01 map units.

Figure 7 is a summary of the above results in a form of a map of polyoma DNA showing the locations of the three late polyoma RNAs and the location of the leader.

DISCUSSION

Poly(A)-containing polyoma cytoplasmic RNA was hybridized with linear double-stranded polyoma DNA and RNA displacement loops (R-loops) were formed (4). The structures were analyzed in the electron-microscope. Three species of late viral RNAs were identified and mapped on the physical map of polyoma DNA. The map locations of the 19S and 16S RNAs are in agreement with previous results (21). We suggest that the three RNA species correspond to 19S, 18S and 16S viral mRNAs described recently by Siddell and Smith (18) and T. Hunter (personal communication) and which code for VP2 VP3 and VP1, respectively. It is worth mentioning that in SV40 infected cells only two viral RNA components, 16S and 19S were identified at late time after infection. Due to overall similarities of gene expression between SV40 and polyoma we believe that the failure to detect the 18S viral RNA species in SV40 infected cells is either because it exists in a much lower concentration than that in polyoma infected cells or that its size is similar to one of the other viral RNA

Fig. 7. Diagram of the circular polyoma DNA indicating the locations of the bodies of the 19S, 18S and 16S late RNAs and the location of the leader. The diagram is a summary of the results presented in the paper. The EcoR₁ site is taken as coordinate 0.0/1.0. One map unit corresponds to 0.01 fractional length of polyoma DNA and the map is read clockwise. Or. is the origin of replication (3). The location of the 5' end of the leader has not been determined.

species. Indeed in our previous studies of R-loops formed between late SV40 RNAs and linear SV40 DNA we have observed a small proportion of R-loop structures that could correspond to the 18S viral RNA (5).

Another conclusion of the present studies is that the three late polyoma RNAs are spliced, namely, the leader sequences at the 5' ends of the RNAs are not coded at a site immediately adjacent to the main portion of the RNAs. It is also evident from the present results in which a unique site was found for the 3' ends of the three leaders, that the leaders share common nucleotide sequences, but it is not known whether all the sequences are identical. We were unable to determine the map location of the

5' end of the leader because a second R-loop pertaining to the leader has not been observed (Fig. 6 shows structures with double-R-loops for SV40 late mRNAs). The reason for this failure is unknown at the present time. Several explanations could be suggested. They include: (i) the leader is shorter than that in SV40; (ii) the length of the leader is heterogeneous; (iii) the leader is segmented and (iv) the leaders contain sequences which are amplified with respect to the viral genome. The occurrence of the last possibility has recently been suggested by R. Kamen (personal communication).

ACKNOWLEDGMENTS

This work was supported by U.S. Public Health Service Research Grant CA 14995. We thank E. Jakobovitz for critical reading of the manuscript.

REFERENCES

1. Salzman, N.P. and Khoury, G. (1974) In: Comprehensive Virology 3, H. Fraenkel-Conrat and R. Wagner, Eds. (New York, Plenum Press), p. 63.
2. Sambrook, J.F. (1975) In: Control in virus multiplication D.C. Burke and W.C. Russels, Eds. (London and New York: Cambridge University Press) p. 153.
3. Griffin, B.E. and Fried, M. (1976) Methods in Cancer Res. Busch, H., Eds. (Academic Press, New York) 12, 49-86.
4. Thomas, M., White, R.L. and Davis, R.W. (1976) Proc. Natl. Acad. Sci. USA 73, 2069-2071.
5. Bratosin, S., Horowitz, M., Laub, O. and Aloni, Y. (1978) Cell, 13, 783-790.
6. Berget, S.M., Moore, C. and Sharp, P.A. (1977) Proc. Natl. Acad. Sci. USA 74, 3171-3175.
7. Aloni, Y., Dhar, R., Laub, O., Horowitz, M. and Khoury, G. (1977) Proc. Natl. Acad. Sci. USA 74, 3686-3690.
8. Klessig, D.F. (1977) Cell 12, 9-21.
9. Fogel, M. and Sachs, L. (1969) Virology 37, 327-334.
10. Penman, S. (1966) J. Mol. Biol. 17, 117-130.
11. Aloni, Y. and Locker, H. (1973) Virology 54, 495-505.
12. Laub, O. and Aloni, Y. (1975) J. Virol. 16, 1171-1183.
13. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
14. Hirt, B. (1967) J. Mol. Biol. 26, 365-369.
15. Radloff, R., Bauer, W. and Vinograd, J. (1967) Proc. Natl. Acad. Sci. USA 57, 1514-1521.
16. Flavell, A.J. and Kamen, R. (1977) J. Mol. Biol. 115, 237-242.
17. Bar-Shavit, R., Laub, O. and Aloni, Y. (1978) J. Gen. Virol. 39, 357-360.
18. Siddell, S.G. and Smith, A.E. (1978) J. Virol. 27, 427-431.
19. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
20. White, R.L. and Hogness, D.S. (1977) Cell 10, 177-192.

21. Kamen, R. and Shure, H. (1976) Cell 7, 361-371.
22. Chow, L.T., Gelinas, R.E., Broker, T.R. and Roberts, R.J. (1977)
Cell 12, 1-8.