MOLECULAR AND CELLULAR BIOLOGY, Apr. 1984, p. 722–729
0270-7306/84/040722-08$02.00/0

Vol. 4, No. 4

Kinetics and Efficiency of Polyadenylation of Late Polyomavirus Nuclear RNA: Generation of Oligomeric Polyadenylated RNAs and Their Processing into mRNA

NICHOLAS H. ACHESON

Department of Microbiology and Immunology, McGill University, Montreal, Quebec H3A 2B4 Canada

Received 12 December 1983/Accepted 19 January 1984

The rate and efficiency of polyadenylation of late polyomavirus RNA in the nucleus of productively infected mouse kidney cells were determined by measuring incorporation of [^{3}H]uridine into total and polyadenylated viral RNAs fractionated by oligodeoxythymidylic acid-cellulose chromatography. Polyadenylation is rapid: the average delay between synthesis and polyadenylation of viral RNA in the nucleus is 1 to 2 min. However, only 10 to 25% of viral RNA molecules become polyadenylated. Polyadenylated RNAs in the nucleus are a family of molecules which differ in size by an integral number of viral genome lengths (5.3 kilobases). These RNAs are generated by repeated passage of RNA polymerase around the circular viral DNA, accompanied by addition of polyadenylic acid to a unique 3′ end situated 2.2 + *n*(5.3) kilobases from the 5′ end of the RNAs (*n* can be an integer from 0 to at least 3). Between 30 and 50% of the sequences in nuclear polyadenylated RNA are conserved during processing and transport to the cytoplasm as mRNA. This is consistent with the molar ratios of nuclear polyadenylated RNAs in the different size classes, and it suggests that most polyadenylated nuclear RNA is efficiently processed to mRNA. Thus, the low overall conservation of viral RNA sequences between nucleus and cytoplasm is explained by (i) low efficiency of polyadenylation of nuclear RNA and (ii) removal of substantial parts of polyadenylated RNAs during splicing. The correlation between inefficient termination of transcription and inefficient polyadenylation of transcripts suggests that these two events may be causally linked.

During the late phase of infection of mouse cells by polyomavirus, only about 5% of the viral RNA synthesized in the nucleus is successfully processed and transported to the cytoplasm as mRNA (2). This low efficiency of conservation of viral RNA during processing can be explained partly by the fact that RNA polymerase II may traverse the circular viral DNA genome (5,292 nucleotide pairs [11, 35]) several times before terminating transcription (1, 3). This leads to production of giant RNAs containing several tandemly repeated copies of the entire nucleotide sequence of one of the strands of the viral DNA. Only a single mRNA body sequence per precursor RNA molecule is conserved during RNA splicing (8, 15, 21); the remaining mRNA sequences are presumably discarded in the nucleus. In addition, two reports have indicated that only a fraction of polyomavirus RNA in the nucleus becomes polyadenylated during the late phase (7, 30); since viral mRNAs in the cytoplasm are polyadenylated (16, 29), presumably only precursor molecules in the nucleus containing polyadenylic acid [poly(A)] give rise to mature mRNA. These considerations led me to examine more closely the polyadenylation of viral RNA in the nucleus by measuring carefully the kinetics of incorporation of [^{3}H]uridine into polyadenylated and nonpolyadenylated RNA.

I have confirmed that only a small fraction (10 to 25%) of nuclear viral RNA is polyadenylated, that it is polyadenylated within less than 2 min of its synthesis, and that the remaining 75 to 90% of nuclear viral RNA is apparently never polyadenylated and cannot therefore serve as a precursor to viral mRNA in the cytoplasm. Nuclear viral RNA which is polyadenylated exists as molecules of discrete lengths of about 2.2 + *n*(5.3) kilobases (kb). These RNAs represent a series which have the same 5′ and 3′ ends but differ in the number of genome traverses made by RNA polymerase II before terminating. These results are consistent with the idea that most polyadenylated nuclear RNAs are successfully processed into mature mRNA and that a major cause of low conservation of viral RNA during processing is inefficient polyadenylation of viral RNA molecules in the nucleus. The intriguing relationship between inefficient termination of transcription and inefficient polyadenylation of transcripts is discussed.

MATERIALS AND METHODS

Labeling of RNA, cell fractionation, and extraction of RNA. Primary baby mouse kidney cells were labeled with an excess of [5-^{3}H]uridine for various lengths of time beginning 28 h after infection with 20 to 40 PFU per cell of polyomavirus. Nuclei and cytoplasm were separated after lysis in cold 10 mM NaCl–1.5 mM $MgCl_2$–10 mM triethanolamine-hydrochloride (pH 8.5)–1% Nonidet P-40. RNA was extracted as described (4). This fractionation method gives relatively high yields of cytoplasmic RNA and therefore relatively pure nuclear RNA, but a small amount of nuclear leakage can be detected in the cytoplasmic extract. For a detailed discussion of these points, see Acheson (1).

Oligo(dT)-cellulose chromatography. Nuclear or cytoplasmic RNA from two 88-mm petri dishes (150 to 300 μg RNA per sample) was suspended in 2 ml of 10 mM triethanolamine-hydrochloride (pH 7.4)–0.5 M NaCl and was passed through a column containing 0.5 g of oligodeoxythymidylic acid (oligo[dT])-cellulose equilibrated with the same buffer. The column was washed with 12 ml of load buffer and then was eluted with 3 ml of 10 mM triethanolamine-hydrochloride (pH 7.4). Unbound (nonpolyadenylated) and eluted (polyadenylated) fractions were pooled and ethanol precipitated.

DNA-RNA hybridization and normalization of results. RNAs were hybridized to an excess of purified, denatured

polyomavirus DNA bound to nitrocellulose filters. Hybridization took place in 4× SSC (SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 65°C for 40 h as described elsewhere (3). All samples were hybridized in duplicate tubes, each tube containing two DNA-loaded filters and a blank. Filters were counted in a toluene-based scintillation fluid by using a narrow tritium channel with a machine background of 3 to 4 cpm. Three 5-min counts were done for each filter. Results were normalized to correct for variable recovery of RNA from sample to sample as described (2), and values were expressed as counts per minute hybridized per 100 μg of total (nuclear plus cytoplasmic) RNA in the original sample.

RESULTS

I wished to determine the relative rates of incorporation of [^{3}H]uridine into polyadenylated and nonpolyadenylated viral RNA during short labeling times to measure the rapidity and efficiency with which late polyomavirus RNA in the nucleus is polyadenylated. Achieving valid results in such experiments requires that certain experimental conditions be met. These include (i) rapid achievement of constant specific activity of nucleotide pools, (ii) adequate separation of nuclei and cytoplasm, (iii) efficient separation of polyadenylated and nonpolyadenylated RNAs by affinity chromatography, and (iv) quantitative determination of virus-specific RNA in each sample by DNA-excess hybridization. Points (i) and (ii) have been discussed in detail previously (2). In polyomavirus-infected baby mouse kidney cells, equilibration of the UTP pool with labeled exogeneous uridine is nearly complete within 3 to 6 min, as shown by the linear incorporation of [^{3}H]uridine into cellular 4S RNA for labeling periods of 5 to 120 min. Cytoplasmic contamination of the nuclear fraction as prepared here was minimal (10 to 20% of total cytoplasm), and nuclear leakage was at an acceptably low level (less than 5% of nuclear RNA). Since only a small fraction of the viral RNA labeled during periods of up to 2 h is found in the cytoplasm, these levels of cytoplasmic contamination of the nuclear extract do not significantly affect the precision of measurement of nuclear viral RNA. Even low levels of nuclear leakage, however, affect the measurement of cytoplasmic RNA levels. But since most nuclear viral RNA is not polyadenylated, measurements of the amount of radioactive polyadenylated viral RNA in the cytoplasm are not seriously affected (see below).

Table 1 shows that oligo(dT)-cellulose chromatography adequately separated polyadenylated and nonpolyadenylated RNAs whether they were viral or cellular, nuclear or cytoplasmic. It has been reported that nuclear polyadenylated RNAs are not efficiently retained by affinity chromatography with polyuridylic acid-Sepharose (23); our results confirm those of others (24, 31) that oligo(dT)-cellulose does bind nuclear polyadenylated RNAs efficiently. Finally, our filter-bound DNA-RNA hybridization method has been shown to be quantitative in a number of previous publications (2–4).

Proportion of pulse-labeled viral RNA in the nucleus which is polyadenylated. Cells were labeled with [^{3}H]uridine for periods of 5 to 60 min during the late phase of infection. Nuclear RNA samples were fractionated by passage through a column of oligo(dT)-cellulose, and portions of each sample were hybridized with an excess of polyomavirus DNA to determine radioactivity in virus-specific RNA. Figure 1 shows that, at all labeling times, only a small percentage of the viral RNA in the nucleus was polyadenylated. Since accumulation of radioactivity in both total and polyadenylated viral RNAs is linear with time up to about 40 min, the ratio of the initial slopes of the two accumulation curves is a measure of the proportion of viral RNA which becomes polyadenylated. That ratio for the experiment shown in Fig. 1 was 0.13; thus, 13% of the viral RNA synthesized was polyadenylated. In two other experiments, 14 and 23% of the viral RNA in the nucleus was polyadenylated. That this result was not due to inefficient binding of polyadenylated RNA to oligo(dT)-cellulose has been shown above. However, if viral RNA is degraded during isolation, the proportion of molecules with poly(A) at their 3′ ends could be artificially low. Results of our previous experiments have shown that intact, high-molecular-weight RNA can be extracted routinely from polyomavirus-infected mouse cells (1, 3). In addition, the size distribution of viral polyadenylated RNA in the nucleus (see below) strongly argues against the possibility that significant degradation, which might lead to a underestimation of the proportion of viral RNA which is polyadenylated, took place.

Interval between synthesis and polyadenylation of viral RNA in the nucleus. It has been shown in several systems that nuclear mRNA precursors are polyadenylated within 1 to 2 min of the completion of their synthesis (26, 32). One possible explanation for the apparently low efficiency of polyadenylation of polyomavirus nuclear RNA could be that polyadenylation is exceptionally slow, so that most viral RNAs are polyadenylated only after a much longer delay. If this were the case, radioactivity would only begin to accu-

TABLE 1. Retention of cellular and viral polyadenylated RNA by oligo(dT)-cellulose[a]

RNA	Retention by oligo(dT)-cellulose	cpm	cpm in poly(A)	% of poly(A) bound
Cytoplasmic				
Cellular	Bound	0.53×10^6	83×10^3	>94
	Unbound	1.0×10^6	$<5 \times 10^3$	
Viral	Bound	77×10^3	10.6×10^3	>99
	Unbound	21.5×10^3	$<0.1 \times 10^3$	
Nuclear				
Cellular	Bound	1.2×10^6	47.8×10^3	>92
	Unbound	4.9×10^6	$<4 \times 10^3$	
Viral	Bound	151×10^3	14.7×10^3	>93
	Unbound	164×10^3	$<1 \times 10^3$	

[a] Cells were labeled by exposure to 100 μCi of [2-^{3}H]adenosine per ml for 1 h at 28 h after infection. Nuclei and cytoplasm were separated as described in the text, except that lysis was done in buffer containing 140 mM instead of 10 mM NaCl. Purified RNAs were fractionated by oligo(dT)-cellulose chromatography. Bound and unbound fractions were preparatively hybridized to denatured polyomavirus DNA bound to nitrocellulose filters, and hybridized RNA was eluted, as described (1). Samples of RNA which did not hybridize (cellular RNA) or which hybridized and were eluted (viral RNA) were treated with DNAse and then digested with 10 U of RNAse T_1 per ml at 37°C for 30 min in 100 mM NaCl–10 mM EDTA–10 mM triethanolamine-hydrochloride (pH 7.4). Digested RNAs were analyzed by electrophoresis in 10% polyacrylamide slab gels in Tris-borate-EDTA buffer (5). Two-millimeter gel slices (60 μl volume) were treated with 250 μl of Protosol (New England Nuclear) at 37°C overnight and then were counted in 2 ml of toluene-based scintillation fluid. Counts per minute of ^{3}H in poly(A) were determined by summing all radioactivity in fractions in the region of the 230-nucleotide poly(A) peak, after subtraction of background radioactivity.

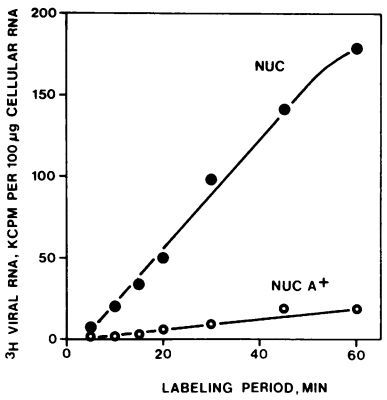

FIG. 1. Incorporation of [^{3}H]uridine into total and polyadenylated viral RNA in the nucleus during the late phase of infection. Cells were labeled with 500 µCi of [^{3}H]uridine per ml per petri dish for the times shown, beginning 28 h after infection. RNA was extracted and purified from nuclei prepared as described and fractionated by oligo(dT)-cellulose chromatography. RNAs were hybridized with an excess of filter-bound viral DNA, and results were normalized as described in the text. Symbols: ●, total viral RNA in nucleus; ○, polyadenylated viral RNA in nucleus.

mulate in polyadenylated viral RNA at a very slow rate, which would increase with labeling time. The results shown in Fig. 1 suggest that this is not the case since the rate of accumulation of [^{3}H]uridine in polyadenylated viral RNA appears to have been constant over 60 min. Furthermore, Fig. 2 shows that the delay between synthesis and polyadenylation was similar to that reported in other systems. In this figure, radioactivity in total and polyadenylated nuclear viral RNA is plotted for labeling periods of 5 to 15 min, and the straight lines most closely approximating the constant rates of incorporation are extrapolated to zero incorporation. The scale for polyadenylated RNA is expanded fivefold over that for total viral RNA, so the two straight lines are nearly parallel in this figure (approximately 20% of the RNA was polyadenylated in this experiment). It can be seen that the rate of labeling of total nuclear viral RNA became constant within 5 min and extrapolated to zero labeling at about 2.5 min. The rate of labeling of polyadenylated viral RNA extrapolated to zero incorporation at about 3.75 min. The two curves are thus separated by a delay of about 1.25 minutes, the average interval between synthesis of a viral RNA molecule and its polyadenylation (in three experiments this value was between 1 and 2 min). The results presented in Fig. 1 and 2 further suggest that if an RNA molecule is not polyadenylated within a short time of its synthesis, it never becomes polyadenylated; if a significant fraction of viral RNA molecules were polyadenylated more slowly, the accumulation curve for polyadenylated RNA would not be a straight line but would curve upwards.

Size of polyadenylated and nonpolyadenylated viral RNA in the nucleus. Polyadenylated and nonpolyadenylated nuclear RNAs were subjected to sucrose density gradient sedimentation, and portions of each gradient fraction were hybridized to an excess of polyomavirus DNA. The results, plotted in Fig. 3, show that polyadenylated viral RNA exists as molecules of discrete sizes, with sedimentation coefficients of about 19S, 33S, 41S, 47S, and perhaps more rapidly sedimenting species. On the other hand, nonpolyadenylated RNA sedimented heterogeneously between 10S and about 70S, no specific peaks being apparent. This agrees well with our previous determinations of the size of total pulse-labeled viral RNA (1, 3), most of which we now know to be nonpolyadenylated.

The lengths of polyadenylated viral RNAs in the different size classes can be estimated from their sedimentation rates relative to those of ^{14}C-labeled 18S and 28S rRNA markers in the same gradients as well as to those of 45S and 32S ribosomal precursor RNAs (data not shown) present in the nonpolyadenylated cellular RNA. Results of these calculations are presented in Table 2. Polyadenylated viral RNAs represent a family of molecules which differ in size by an integral number of viral genome lengths (about 5.3 kb). This is most easily explained if these RNAs all have their 5′ and 3′ ends at the positions on the viral genome where the major capped 5′ and polyadenylated 3′ ends of viral L-strand RNA have been mapped (nucleotides 5076 to 5127 and nucleotide 2900, respectively [13, 15, 37, 38]) but contain, in addition, 0 (19S), 1 (33S), 2 (41S), 3 (47S) or more full transcripts of the circular viral genome. These transcripts could arise if a certain proportion of the RNA polymerase molecules which transcribe the L strand did not terminate after passing the polyadenylation site but continued to transcribe the L strand until encountering the polyadenylation signal a second, third, or fourth time or more. I have previously shown that, in fact, many viral RNAs in the nucleus contain multiple, tandemly repeated copies of the nucleotide sequence in the entire L strand and that these copies probably arise by inefficient termination of transcription on this DNA strand (1).

Knowledge of the number of polyadenylated RNA molecules of different lengths made in the cell nucleus should reveal the frequency of transcription termination, or alternatively, of polyadenylation, in this system (see below). For this purpose, the total radioactivities in RNAs of the different size classes were estimated from graphs of the results shown in Fig. 3B and C by distributing the counts per minute between each set of peaks to one or the other species, depending on relative peak height. The results, shown also in Table 2, were divided by the estimated length of the RNA in

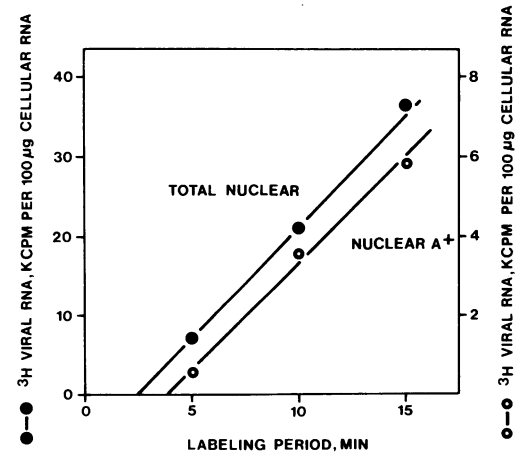

FIG. 2. Interval between synthesis and polyadenylation of viral RNA in the nucleus. Data from an experiment similar to that shown in Fig. 1 are plotted so that the initial slopes of the two curves are equal (by setting the scale for total RNA five times that for polyadenylated RNA). Lines drawn through the experimental points extrapolate to zero incorporation at about 2.5 min (●, total viral RNA) and about 3.75 min (○, polyadenylated viral RNA).

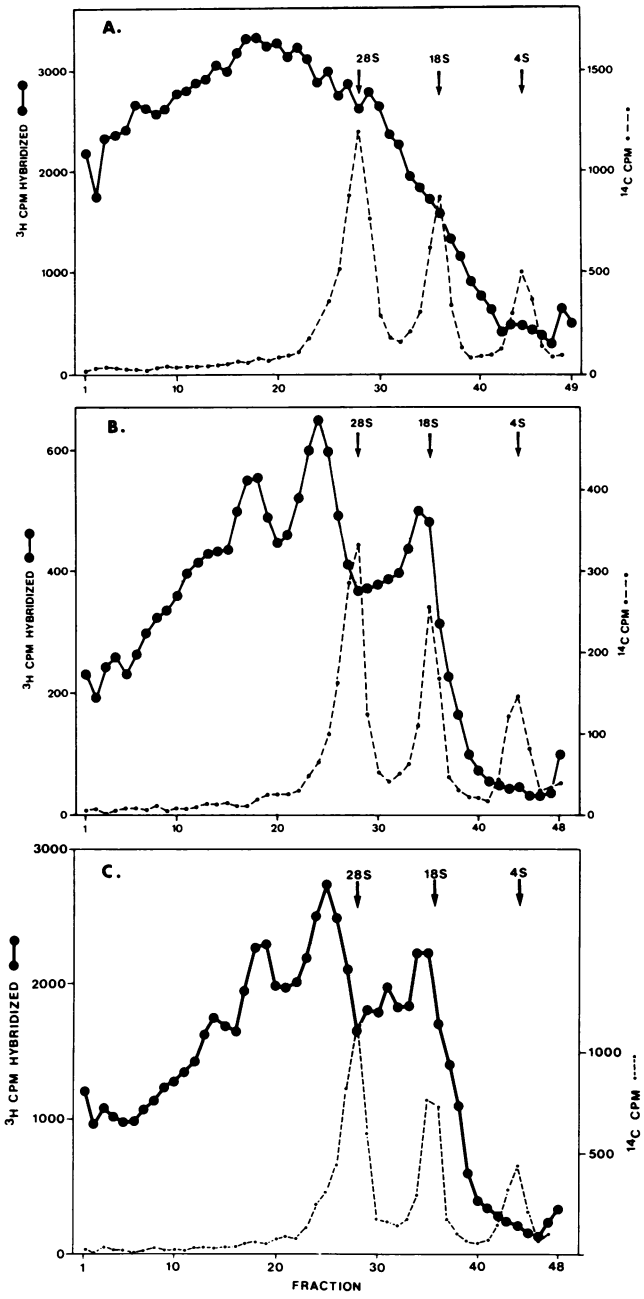

FIG. 3. Size of polyadenylated and nonpolyadenylated viral RNA in the nucleus. Cells were labeled with 500 μCi [^{3}H]uridine per ml per petri dish, beginning 28 h after infection. Nuclei were prepared by lysis of cells in 25 mM NaCl–5 mM $MgCl_2$–10 mM triethanolamine-hydrochloride (pH 7.4)–250 mM sucrose–1% Nonidet P-40–0.5% sodium deoxycholate. RNA was extracted and purified and then was fractionated by oligo(dT)-cellulose chromatography. RNAs were denatured by heating at 55°C for 3 min in 80% (vol/vol) dimethylsulfoxide–0.2 mM EDTA–2 mM triethanolamine-hydrochloride (pH 7.4) and then were ethanol precipitated. RNAs were suspended in 100 μl of 10 mM triethanolamine-hydrochloride (pH 7.4)–50 mM LiCl–0.1% sodium dodecyl sulfate to which was added [^{14}C]uridine-labeled mouse kidney cell RNA as a sedimentation marker. RNA were then layered onto 3.8-ml 15 to 30% (wt/wt) linear sucrose gradients in the same buffer. Sedimentation was for 2 h at 55,000 rpm in a Spinco SW-60 rotor at 20°C. Fractions were collected from the tube bottom. Samples were spotted onto filter paper disks which were washed in 5% trichloroacetic acid and counted as described (4). The remainder of each fraction was hybridized with filter-bound polyomavirus DNA as described in the text. Results are expressed as counts per minute per fraction. (A) Nonpolyadenylated nuclear RNA labeled for 20 min; (B) polyadenylated nuclear RNA labeled for 20 min; (C) polyadenylated nuclear RNA labeled for 60 min. Solid lines, ^{3}H counts per minute in viral RNA; broken lines, ^{14}C counts per minute in mouse 4S, 18S, and 28S RNA.

each peak, and the radioactivities per unit of length were normalized to a value of 1.00 for 19S RNA. These normalized values should reflect the relative numbers of molecules labeled in each size class of polyadenylated viral RNA. It can be seen from the numbers in the righthand column in Table 2 that each size class of RNA contains about 0.4 times as many molecules as the size class immediately preceding it, one genome length smaller. This is consistent with a model in which each RNA polymerase molecule has a probability of terminating of about 0.6 per genome length of DNA traversed during the transcription of the L strand, assuming that each terminated RNA molecule has an equivalent probability of being polyadenylated (see below for a detailed justification of this model).

It is notable that the profiles of polyadenylated viral RNA shown in Fig. 3 and the profiles of nonpolyadenylated cellular RNA, which contains 45S and 32S ribosomal precursor RNAs (data not shown), strongly indicate that little or no degradation of the RNA occurred during isolation in these experiments. Significant degradation would give rise to smearing of peaks and to accumulation of low-molecular-weight RNAs, which were not seen.

Transport of polyadenylated viral RNA from nucleus to cytoplasm. Figure 4 shows the accumulation of [^{3}H]uridine in polyadenylated viral RNA in the cytoplasm and in the nucleus as a function of labeling time. It can be seen that polyadenylated viral RNA entered the cytoplasm as two kinetically distinct components, a minor component which entered with no delay and a major component which enters after a significant delay. The minor component represents about 5% of the total radioactivity in polyadenylated viral RNA in the cell; it is likely that most of this component represents nuclear leakage. Previous experiments showed that 5 to 6% of unfractionated viral RNA was found in the cytoplasmic extract after a 15-min labeling period. Significantly lower levels of viral RNA (<1%) entered the cytoplasmic extract when cell lysis was carried out in an isotonic pH 7.4 buffer, but yields of cytoplasm were also lower (2). In the present work, 5% of nonpolyadenylated viral RNA in the cell was found in the cytoplasmic extract after labeling for periods of 5 to 60 min (data not shown). I conclude that about 5% of viral RNA, both polyadenylated and nonpolyadenylated, leaks from nuclei of cells lysed as in this experiment. However, it is not excluded that a small amount of viral mRNA enters the cytoplasm more rapidly than the major component does.

This major component began to enter the cytoplasm about 26 min after the start of labeling, or about 22 min (25 min in a second experiment) after the first appearance of label in nuclear polyadenylated RNA (by extrapolation; see above). These values agree with those determined previously for the interval between synthesis of total late viral RNA and its transport to the cytoplasm as mRNA (2).

The data in Fig. 4 can be used to calculate the proportion of sequences in polyadenylated viral RNA which are conserved during processing and transport to the cytoplasm,

TABLE 2. Relative molar amounts of different size classes of viral polyadenylated RNAs in the nucleus[a]

Labeling time (min)	Sedimentation coefficient	Estimated length (kb)	cpm	cpm/length, normalized[b]	Ratios, cpm/length[c]
20	19S	2.2	2,400	1.00	
	33S	7.5	3,500	0.43	0.43
	41S	12.8	2,200	0.16	0.37
60	19S	2.2	12,600	1.00	
	33S	7.5	17,950	0.41	0.41
	41S	12.8	11,950	0.16	0.39
	47S	18	7,425	0.07	0.44

[a] Values determined from experiment shown in Fig. 3, as described in the text.

[b] Values of the total radioactivity in each peak were divided by the estimated length of molecules in that size class, and the results were normalized, taking the value for the 19S RNA as 1.00. These normalized values are proportional to the number of molecules in each size class.

[c] Counts per minute per unit of length for the size class denoted, divided by counts per minute per unit of length for the size class 5.3 kb smaller.

assuming that all cytoplasmic polyadenylated mRNA is derived from nuclear polyadenylated RNA. This proportion is calculated by dividing the rate of entry of radioactivity into polyadenylated mRNA in the cytoplasm by the rate of incorporation of radioactivity into polyadenylated RNA in the nucleus. These values are proportional to the initial slopes of the two incorporation curves of Fig. 4, after correction of the major cytoplasmic component by subtraction of the slope of the minor component, assumed to be nuclear leakage. Two independent experiments yielded values of 0.29 and 0.35 for the proportion of nuclear polyadenylated RNA transformed into mRNA. These are probably underestimates since not all of the cytoplasm is recovered in the cytoplasmic extract; a maximum of about 20% is present in the nuclear extract (2). Correction for this factor gives values of 0.36 and 0.44 for the maximum conservation of sequences in polyadenylated viral RNA during processing. The major source of error in the calculation of these values, besides the correction just referred to, is the determination of the slopes of the incorporation curves, which is precise to within about 20%. These results, therefore, suggest that a minimum of 30% and a maximum of 50% of the sequences in polyadenylated viral RNA are conserved during processing and transport to the cytoplasm.

Size of polyadenylated viral RNA in the cytoplasm. Samples of cytoplasmic polyadenylated RNA labeled for different times were sedimented on sucrose density gradients, and portions of each fraction were hybridized with polyomavirus DNA. The results are shown in Fig. 5, along with sedimentation profiles of ^{14}C-labeled 4S, 18S, and 28S mouse cell RNA markers present in each gradient. Two observations can be made from these results. First, a significant proportion of the radioactivity in the samples labeled for 15, 20, and 30 min sedimented faster than the 18S-plus-19S and the 16S species, which are the known mRNA species produced in the late phase of productive infection (8, 15, 16). These faster-sedimenting RNAs resemble nuclear polyadenylated viral RNAs (see Fig. 3) and most probably represent nuclear leakage into the cytoplasmic extract. Even in the sample labeled for 15 min, these RNAs accounted for less than half the radioactivity in polyadenylated viral RNA in the cytoplasm. Thus, some processed viral mRNA, both 16S and 18S plus 19S, reaches the cytoplasm as soon as 15 min after the beginning of labeling.

Second, 18S-plus-19S species predominated in samples labeled for short times, but 16S RNA became a larger and larger proportion of the total mRNA population as labeling time increased. Others have shown that most late polyomavirus mRNA is derived from the late region of the L DNA strand and that an "18S" mRNA (which codes for capsid protein VP3) and a "19S" mRNA (which codes for VP2) are both present in the 18S-plus-19S region in sucrose gradients (15, 34). The 16S mRNA, which codes for VP1, is the major viral mRNA species at steady state or in RNAs labeled (8, 16, 28) for long periods. Piper et al. (28) have shown that all three L-strand mRNAs have long half-lives (>10 h). My results, therefore, suggest that 18S, or 19S mRNA or both begin to be transported to the cytoplasm with a shorter delay, but are produced at a lower overall rate, than 16S mRNA, thus accounting for the predominance of 16S mRNA at steady state. A more detailed analysis of the arrival times of the three L-strand mRNAs could be achieved by S1 nuclease mapping (6, 28) of pulse-labeled cytoplasmic RNAs; I have not yet attempted such experiments.

DISCUSSION

Two previous reports have suggested that only a small proportion of late polyomavirus RNA in the nucleus is polyadenylated (7, 30). However, in neither report was it demonstrated that the methods used for assay of polyadenylated viral RNA were quantitative. This is a potential problem, especially when dealing with high-molecular-weight nuclear RNAs (23, 31). In addition, relatively long labeling times were employed in both reports (30 min in reference 30, 5 h in reference 7), which left some doubt whether their results reflected actual rates of polyadenylation or were perhaps a result of processing or degradation of RNAs. It is known, for example, that mRNA precursors are polyadenylated within 1 to 2 min of their synthesis and that export of

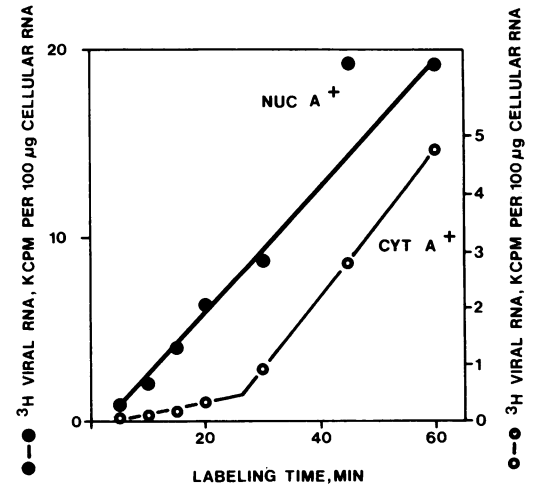

FIG. 4. Incorporation of [^{3}H]uridine into polyadenylated viral RNA in the nucleus and the cytoplasm. See legend to Fig. 1 for details. Symbols: ●, polyadenylated viral RNA in nucleus (scale to left of figure); ○, polyadenylated viral RNA in cytoplasm (expanded scale, to right of figure).

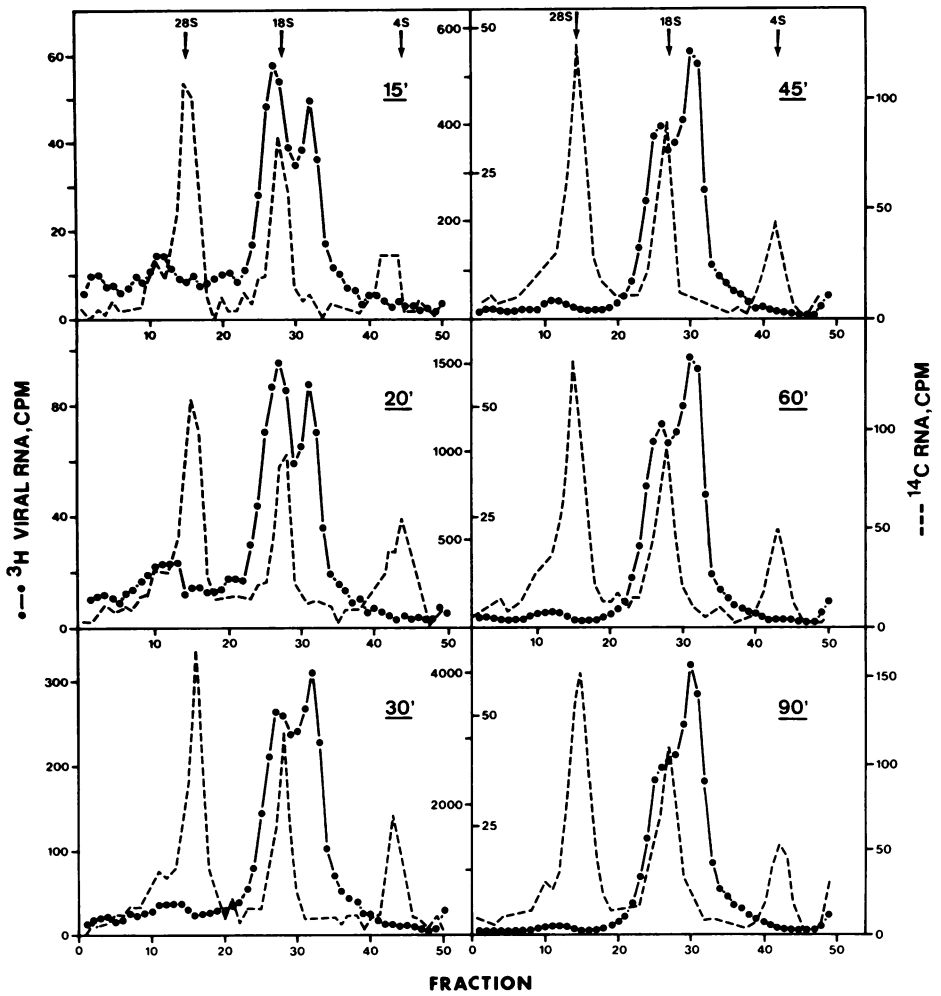

FIG. 5. Sizes of polyadenylated viral RNA in the cytoplasm after labeling for different lengths of time. RNA was labeled, extracted, and fractionated as described in the legend to Fig. 3. RNAs were denatured by heating at 80°C for 2 min in 100 μl of 10 mM triethanolamine-hydrochloride (pH 7.4)–50 mM LiCl–0.1% sodium dodecyl sulfate and were sedimented as described in the legend to Fig. 3 for 3 h and 15 min at 59,000 rpm. Portions of each fraction were counted and hybridized as described. RNAs were labeled for the times indicated.

mRNAs to the cytoplasm can occur within 5 to 10 min after synthesis of their precursors (2, 26, 32, 39).

My results (Fig. 1 and 2) indeed demonstrate that only 10 to 25% of viral RNA in the nucleus is polyadenylated. Greater than 90% of polyadenylated viral and cellular RNAs are recovered by oligo(dT)-cellulose chromatography (Table 1); degradation of RNA during extraction and preparation is minimal (Fig. 3 and 5). Furthermore, the kinetics of incorporation of [^{3}H]uridine into total and polyadenylated viral RNA in the nucleus show that viral RNAs are polyadenylated within 1 to 2 min of their synthesis, as has been shown for other mRNA precursors (26, 32). These results also suggest that if an RNA molecule does not become polyadenylated within a short time after its synthesis, it never becomes polyadenylated. If slow polyadenylation of a significant fraction of viral RNA took place, the proportion of polyadenylated to total viral RNA would increase with time; the two curves are parallel for at least 40 min.

Polyadenylated viral RNAs in the nucleus are a family of molecules with a length of approximately 2.2 + n(5.3) kb, where n can be an integer from 0 to at least 3. Since the major 5′ ends of L-strand RNAs are near nucleotide 5100 and the only known polyadenylated 3′ end is near nucleotide 2900, 2,200 nucleotides downstream (13, 15, 37, 38), this family must be constituted of RNAs which begin near nucleotide 5100 and end either the first time the polyadenylation site is encountered on the L DNA strand (n = 0) or after an integral number of additional genome traverses (n = 1, 2, 3 . . .). This is formally analogous to cellular transcription units which have more than one polyadenylation site and where poly(A) addition can take place either at the first site or at sites farther downstream (18, 26, 36). Two constraints necessarily hold true in such a miltisite polyadenylation system. First, RNA polymerase must not terminate efficiently in the region between adjacent polyadenylation sites. Second, the polyadenylation mechanism, which may include cleavage of a larger precursor RNA (19, 20, 25, 26), must not always act as the first possible polyadenylation site becomes available on a growing RNA chain.

Table 2 shows that each size class of RNA contains about 0.4 times as many molecules as the size class one genome traverse smaller. Such a frequency distribution could be generated if either transcription termination (followed by polyadenylation near the 3′ terminus) or the polyadenylation

reaction itself took place with a probability of 0.6 each time a polyadenylation site, or some associated signal, is encountered on growing RNA chains. A general formulation for the distribution of molecules generated by such a system is $f_i = T \times (1 - T)^{(i-1)}$ where f_i is the fraction of total molecules produced which are polyadenylated at the ith site and T is the frequency of transcription termination or polyadenylation at each site as it is encountered ($0 < T < 1$) (Acheson, manuscript in preparation). For example, if $T = 0.6$, then $f_1 = 0.6$ (that is, 60% of the molecules will have their 3′ ends at the first site), $f_2 = 0.6 \times (1 - 0.6) = 0.24$ (that is, 24% of the molecules will have their 3′ ends at the second site), and so on. It is evident that each size class contains $(1 - T)$ times as many molecules as the size class immediately smaller; in this example, $(1 - T) = 1 - 0.6 = 0.4$. Admittedly, this is only one way in which such a family of molecules could be generated, but it has the advantage of being relatively simple, and it is consistent with the data I have. It is an unproven assumption that the distribution of RNAs in each size class reflects their relative rates of formation and is not perturbed by differential processing (e.g., splicing and transport to the cytoplasm) of certain species of polyadenylated RNA. The fact that this distribution is similar in RNA labeled for 20 min, when fully processed mRNA is just beginning to appear in the cytoplasm, and in RNA labeled for 60 min argues in favor of this assumption.

The low efficiency of polyadenylation of viral RNAs can adequately explain the low conservation of RNA during processing in the nucleus and transport to the cytoplasm as mRNA. Only about 5% of total viral RNA synthesized becomes mRNA during the late phase of productive infection (2). If only about 15% of nuclear RNA is polyadenylated, and is therefore a potential precursor to mRNA, a maximum of 15% conservation could be expected. But only one mRNA body region per precursor molecule is converted into mature mRNA, the remaining sequences being discarded during RNA splicing. This is based on the observation that each mRNA molecule contains a single mRNA body region, which is linked at its 5′ end to one or more tandemly joined 57-nucleotide leader RNAs (17, 21, 37). These multiple copies of the leader must arise by splicing between leader sequences present in precursor RNAs which contain multiple copies of the entire L DNA strand since only one leader region is present in the viral DNA. Thus, leader-to-leader splicing removes nearly an entire genome-length of RNA from a precursor RNA molecule. In addition, body-to-leader splicing, in the case of 16S and 18S L-strand mRNAs, removes additional sequences (15). Assuming that two molecules of 16S mRNA are produced per molecule of 18S-plus-19S RNA (reference 28 and see Fig. 5), splicing of the family of polyadenylated precursor RNAs described in this paper would lead to removal of about 70% of the mass of RNA sequences. Thus, only 30% of 15%, or 4.5% of the total viral RNA synthesized, would be expected to be conserved, and this is close to what was found. Thus, it would appear that those viral RNAs which do become polyadenylated are spliced and transported to the cytoplasm efficiently.

Miller et al. (22) found that transcription of simian virus 40 DNA injected into *Xenopus* oocytes also led to the production of large, heterogeneous viral RNAs, only a fraction of which are polyadenylated. It appears that termination of transcription and polyadenylation of viral transcripts are both inefficient in that system. (In fact, the authors ascribe their results to low rates of processing, but it is likely that in their system processing takes place at normal rates but at reduced efficiencies. Certainly in polyomavirus-infected mouse cells, polyadenylation, splicing, and transport of RNA to the cytoplasm all take place at rates similar to those determined in a variety of mammalian cells growing at 37°C [10].) It is intriguing that inefficient termination of transcription and inefficient polyadenylation occur together in these two systems. In contrast, simian virus 40 late transcription is efficiently terminated in productively infected monkey cells (12), and conservation of viral RNA sequences as mRNA is high, indicating efficient polyadenylation of nuclear RNAs (9). The same can be said for early polyomavirus transcription (4).

These considerations lead me to speculate that there may be a necessary relationship between transcription termination and polyadenylation. This relationship could be either of two kinds. On the one hand, cleavage and polyadenylation of a growing RNA chain might be necessary for efficient termination to take place subsequently. One could imagine that once the nascent RNA chain is cleaved, the RNA polymerase would become destabilized and would fall off the DNA template. This could happen within a minute or so of the passage of the RNA polymerase over the polyadenylation site, at which time the RNA polymerase might be 1,000 to 2,000 nucleotides downstream from this site. It is of interest that termination takes place efficiently within this distance downstream of the polyadenylation sites both on the simian virus 40 L strand (12) and on the mouse B-globin gene (14), although there is no evidence that termination occurs at a specific nucleotide. RNA polymerase does not terminate shortly downstream from polyadenylation sites in the major late transcription unit of adenovirus type 2 DNA (26). This may, however, be a special case since termination is much more efficient in this transcription unit in the early than in the late phase of infection (10, 27, 33) and thus may be under control of a viral product. In any event, if polyadenylation must precede termination, inefficient polyadenylation would result in inefficient termination. Such a mechanism would not, however, explain the production of large amounts of nonpolyadenylated polyomavirus RNA for these RNAs appear to be terminated without being polyadenylated.

Alternatively, termination could be necessary step before an RNA can be polyadenylated. In this case, an RNA would be a substrate for cleavage and polyadenylation only after a free 3′ end is created by termination of transcription and release from the RNA polymerase and the template DNA. This mechanism could explain the production of RNAs with multiple polyadenylation sites, of which only the most distal is utilized (18, 36). (Again, adenovirus type 2 late RNAs would be an exception.) Cleavage and polyadenylation would take place at the first site encountered upstream of the 3′ end created by termination. The model is also consistent with the occurrence of termination some distance downstream from the polyadenylation sites in simian virus 40 and mouse β-globin genes. Inefficient polyadenylation of polyomavirus RNAs could be explained if termination is required to take place at certain sites or a certain distance downstream from the polyadenylation site; RNAs terminated elsewhere would escape polyadenylation.

For both of these models, polyadenylation would be expected to be rapid, either taking place on the nascent RNA chain (in the first model) or taking place soon after termination occurs (in the second model). Once the RNA polymerase passes a certain distance beyond a polyadenylation site without cleavage and polyadenylation (the first model) or termination (the second model) taking place, that site would no longer be available to the polyadenylation mechanism.

ACKNOWLEDGMENTS

This work was begun in the Department of Virology at the Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland (supported by the Swiss National Science Foundation, grants 3.691.76 and 3.299.78) and was finished at McGill University (supported by the Medical Research Council, grant MA-7281).

I thank Francoise Miéville for excellent technical assistance.

LITERATURE CITED

1. **Acheson, N. H.** 1978. Polyoma giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proc. Natl. Acad. Sci. U.S.A. **75**:4754–4758.
2. **Acheson, N. H.** 1981. Efficiency of processing of viral RNA during the early and late phases of productive infection of polyoma virus. J. Virol. **37**:628–635.
3. **Acheson, N. H., E. Buetti, K. Scherrer, and R. Weil.** 1971. Transcription of the polyoma virus genome: synthesis and cleavage of giant late polyoma-specific RNA. Proc. Natl. Acad. Sci. U.S.A. **68**:2231–2235.
4. **Acheson, N. H., and F. Miéville.** 1978. Extent of transcription of the E strand of polyoma virus DNA during the early phase of productive infection. J. Virol. **28**:885–894.
5. **Adesnik, M., and J. E. Darnell.** 1972. Biogenesis and characterization of histone messenger RNA in HeLa cells. J. Mol. Biol. **67**:397–406.
6. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
7. **Birg, F., J. Favaloro, and R. Kamen.** 1977. Analysis of polyoma virus nuclear RNA by mini-blot hybridization. Proc. Natl. Acad. Sci. U.S.A. **74**:3138–3142.
8. **Buetti, E.** 1974. Characterization of late polyoma mRNA. J. Virol. **14**:249–260.
9. **Chiu, N. H., M. F. Radonovich, M. M. Thoren, and N. P. Salzman.** 1978. Selective degradation of newly synthesized nonmessenger simian virus 40 transcripts. J. Virol. **28**:590–599.
10. **Darnell, J. E., Jr.** 1982. Variety in the level of gene control in eukaryotic cells. Nature (London) **297**:365–371.
11. **Deininger, P. L., A. Esty, P. LaPorte, H. Hsu, and T. Friedmann.** 1980. The nucleotide sequence and restriction enzyme sites of the polyoma genome. Nucleic Acids Res. **8**:855–860.
12. **Ford, J. P., and M.-T. Hsu.** 1978. Transcription pattern of in vivo-labeled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3′ terminus. J. Virol. **28**:795–801.
13. **Heiser, W. C., and W. Eckhart.** 1982. Polyoma virus early and late mRNAs in productively infected mouse 3T6 cells. J. Virol. **44**:175–188.
14. **Hofer, E., R. Hofer-Warbinek, and J. E. Darnell, Jr.** 1982. Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell **29**:887–893.
15. **Kamen, R., J. Favaloro, and J. Parker.** 1980. Topography of the three late mRNA's of polyoma virus which encode the virion proteins. J. Virol. **33**:637–651.
16. **Kamen, R., and H. Shure.** Topography of polyoma virus messenger RNA molecules. Cell **7**:361–371.
17. **Legon, S., A. J. Flavell, A. Cowie, and R. Kamen.** 1979. Amplification of the leader sequence of "late" polyoma virus mRNAs. Cell **16**:373–388.
18. **Maki, R., W. Roeder, A. Traunecker, C. Sidman, M. Wabl, W. Raschke, and S. Tonegawa.** 1981. The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. Cell **24**:353–365.
19. **Manley, J. L.** 1983. Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. Cell **33**:595–605.
20. **Manley, J. L., P. A. Sharp, and M. L. Gefter.** 1982. RNA synthesis in isolated nuclei: processing of adenovirus serotype 2 late messenger RNA precursors. J. Mol. Biol. **159**:581–599.
21. **Manor, H., M. Wu, N. Baran, and N. Davidson.** 1979. Electron microscopic mapping of RNA transcribed from the late region of polyoma virus DNA. J. Virol. **32**:293–303.
22. **Miller, T. J., D. L. Stephens, and J. E. Mertz.** 1982. Kinetics of accumulation and processing of simian virus 40 RNA in *Xenopus laevis* ooctyes injected with simian virus 40 DNA. Mol. Cell. Biol. **2**:1581–1594.
23. **Molloy, G. R., W. Jelinek, M. Salditt, and J. E. Darnell, Jr.** 1974. Arrangement of specific oligonucleotides within poly(A) terminated mRNA molecules. Cell **1**:45–53.
24. **Nakazoto, H., and E. Edmonds.** 1974. Purification of messenger RNA and heterogeneous nuclear RNA containing poly(A) sequences. Methods Enzymol. **29**:431–443.
25. **Nevins, J. R., J.-M. Blanchard, and J. E. Darnell, Jr.** 1980. Transcription units of adenovirus type 2. termination of transcription beyond the poly(A) addition site in early regions 2 and 4. J. Mol. Biol. **144**:377–386.
26. **Nevins, J. R., and J. E. Darnell, Jr.** 1978. Steps in the processing of Ad2 mRNA: poly(A^+) nuclear sequences are conserved and poly(A) addition precedes splicing. Cell **15**:1477–1493.
27. **Nevins, J. R., and M. C. Wilson.** 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. Nature (London) **290**:113–118.
28. **Piper, P., J. Wardale, and F. Crew.** 1979. Splicing of the late mRNAs of polyoma virus does not occur in the cytoplasm of the infected cell. Nature (London) **282**:686–691.
29. **Rosenthal, L. J.** 1976. Isolation and characterization of poly(A)-containing polyoma "early" and "late" messenger RNAs. Nucleic Acids Res. **3**:661–676.
30. **Rosenthal, L. J., C. Salomon, and R. Well.** 1976. Isolation and characterization of poly(A)-containing intranuclear polyoma-specific "giant" RNAs. Nucleic Acids Res. **3**:1167–1183.
31. **Salditt-Georgieff, M., and J. E. Darnell, Jr.** 1982. Further evidence that the majority of primary nuclear RNA transcripts in mammalian cells do not contribute to mRNA. Mol. Cell. Biol. **2**:701–707.
32. **Salditt-Georgieeff, M., M. Harpold, S. Sawicki, J. Nevins, and J. E. Darnell, Jr.** 1980. The addition of poly(A) to nuclear RNA occurs soon after RNA synthesis. J. Cell Biol. **86**:844–848.
33. **Shaw, A. R., and E. B. Ziff.** 1980. Transcripts from the adenovirus-2 major late promoter yield a single early family of 3′ coterminal mRNAs and five late families. Cell **22**:905–916.
34. **Siddell, S. G., and A. E. Smith.** 1978. Polyoma virus has three late mRNA's: one for each virion protein. J. Virol. **27**:427–431.
35. **Soeda, E., J. A. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**:445–453.
36. **Tosi, M., R. A. Young, O. Hagenbuchle, and U. Schibler.** 1981. Multiple polyadenylation sites in a mouse α-amylase gene. Nucleic Acids Res. **9**:2313–2323.
37. **Treisman, R.** 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. **8**:4867–4888.
38. **Treisman, R., and R. Kamen.** 1981. Structure of polyoma virus late nuclear RNA. J. Mol. Biol. **148**:273–301.
39. **Wilson, M. C., S. G. Sawicki, M. Salditt-Georgieff, and J. E. Darnell.** 1978. Adenovirus type 2 mRNA in transformed cells: map positions and difference in transport time. J. Virol. **25**:97–103.