Proc. Natl. Acad. Sci. USA
Vol. 75, No. 10, pp. 4754–4758, October 1978
Biochemistry

Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome

(DNA·RNA hybrids/primary transcripts/termination of transcription/mRNA precursors/RNA splicing)

NICHOLAS H. ACHESON

Department of Virology, Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland

Communicated by Igor Tamm, July 17, 1978

ABSTRACT **The bulk of late virus-specific RNA synthesized in polyoma virus-infected mouse cells is larger than a single strand of polyoma DNA. The arrangement of viral nucleotide sequences in these giant polyoma RNAs was studied by electron microscopy of hybrids between purified high molecular weight viral RNA and the *Hin*dII-1 fragment of polyoma DNA, which contains 91% of the viral genome. Hybrid molecules containing a short single-stranded gap (corresponding to the 9% of viral sequences not present in *Hin*dII-1), flanked by double-stranded regions, were photographed and measured. The majority of hybrid molecules contained no single-stranded loops or branches, showing that all viral sequences are transcribed contiguously and that no nonviral sequences are present in the RNA. Hybrid molecules, containing RNA up to 3.5 times the genome length, had a repeating structure of single-stranded gaps 8% of genome length interspersed with double-stranded regions 89% of genome length, showing that giant polyoma RNAs contain tandem repeats of the nucleotide sequence of the entire viral DNA. A small proportion of hybrid molecules contained single-stranded branches or deletion loops in characteristic positions, indicating that RNA "splicing" may occur on high molecular weight nuclear polyoma RNA.**

Transcription of polyoma DNA during the late phase of productive infection (for review, see ref. 1) gives rise to "giant" RNA molecules up to 3–4 times the size of the viral genome (2–5). These giant polyoma-specific RNAs are limited to the nucleus (6) where transcription takes place; viral mRNAs in the cytoplasm are one-quarter to one-half the size of the viral DNA (6–8). Most late nuclear polyoma-specific RNA is transcribed from the L strand of viral DNA (9); all sequences of the L strand are represented in late nuclear transcripts (10, 11).

To try to understand the mechanism by which these giant RNA molecules are synthesized, I have hybridized this RNA with a fragment (*Hin*dII-1) of polyoma DNA consisting of 91% of the genome (12, 13) and have examined these hybrids by electron microscopy. The results show that polyoma giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome and that they contain no detectable nonviral sequences.

MATERIALS AND METHODS

Preparation of RNA. Primary mouse kidney cell cultures were infected with plaque-purified polyoma virus at a multiplicity of 20–40 plaque-forming unts per cell. At 24 hr after infection, cells were labeled for 3 hr with 333 μCi of [5-^{3}H]-uridine (Amersham, 25 Ci/mmol) per culture dish in 1 ml of reinforced Eagle's medium. Cells were scraped into cold 10 mM triethanolamine, pH 8.5/10 mM NaCl/1.5 mM $MgCl_2$ and lysed by addition of Nonidet P40 (Shell) to 0.5% final concentration. Nuclei were pelleted and resuspended in the same buffer. RNA was extracted by addition of 1 vol of 10 mM triethanolamine (pH 8.5)/100 mM NaCl/10 mM EDTA/4% triisopropylnaphthylenesulfonate (Serva) (14, 15) and 2 vol of phenol/chloroform/isoamyl alcohol, 50:50:1 (vol/vol). After three successive extractions at room temperature, nucleic acids in the aqueous phase were precipitated three times in succession with ethanol and then incubated at 0° for 20 min with DNase I (RNase-free; Worthington; 10 μg/ml) in 10 mM Na acetate, pH 5.1/100 mM NaCl/2 mM $MnCl_2$. RNA was extracted with phenol/chloroform in the presence of 1% sodium dodecyl sulfate ($NaDodSO_4$) and ethanol-precipitated two times.

Preparative Hybridization of RNA to Polyoma DNA. Polyoma form I DNA was denatured and bound to nitrocellulose filters as described (2), at an input of 20 μg of DNA per 15-mm-diameter filter. Nuclear RNA was dissolved in 50% (vol/vol) formamide/300 mM NaCl/30 mM Na citrate/0.1% $NaDodSO_4$ (16, 17) and incubated with DNA-containing filters at 37° on a shaking water bath for 16–20 hr. Filters were washed twice for 15 min each in the same buffer at 37°, twice for 5 min each in 15 mM NaCl/1.5 mM Na citrate/0.1% $NaDodSO_4$ at room temperature, and twice for 2.5 min each in the same buffer at 65°. The last washing step was found necessary to remove nonviral RNA which binds to filters in presence of viral DNA·RNA hybrids (unpublished results). Hybridized RNA was eluted by incubating the filters three times for 2.5 min each in 90% formamide/10 mM triethanolamine, pH 7/0.1% $NaDodSO_4$ at 65°. Carrier *Escherichia coli* tRNA was added, and RNA was ethanol precipitated two times.

Dimethyl Sulfoxide/Sucrose Gradient Sedimentation of Eluted Polyoma-Specific RNA. RNA was resuspended in 50 μl of 1 mM EDTA, to which was added 450 μl of 99% dimethyl sulfoxide/1 mM EDTA, pH 7. The sample was incubated at 50° for 5 min, then layered onto a 99% dimethyl sulfoxide/sucrose gradient (2, 18) and centrifuged at 23,000 rpm for 38 hr at 27° in a Spinco SW-40 rotor.

Hybridization of Polyoma-Specific RNA with DNA Fragment *Hin*dII-1. Approximately 25 ng of purified, high molecular weight RNA was ethanol precipitated along with 200 ng of electrophoretically purified DNA fragment *Hin*dII-1 (a 5- to 10-fold excess of DNA over RNA), pelleted, and resuspended in 10 μl of 70% formamide/300 mM NaCl/10 mM Tris, pH 8.5/1 mM EDTA. The solution was incubated at 65° for 5 min (to denature DNA), then at 45° for 4 hr (to allow DNA·RNA hybridization), and finally at 58° for 5 min (to denature DNA·DNA duplexes without denaturing DNA·RNA hybrids) (19, 20).

Spreading and Electron Microscopy of DNA·RNA Hybrids. Hybridized nucleic acids were diluted 10- to 25-fold into 50 μl (final volume) of 50% formamide/80 mM Tris, pH 8.5/8 mM EDTA containing 50 μg of cytochrome *c* per ml. This solution was spread onto a hypophase of distilled water. Grids were stained and shadowed as described by Wellauer and Dawid (20). Molecules were projected to a magnification of ×204,000 and were measured with a Hewlett–Packard map reader and computer.

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: $NaDodSO_4$, sodium dodecyl sulfate; ss, single-stranded; ds, double-stranded; kb, kilobases.

RESULTS

Design of the Experiment. If giant polyoma-specific RNA is transcribed by continuous transit of RNA polymerase around circular viral DNA, the RNA should contain tandem repeats of the nucleotide sequence of viral DNA (Fig. 1*A*). Hybrids with a fragment of polyoma DNA such as *Hin*dII-1, which contains 90.6% of the sequences in the viral DNA (12, 13), should reveal a repeated structure of double-stranded (ds) regions approximately 91% of genome-length separated by single-stranded (ss) regions 9% of genome-length. If, on the other hand, giant RNA is transcribed from viral DNA singly integrated into the host cell chromosome (Fig. 1*B*), it should contain only one genome equivalent of viral sequences, the remainder being host sequences. Resultant hybrids with *Hin*dII-1 DNA should show one or two ds regions (depending on where the DNA is broken before integration) whose summed length is 91% of the genome length, plus ss forks at one or both ends. If certain regions of the viral DNA are not transcribed at all—e.g., by "jumping" of the RNA polymerase (Fig. 1*C*)—regularly spaced deletion loops or branches of ss DNA should be found in DNA·RNA hybrids.

Purification of High Molecular Weight Polyoma-Specific Nuclear RNA. Viral RNA was purified from cellular RNA by preparative hybridization in formamide to decrease degradation of the RNA. Exhaustive rehybridization of purified RNA showed that at least 70% was able to form RNase-resistant hybrids with viral DNA (data not shown). The purified RNA was sedimented on a 99% dimethyl sulfoxide/sucrose gradient to separate high molecular weight RNA from eluted DNA and degraded RNA (Fig. 2). The mean sedimentation coefficient of the eluted RNA was about 28 S. RNAs with sedimentation coefficients up to 45 S were detected; however, the average size of the RNA was smaller than that of 20-min pulse-labeled RNA directly analyzed on dimethyl sulfoxide/sucrose gradients (2). This is probably a result of some degradation during preparative hybridization of the RNA, as well as the longer labeling time (3 hr) used in this experiment which could have allowed normal intracellular processing events to take place on labeled RNA.

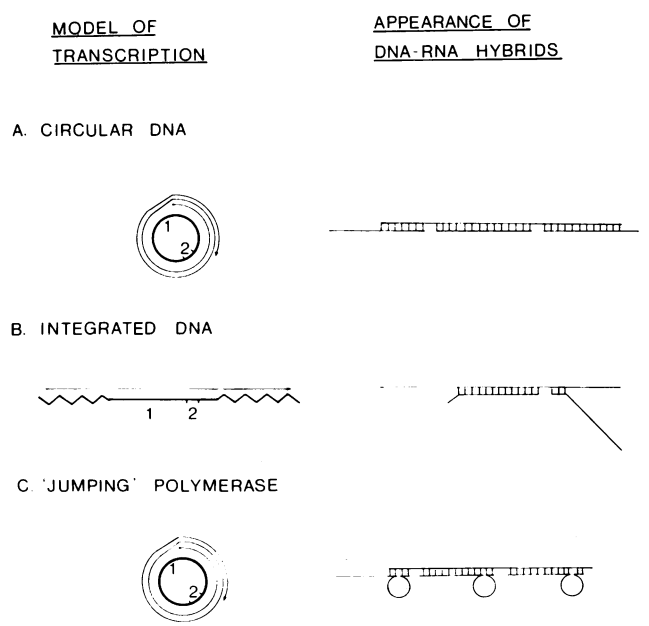

FIG. 1. Models for the mechanism of synthesis of late giant polyoma-specific RNA, and appearance of hybrids between resulting transcripts and fragment *Hin*dII-1 of polyoma DNA. Numbers 1 and 2 refer to *Hin*dII fragments of polyoma DNA. Viral DNA, —; host DNA, ~~; primary transcripts, —; DNA·RNA hybrids, ▬. See text for details.

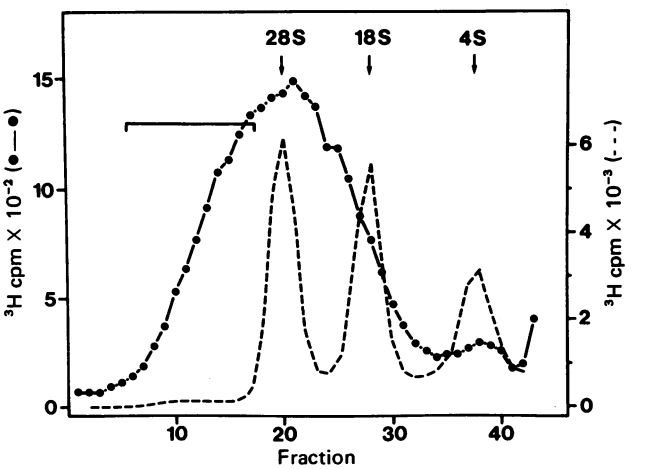

FIG. 2. Dimethyl sulfoxide/sucrose gradient profile of [^{3}H]uridine-labeled polyoma-specific late nuclear RNA (●) purified by preparative hybridization. [^{3}H]Uridine-labeled poly(A) − cytoplasmic RNA (- - -) from the same cells was centrifuged in a parallel gradient in the same rotor, as a sedimentation standard. The positions of 28S and 18S rRNA and 4S RNA are marked by vertical arrows. The horizontal bar marks the fractions pooled for subsequent hybridization and electron microscopy.

Fractions containing RNA larger than approximately 30 S ($M_r > 2 \times 10^6$) were pooled (horizontal bar in Fig. 2).

Hybridization of RNA to *Hin*dII-1 Fragment of Polyoma DNA and Electron Microscopy of Hybrids. High molecular weight polyoma-specific RNA was hybridized with an excess of purified *Hin*dII-1 DNA under conditions (70% formamide, 45°–58°) that minimize thermal breakage of RNA and that allow DNA·RNA hybrids to form but denature most renatured DNA molecules, thus facilitating the search for hybrid molecules by electron microscopy. Examination of the hybridization mixture after spreading revealed the following classes of molecules: (*i*) ss DNA and ds DNA of the size expected for the *Hin*dII-1 fragment [4.9 kilobases (kb)]; these were used as internal length standards. (*ii*) molecules containing a ds region flanked at one or both ends by ss tails but not by ss forks (these molecules, which probably represent hybrids between DNA and RNA fragments shorter than one genome length, were not further analyzed); and (*iii*) molecules containing a short ss "gap" flanked by ds regions of variable length, followed by additional ss and ds regions. As discussed in detail below, these molecules probably are hybrids between one RNA molecule and two or more *Hin*dII-1 DNA fragments. The ss gap represents RNA joining two DNA·RNA hybrid regions. Several examples of hybrid molecules are shown in Fig. 3, and the results of measuring 40 such molecules are displayed in Fig. 4. All these molecules have been aligned with one end of the ss gap at the same position.

Both linear and circular hybrid molecules containing short ss gaps were seen. Circular molecules result from hybridization of a single DNA molecule with both ends of an RNA molecule. If the RNA is longer than one genome length, a circular hybrid molecule will carry a ss RNA tail (Figs. 3*E* and 4, molecule 36c); circular hybrid molecules containing RNA shorter than genome length contain a second internal ss region (Fig. 4, molecules 34c, 35c, 39c, and 40c). Double-length circles, which may result from hybridization of RNA longer than one genome length with two DNA molecules, also were present (molecules 9c and 34c).

Most hybrid molecules contained no internal ss branches or loops. These molecules therefore contained RNA that is a continuous transcript of the entire region hybridized. In those hybrid molecules that contained contiguous DNA·RNA hybrid regions longer than single genome length, therefore, *all* the

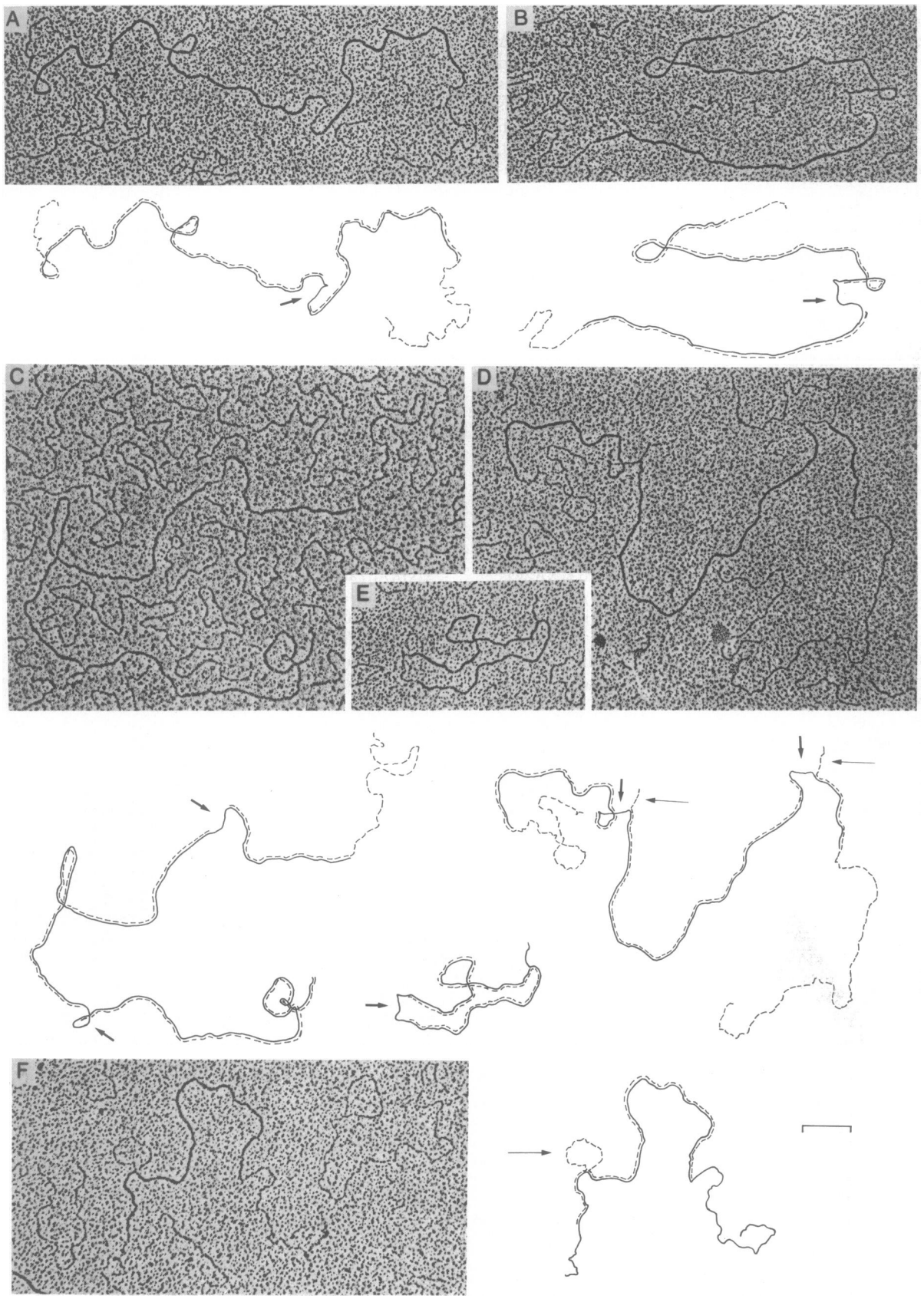

FIG. 3. Electron micrographs and diagrams of hybrid molecules between high molecular weight nuclear polyoma-specific RNA and polyoma DNA fragment *Hin*dII-1. (*A* and *B*) Hybrid molecules each containing two *Hin*dII-1 DNA fragments linked by an RNA molecule. Arrow, ss RNA gap. (*C*) Hybrid molecule containing three *Hin*dII-1 DNA·RNA hybrid regions and two ss RNA gaps. (*D*) Molecule similar to that in *C* but with short ss branches (thin arrows) at one end of each ss RNA gap. (*E*) Circular hybrid molecule with a ss RNA gap and a ss RNA tail. (*F*) Hybrid molecule containing a ss deletion loop internal to a hybrid region. In the interpretative drawings, RNA is represented as a continuous line (—) and DNA as a dashed line (- - -). Scale (lower right corner) = 0.5 kb.

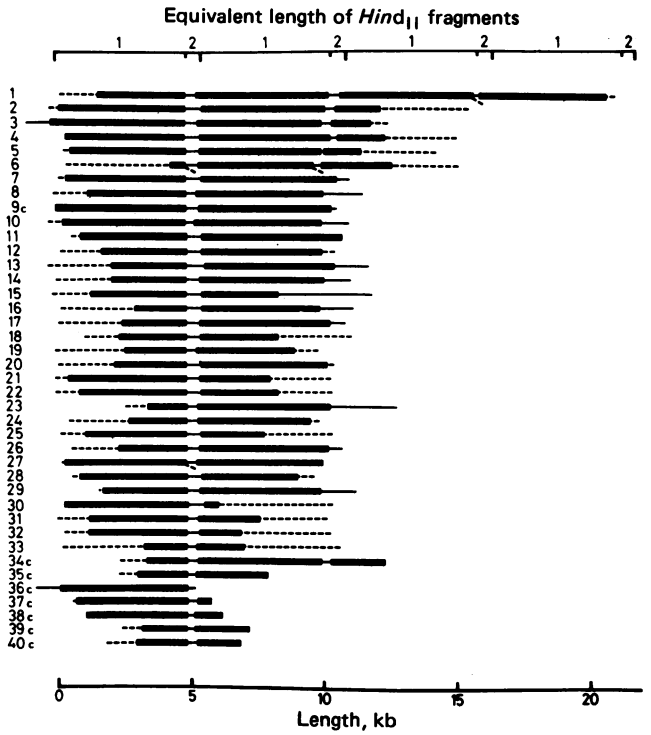

FIG. 4. Length of ss and ds regions in hybrid molecules that contain at least one ss RNA gap. ss regions at ends of hybrid molecules that terminate approximately one *Hin*dII-1 fragment length from a ss RNA gap are interpreted as being ss DNA (- - -). All other terminal ss regions, as well as ss gaps of approximately one *Hin*dII-2 fragment length, are interpreted as being ss RNA (—). ds regions in hybrid molecules are interpreted as being DNA·RNA hybrids (▬). Short ss branches projecting at one end of gaps in molecules 1, 6, and 27 are considered to be ss DNA (see text). Molecules 9c and 34c–40c are circular and have been displayed as linear by cutting at one of the presumed ends of the RNA. ss DNA (but not RNA) lengths have been corrected to linear density of ds DNA by multiplying by the ratio of the length of renatured ds *Hin*dII-1 DNA to the length of ss *Hin*dII-1 DNA measured from photographs taken from the same spreadings. Lengths are expressed on the upper scale as unit lengths of polyoma *Hin*dII fragments and on the lower scale as kilobases (taking 5.4 kb per polyoma genome).

nucleotide sequences of one strand of the viral DNA were represented at least once in the RNA. It cannot be excluded, however, that one or more short segments of less than 50 nucleotides might be missing from these transcripts, because such short loops cannot be detected by electron microscopy.

Lengths of ss and ds Regions in Hybrid Molecules. The distance between two ss gaps (in molecules with more than one gap), or between a gap and the end of the DNA within a hybrid molecule, should be close to the length of the *Hin*dII-1 fragment—i.e., 91% of the viral DNA length (about 4.9 kb). Fig. 5 shows the distribution of these lengths for all molecules displayed in Fig. 4. The small number of gap-to-gap and gap-to-end distances that are significantly smaller than this length are probably a result of breakage of DNA during hybridization or spreading. The mean (±SD) length of all gap-to-gap and gap-to-end distances, excluding the five abnormally short ends, is 4.8 ± 0.2 kb, or 88% of the viral DNA length. If only gap-to-gap distances (i.e., DNA·RNA hybrids the full length of the *Hin*dII-1 fragment) are considered (Table 1), their mean length is 89% of polyoma DNA length. The mean value of the lengths of all ss gaps is also shown in Table 1. This value (8% of polyoma DNA length) is close to that expected (9.4%) if these gaps represent the sequences present in DNA fragment *Hin*dII-2.

RNA lengths in hybrid molecules can be deduced from the length of the ds regions extending in each direction from a ss

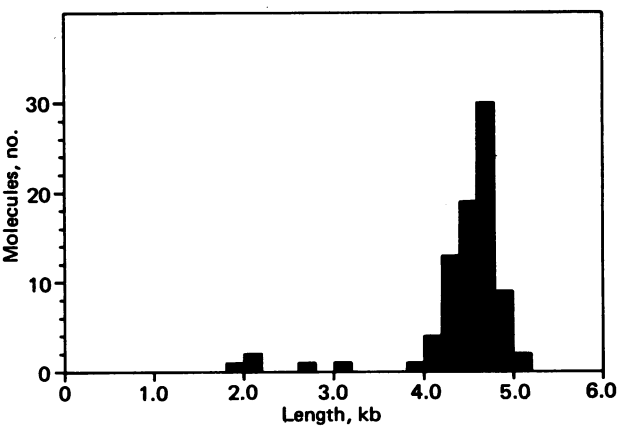

FIG. 5. Distribution of lengths of all internal ds regions flanked by two ss gaps (i.e., fully hybrid regions), and all terminal ds regions, or ds plus ss regions, whichever comes closer to one *Hin*dII-1 fragment length. Measurements were taken from corrected data shown in Fig. 4.

gap. The length of RNA in the 40 measured hybrid molecules varied from less than one genome length to 3.5 times the genome length (see Fig. 4).

Branches and Loops. Three of the molecules shown in Fig. 4 (no. 1, 6, and 27) contained ss branches located at one end of the ss gap region. In molecule 6 (see Fig. 3*D*), both gaps had short ss branches, each at the same end of the gap. Because the gap regions consist of ss RNA that joins two adjacent DNA·RNA hybrid regions, these branches must be DNA at one of the ends of the *Hin*dII-1 fragment. This implies that a short region of approximately 200 nucleotides is missing at this position in these RNA molecules. This could result from transcription of defective DNA that lacks sequences at this position, or it could result from "splicing" of transcripts of nondefective DNA. Because one of the ends of the *Hin*dII-1 fragment is near the position where the 3′ termini of L-strand mRNAs are located (7, 13), a site for cleavage of the primary transcript may well be located in this region. Further experiments are needed to clarify the nature of these molecules.

Finally, two hybrid molecules (one shown in Fig. 3*F*) contained internal ss loops. The loops in both molecules were of the same size (0.97 kb) and were located at the same position with respect to the ends of the *Hin*dII-1 fragment. It is likely that these looped molecules result from hybridization of full-length *Hin*dII-1 DNA with RNA in which this 0.97-kb region is deleted. Interestingly, this deletion corresponds in position and length to the region in simian virus 40 and polyoma DNA that is deleted between the main body of the 16S L-strand mRNA

Table 1. Lengths of *Hin*dII-1 ds DNA molecules and of full-length DNA·RNA hybrid regions and ss RNA "gaps" in hybrid molecules

Type of molecule	Molecules, no.	Length, μm (mean ± SD)	Fraction of polyoma DNA length
ds *Hin*dII-1 DNA	55	1.63 ± 0.03	0.91
DNA·RNA hybrid regions*	18	1.60 ± 0.06	0.89
ss RNA "gaps"†	46	0.14 ± 0.04	0.08

* ds regions flanked on both sides by ss RNA gaps, or in circular molecules with one or more ss RNA gaps.

† ss regions flanked on both sides by ds (DNA·RNA hybrid) regions and separated by one *Hin*dII-1 fragment length from an adjacent ss region.

and its 5′-end "leader" sequences (21–23; unpublished results).

DISCUSSION

The experiment described here demonstrates that most high molecular weight nuclear polyoma-specific RNAs contain at least one complete copy of the nucleotide sequences present in one strand of the viral DNA. Hybrid molecules between RNA and the *Hin*dII-1 DNA fragment reveal the presence, in long RNA molecules, of 1 to 3½ copies of viral sequences, organized in tandem repeats of genome length. The majority of hybrid molecules contain no deletion loops or branches; thus, *all* sequences in a strand of viral DNA are present in these transcripts. Also, hybrid molecules do not contain substitution loops or terminal ss forks which would reveal the presence of nonviral sequences in the RNA. Thus, all sequences in these RNAs are virus-specified.

Because most polyoma-specific late RNA labeled during a brief exposure of infected cells to [^{3}H]uridine is of greater than genome length (2), the molecules analyzed in this study are probably representative of the majority of primary transcripts of polyoma DNA. Because more than 90% of late transcription is specific for the L strand of viral DNA (9), the present results are probably relevant only to L-strand transcripts and say nothing about the size or sequence content of E-strand transcripts.

The results of this study rule out the possibility that a significant fraction of late viral RNA is transcribed from viral DNA singly integrated into the host cell chromosomes, because such transcripts would contain only one genome length of viral sequences linked at one or both ends to nonviral sequences. Experiments by Lev and Manor (4), as well as my own unpublished results, had already indicated that most sequences in giant polyoma-specific RNA are viral and not host. However, they did not rule out the presence of some nonviral sequences either interspersed within or at the extremities of viral sequences.

The simplest model consistent with these results is that monomeric circular viral DNA is the template for late transcription (Fig. 1A). Analysis of viral transcription complexes (24–26) has shown that at least some late viral transcription takes place on monomeric circular DNA. It remains theoretically possible that synthesis of giant RNA is directed by oligomers of viral DNA, which represent 1–2% of viral DNA in infected cells (27–29).

In either case, termination of transcription must be rare or nonexistent on the L strand of polyoma DNA, allowing the RNA polymerase to transit several cycles around the viral DNA without releasing the nascent RNA chain. This is in contrast to transcription of the (complementary) E strand of polyoma DNA during the early phase of virus infection; most E-strand transcripts appear to be terminated near the end of the "early" region of the genome, giving rise to RNA about half the length of the viral DNA (10; unpublished results).

Polyoma L-strand mRNAs (unpublished results), like the analogous simian virus 40 mRNAs (21–23), contain sequences from noncontiguous regions of the viral DNA and thus are spliced. Because most of the polyoma giant RNA molecules analyzed in this study are not lacking sequences present in the viral DNA, these results suggest that splicing is a post-transcriptional event. The two anomalous classes of hybrid molecules detected, which appear to lack specific sequences at the boundaries of mRNA regions, suggest that splicing may take place on high molecular weight RNA before final cleavage at the 5′ and 3′ ends of the mRNAs.

I thank Peter Wellauer for introducing me to electron microscopy of nucleic acids, Françoise Miéville for excellent technical assistance, and Peter Beard, Bernhard Hirt, and Peter Wellauer for critical comments on the manuscript. This work was supported by the Swiss National Science Foundation, Grant 3.691.76.

1. Acheson, N. H. (1976) *Cell* **8**, 1–12.
2. Acheson, N. H., Buetti, E., Scherrer, K. & Weil, R. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2231–2235.
3. Rosenthal, L. J., Salomon, C. & Weil, R. (1976) *Nucleic Acids Res.* **3**, 1167–1183.
4. Lev, Z. & Manor, H. (1977) *J. Virol.* **21**, 831–842.
5. Birg, F., Favaloro, J. & Kamen, R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3138–3142.
6. Buetti, E. (1974) *J. Virol.* **14**, 249–260.
7. Kamen, R. & Shure, H. (1976) *Cell* **7**, 361–371.
8. Türler, H., Salomon, C., Allet, B. & Weil, R. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1480–1484.
9. Flavell, A. J. & Kamen, R. (1977) *J. Mol. Biol.* **115**, 237–242.
10. Beard, P., Acheson, N. H. & Maxwell, I. N. (1976) *J. Virol.* **17**, 20–26.
11. Kamen, R., Lindstrom, D., Shure, H. & Old, R. W. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **39**, 187–198.
12. Folk, W. R., Fishel, B. R. & Anderson, D. M. (1975) *Virology* **64**, 277–280.
13. Griffin, B. E. (1977) *J. Mol. Biol.* **117**, 447–471.
14. Kirby, R. S. (1965) *Biochem. J.* **96**, 266–269.
15. Fraser, R. S. S. & Loening, U. E. (1973) *Eur. J. Biochem.* **34**, 153–158.
16. Weinberg, R. A., Warnaar, S. O. & Winocour, E. (1972) *J. Virol.* **10**, 193–201.
17. McConaughy, B. L., Laird, C. D. & McCarthy, B. J. (1969) *Biochemistry* **8**, 3289–3295.
18. Strauss, J. H., Jr., Kelly, R. B. & Sinsheimer, R. L. (1968) *Biopolymers* **6**, 793–807.
19. Casey, J. & Davidson, N. (1977) *Nucleic Acids Res.* **4**, 1539–1552.
20. Wellauer, P. K. & Dawid, I. B. (1977) *Cell* **10**, 193–212.
21. Aloni, Y., Dhar, R., Laub, O., Horowitz, M. & Khoury, G. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3686–3690.
22. Hsu, M.-T. & Ford, J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4982–4985.
23. Lavi, S. & Groner, Y. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5323–5327.
24. Green, M. H. & Brooks, T. L. (1976) *Virology* **72**, 110–120.
25. Condit, R. C., Cowie, A., Kamen, R. & Birg, F. (1977) *J. Mol. Biol.* **115**, 215–235.
26. Gariglio, P. & Mousset, S. (1975) *FEBS Lett.* **56**, 149–155.
27. Cuzin, F., Vogt, M., Dieckmann, M. & Berg, P. (1970) *J. Mol. Biol.* **47**, 317–333.
28. Jaenisch, R. & Levine, A. (1971) *Virology* **44**, 480–493.
29. Shane, M., Seidman, M. & Salzman, N. P. (1977) *Virology* **83**, 27–29.