JOURNAL OF VIROLOGY, Nov. 1979, p. 536-545
0022-538X/79/11-0536/10$02.00/0

Vol. 32, No. 2

Evidence for the Identity of Shared 5'-Terminal Sequences Between Genome RNA and Subgenomic mRNA's of B77 Avian Sarcoma Virus

C. MARTIN STOLTZFUS†* AND LEE K. KUHNERT

Department of Microbiology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received for publication 10 May 1979

The polyribosomal fraction from chicken embryo fibroblasts infected with B77 avian sarcoma virus contained 38S, 28S, and 21S virus-specific RNAs in which sequences identical to the 5'-terminal 101 bases of the 38S genome RNA were present. The only polyadenylic acid-containing RNA species with 5' sequences which was detectable in purified virions had a sedimentation coefficient of 38S. This evidence is consistent with the hypothesis that a leader sequence derived from the 5' terminus of the RNA is spliced to the bodies of the 28S and 21S mRNA's, both of which have been shown previously to be derived from the 3' terminal half of the 38S RNA. The entire 101-base 5' terminal sequence of the genome RNA appeared to be present in the majority of the subgenomic intracellular virus-specific mRNA's, as established by several different methods. First, the extent of hybridization of DNA complementary to the 5'-terminal 101 bases of the genome to polyadenylic acid-containing subgenomic RNA was similar to the extent of its hybridization to 38S RNA from infected cells and from purified virions. Second, the fraction of the total cellular polyadenylic acid-containing RNA with 5' sequences was similar to the fraction of RNA containing sequences identical to the extreme 3' terminus of the genome RNA when calculated by the rate of hybridization of the appropriate complementary DNA probes. This suggests that most intracellular virus-specific RNA molecules contain sequences identical to those present in the 5'-terminal 101 bases of the genome. Third, the size of most of the radioactively labeled DNA complementary to the 5'-terminal 101 bases of the genome remained unchanged after the probe was annealed to either intracellular 38S RNA or to various size classes of subgenomic RNA and the hybrids were digested with S1 nuclease and denatured with alkali. However, after this procedure some DNA fragments of lower molecular weight were present. This was not the case when the DNA complementary to the 5'-terminal 101 bases of the genome was annealed to 38S genome RNA. These results suggest that, although the majority of the intracellular RNA contains the entire 101-base 5'-terminal leader sequence, a small population of virus-specific RNAs exist that contain either a shortened 5' leader sequence or additional splicing in the terminal 101 bases.

The genome of avian sarcoma viruses (ASV) is composed of two apparently identical 38S RNA subunits (2, 6). Within infected cells, however, a number of different size classes of virus-specific RNA molecules have been reported, namely 38S, 28S, and 21S (10, 16, 32). In addition, the presence of 10S to 12S virus-specific RNA has also been noted by some investigators (10). The intracellular 38S RNA contains most, if not all, of the sequences in the viral genome and probably is translated to form the non-glycosylated structural proteins of the virion core (*gag* proteins) (22) and the viral reverse transcriptase (*pol* protein) (20). The subgenomic 28S RNA contains the sequences coding for the virion envelope protein (*env*) and the sarcoma protein (*src*), whereas the 21S RNA contains only *src* sequences (16, 32). The 28S RNA is probably translated to form the envelope proteins (21). The gene product of *src* is responsible for the neoplastic transformation of fibroblasts and is presumably the translation product of the 21S RNA. *env* and *src* are contained in the 3' half and third of the ASV genome, respectively (18, 31).

Recent experiments have supported the hy-

† Present address: Department of Microbiology, University of Iowa, Iowa City, IA 52242.

pothesis that the subgenomic RNAs of avian retroviruses are formed by a cutting and splicing mechanism similar to that proposed for the generation of the mRNA's of adenoviruses (4, 11), papovaviruses (1), mouse globin (17), and ovalbumin (7). First, it was found that DNA complementary to the 5'-terminal 101-base sequence of the ASV genome RNA (5'-cDNA) hybridizes to RNA from each mRNA size class. This suggested that at least part of the 5'-terminal 101-base genome sequence is present in the smaller RNAs which contain sequences derived from the 3' half of the genome RNA (32). Second, it was shown that a 21-base 5'-terminal cap-containing T1 oligonucleotide, which is identical to that present at the 5' terminus of the genome RNA is found in at least some of the subgenomic virus-specific intracellular RNAs (19). This oligonucleotide was also present in the sequences which were protected from digestion by RNase T_1 when the subgenomic RNA was hybridized to 5'-cDNA. Several other oligonucleotides characteristic of the 5'-terminal 101 bases of the genome RNA were also present in the protected sequence (12). These data provide strong suggestive evidence that the smaller RNAs contain a 101-base 5'-terminal sequence identical to that at the 5' end of the genome RNA.

It remained a possibility, however, that there exist deletions of the terminal 101-base sequence in the smaller mRNA's which may not be detected by the oligonucleotide fingerprinting procedure. For instance, an initiation site for protein synthesis could be created by the splicing out of one or a few bases. It also was possible that the alleged spliced RNAs represent only a minor fraction of the total virus-specific RNA in the infected cell. Our approaches to answering these questions were the following: (i) to determine the rate and extent of hybridization of radioactive 5'-cDNA to polyadenylic acid [poly (A)]-containing RNAs of various sizes which were isolated from B77 sarcoma virus-infected cells; and (ii) to determine whether the 101-base 5'-cDNA which is annealed to virus-specific cellular RNA can be recovered intact after treatment with single-stranded nuclease S1 followed by alkali denaturation. We provide evidence here that most of the virus-specific intracellular RNA molecules contain the entire 101-base 5'-terminal genome sequence with no interruptions in this sequence.

MATERIALS AND METHODS

Virus propagation and purification. B77 ASV was propagated in secondary cultures of gs^- chf^- chicken embryo fibroblasts (SPAFAS, Inc., Roanoke, Ill.) as previously described (28). After the removal of cell debris, the virus was concentrated from the culture medium by using the polyethylene glycol-pronase procedure (8, 9) and purified on sucrose gradients (28).

Isolation and purification of viral and cellular RNA. Methods for isolation and purification of B77 sarcoma virus RNA have been described previously (28). To isolate total cellular RNA from infected cells, confluent monolayers of cells were detached from 10 to 20 100-mm plastic petri dishes by the addition of 0.08% (wt/vol) trypsin in a solution containing 140 mM NaCl, 5 mM KCl, 0.6 mM Na_2HPO_4, 5.5 mM glucose, and 25 mM Tris-hydrochloride, pH 7.2. These cells were labeled with 20 μCi of [^{3}H]uridine per ml for 12 h before the cells were harvested. The cells were sedimented for 5 min at 500 × *g* in a clinical centrifuge. To the sedimented cells was added 2 ml of a solution containing 50 mM sodium acetate, pH 5.1, 100 mM NaCl, 10 mM EDTA, 0.5% sodium dodecyl sulfate (SDS), and 50 U of sodium heparin per ml (NENSH buffer); this buffer also contained 250 μg of proteinase K per ml. After incubation at 37°C for 30 min, the disrupted cell preparation was extracted at room temperature with an equal volume of $CHCl_3$–phenol (1:1, vol/vol). After centrifugation at 5000 × *g*, the aqueous phase was removed. An equal volume of NENSH buffer was added to the phenol phase, and the extraction procedure was repeated. After precipitation with 2 volumes of 95% ethanol at −20°C, the nucleic acid was pelleted at 10,000 × *g* for 30 min, redissolved in 2 ml of a solution containing 10 mM sodium acetate, pH 5.1, 3 mM $MgCl_2$, and 50 U of sodium heparin per ml, and digested with 25 μg of DNase I per ml at 10°C for 30 min. After the addition of SDS to a final concentration of 1.0%, the RNA-containing aqueous phase was extracted with $CHCl_3$–phenol, and the RNA was precipitated with 2 volumes of 95% ethanol at −20°C. The RNA precipitate was pelleted at 10,000 × *g* and redissolved in 10 mM Tris-hydrochloride (pH 7.5)–1 mM EDTA (TE buffer). To the dissolved RNA was added an equal volume of 3 M LiCl, and the solution was incubated at 4°C for 12 to 24 h. The precipitate, containing primarily rRNA and mRNA, was pelleted at 10,000 × *g* for 30 min. The RNA pellet was washed with 70% ethanol containing 0.1 M NaCl and then dissolved in 1.0 ml of TE buffer, heated for 2 min at 80°C, quenched in ice, made 0.5 M NaCl, and applied to a column (5 by 20 mm) of oligodeoxythymidylic acid [oligo (dT)]-cellulose (type 2; Collaborative Research, Waltham, Mass.). The column was then washed with TE buffer containing 0.5 M NaCl to elute poly(A)$^-$ RNA and then with TE buffer to remove poly(A)$^+$ RNA. The fractions containing poly(A)$^+$ RNA were precipitated with 2 volumes of 95% ethanol. The precipitate was then dissolved in 0.2 ml of TE buffer containing 0.25% SDS, applied to 5 to 30% glycerol gradients, and centrifuged for 16 h at 23,000 rpm and 4°C in an SW41 rotor and a Beckman L5-65 centrifuge. Fractions (0.5 ml) were collected from the bottom of the centrifuge tube. A sample containing radioactive 18S and 28S rRNA and 38S B77 ASV genome RNA was also centrifuged in a parallel gradient and fractionated in a manner identical to that used for the poly(A)$^+$ cellular RNA. Fractions corresponding to various RNA sizes (see below) were pooled, precipitated with 2 volumes of 95% ethanol, and subjected to oligo(dT)-cellulose chromatography

as described above in order to remove fragments derived from the 5′ terminus of the 38S RNA. The poly(A)$^+$ RNA from each pool was then precipitated with ethanol, dissolved in 0.2 ml of sterile distilled water, and stored at −70°C. The specific activity of the RNA was estimated by measuring the amount of total RNA radioactivity in a known number of units of optical density at 260 nm.

Isolation of polyribosomal RNA. Confluent 100-mm plates of B77 ASV-infected cells were washed with ice-cold 0.25 M sucrose in a solution containing 25 mM NaCl, 100 mM Tris-hydrochloride, pH 7.1, 5 mM $MgCl_2$, and 80 U of sodium heparin per ml (TNM buffer). The cells were then resuspended in TNM buffer containing 1% (vol/vol) Nonidet P-40 and 1% (wt/vol) sodium deoxycholate and disrupted with 10 strokes of a tight pestle in a Dounce homogenizer. The preparation was centrifuged at 12,000 × *g* for 10 min to remove nuclei and cell debris. The cytoplasmic supernatant was layered over discontinous sucrose gradients (2 ml of 2.5 M sucrose in TNM buffer and 4 ml of 1 M sucrose in the same buffer). Centrifugation was carried out for 2.5 h at 35,000 rpm in an SW41 rotor and a Beckman L5-65 centrifuge. The polysomes, which accumulated at the interface of the 1 and 2.5 M sucrose solutions, were removed from the side of the tube with a syringe, diluted with an equal volume of 100 mM NaCl–1 mM EDTA–10 mM Tris-hydrochloride, pH 7.5, and brought to a concentration of 20 mM EDTA, 0.5% SDS, and 250 μg of proteinase K per ml. The sample was incubated at 37°C for 30 min and extracted with phenol–$CHCl_3$ (1:1, vol/vol) at room temperature, and the RNA was precipitated with 2 volumes of 95% ethanol.

Preparation of 5′-cDNA. B77 sarcoma virus complementary DNA (cDNA) was synthesized in a reaction mixture containing 70 mM NaCl, 57 mM Tris-hydrochloride, pH 8.2, 9 mM magnesium acetate, 2 mM dithiothreitol, 0.2% (vol/vol) Triton X-100, 0.5 mM each dCTP, dGTP, and TTP, 0.5 μM [α-^{32}P]-dATP (200 Ci/mmol), 0.1 mg of actinomycin D per ml, and B77 sarcoma virus at a concentration of 0.2 mg/ml. [^{3}H]dCMP-labeled 5′-cDNA was prepared as described above, except that dATP was present at a concentration of 0.5 mM and [5-^{3}H]dCTP was present at a concentration of 1 μM. The reaction was carried out at 37°C for 2 h. The reaction was stopped by the addition of SDS to a concentration of 1% (wt/vol). The solution was made 20 mM EDTA and was extracted with 90% phenol–$CHCl_3$ (1:1, vol/vol). The nucleic acid in the aqueous phase was precipitated with 2 volumes of 95% ethanol at −20°C. The precipitate was then centrifuged at 10,000 × *g* for 30 min, redissolved in 0.4 ml of 0.3 M KOH, and incubated for 18 h at 37°C. The sample was neutralized with an appropriate volume of 1 M HCl and chromatographed on a Sephadex G-50 column (60 by 0.9 cm) equilibrated with 100 mM NaCl–10 mM Tris-hydrochloride (pH 7.5)–1 mM EDTA. The cDNA, which eluted in the void volume of the column, was precipitated with 2 volumes of 95% ethanol. The DNA precipitate was dissolved in a solution containing 4 mM Tris-hydrochloride, pH 7.2, 2 mM sodium acetate, 1 mM EDTA, 25% glycerol, and 0.2% SDS (0.1× E buffer) and electrophoresed on a 10-cm 10% polyacrylamide gel containing 1× E buffer at 100 V for 3.5 h. The gel was then fractionated into 2-mm slices, eluted with 0.3 M sodium acetate containing 50% formamide, and incubated overnight at 37°C. Samples from the fractions were counted for radioactivity, and the fractions containing 5′-cDNA, which was present as a peak at a position corresponding to its size as previously described (14, 15), were filtered to remove gel particles, made 0.1 M NaCl, and precipitated with 2 volumes of 95% ethanol after the addition of 50 μg of yeast tRNA. The 5′-cDNA was stored as a precipitate at −20°C, and an appropriate amount of the material was pelleted before hybridization. The pellets were dissolved in sterile distilled water. A modification of the above procedure for the isolation of 5′-cDNA was also used. The cDNA was chromatographed on Sephadex G-50, heated to 80°C for 2 min, and electrophoresed on 10% polyacrylamide gels before alkaline hydrolysis in order to isolate the 5′-cDNA covalently attached to the tryptophan tRNA primer (Fig. 1A). This material was then treated with 0.3 M KOH in order to digest the primer tRNA and again electrophoresed on polyacrylamide gels to isolate the 5′-cDNA (Fig. 1B). No significant differences in the experimental results were obtained when 5′-cDNA was prepared by the two methods.

Preparation of ^{32}P-labeled 3′-cDNA. Oligo(dT)-primed ^{32}P-labeled cDNA was synthesized in a reaction mixture containing the following. 57 mM Tris-hydrochloride, pH 8.2, 9 mM magnesium acetate, 14 mM NaCl, dCTP, and TTP at concentrations of 0.5 mM, [α-^{32}P]dATP (200 Ci/mmol) at a concentration of 0.5 μM, 12 μg of oligo(dT) per ml, 2 mM dithiothreitol, 0.1 mg of actinomycin D per ml, avian myeloblastosis virus reverse transcriptase at a concentration of 40 U/ml, and 6 μg of B77 38S RNA per ml. The reaction was carried out at 37°C for 2 h. The cDNA was then subjected to alkaline hydrolysis and Sephadex G-50 chromatography as described above for the 5′-cDNA. The cDNA which contained oligo(dT) (approximately 70 to 80% of the product) was then separated from non-oligo(dT)-containing material by oligo(dT)-cellulose chromatography after annealing to poly(A) (5 μg/ml) at room temperature for 60 min (29).

Hybridization conditions. Hybridization of cDNA and RNA was carried out in a solution containing 0.3 M NaCl, 0.01 M Tris-hydrochloride, pH 7.5, 0.001 M EDTA, and 60 μg/ml of yeast tRNA per ml. Approximately 1,000 cpm (5 pg) of cDNA and various concentrations of RNA were used for each reaction. Polyuridylic acid (10-fold excess with respect to viral or cellular RNA) was added to the 3′-cDNA hybridization to avoid artifacts in the determination of the rate of annealing (29). The samples were incubated at 68°C for 18 h in sealed capillaries. The samples were then diluted into 0.4 ml of 30 mM sodium acetate–1 mM $ZnSO_4$–100 mM NaCl–5% glycerol containing 30 μg of calf thymus DNA and divided into two equal portions. To half of the sample was added 3,200 U of S1 nuclease; the sample was digested at 43°C for 30 min and precipitated with 5% CCl_3COOH. The other half of the sample was incubated at 43°C for 30 min and precipitated. Both samples were filtered onto nitrocellulose disks, and the filters were counted for

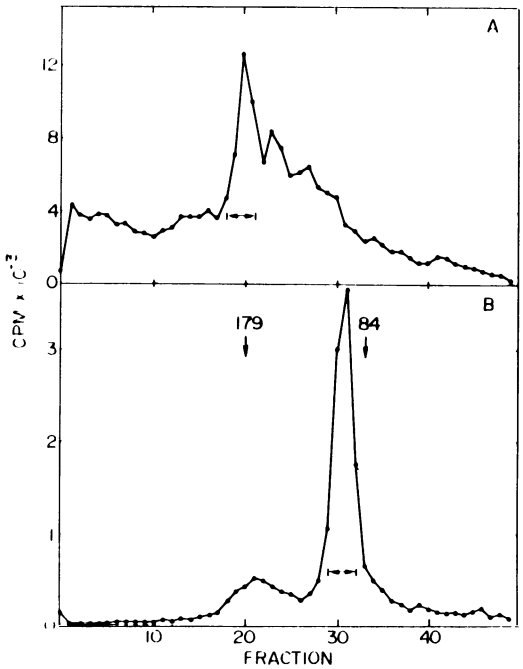

FIG. 1. *Isolation of 5'-cDNA from polyacrylamide gels. (A) B77 sarcoma virus cDNA was electrophoresed on 10% polyacrylamide gels and fractionated as described in the text. Samples from each eluted fraction were counted for radioactivity, the fractions indicated in the figure were combined, and the gel particles were removed by filtration. The cDNA was precipitated with 2 volumes of 95% ethanol after the addition of 50 μg of yeast tRNA. (B) The precipitated cDNA was dissolved in 0.4 ml of 0.3 M KOH, digested for 18 h, neutralized, and precipitated with 2 volumes of 95% ethanol after the addition of 50 μg of yeast tRNA. The sample was then subjected to gel electrophoresis and fractionated as described in the text. The peak fractions of radioactivity were combined, and the gel particles were removed by filtration. The cDNA was precipitated with 2 volumes of 95% ethanol after the addition of 50 μg of yeast tRNA. The positions of alkali-denatured HaeIII restriction fragments of ColE1 DNA (179 and 84 nucleotides) are shown on the figure.*

radioactivity. The percentage of the cDNA present in RNA-DNA hybrids was determined by the percentage of the radioactivity which was S1 resistant. The percent S1 resistance of cDNA without added viral or chicken embryo fibroblast RNA (5 to 7%) was subtracted as a blank.

Materials. [α-^{32}P]dATP (>1,000 Ci/mmol) was prepared by the enzymatic procedure of Walseth and Johnson (30a). [5-^{3}H]-dCTP (25 Ci/mmol) was purchased from ICN, Fullerton, Calif. [5,6-^{3}H]uridine (43 Ci/mmol) was purchased from Amersham Corp., Arlington Heights, Ill. S1 nuclease (840 × 10^3 u/mg) was from Miles Laboratories, Inc., Elkhart, Ind.; DNase I was from Boehringer Mannheim Corp., Indianapolis, Ind.; proteinase K was from EM Laboratories, Darmstadt, Germany. poly(A), polyuridylic acid, and oligo(dT)$_{12-18}$ were purchased from P-L Biochemicals, Milwaukee, Wis.

RESULTS

Presence of 5'-terminal genome sequences in poly(A)-containing polysomal RNA from B77 sarcoma virus-infected cells. To establish that viral RNA molecules containing sequences related to the 5'-terminal genome sequences of B77 ASV are utilized as mRNA's in infected cells, we isolated poly(A)-containing RNA from polyribosomes, separated it into size classes by sedimentation on glycerol

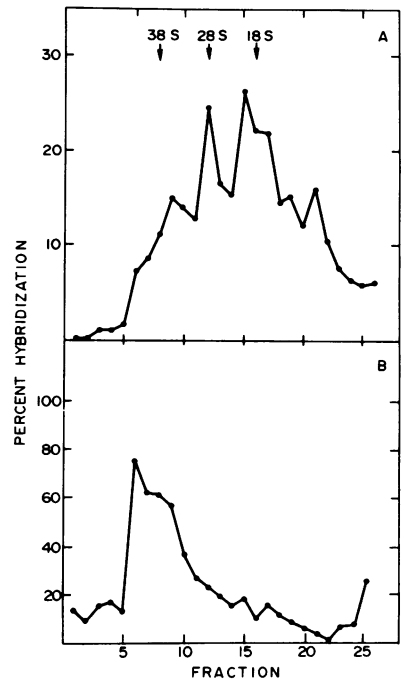

FIG. 2. *Hybridization of B77 sarcoma virus 5'-cDNA to polysomal RNA and genome RNA. Poly(A)-containing RNAs isolated by methods described in the text from the polyribosomes of B77 sarcoma virus-infected cells (A) and from B77 sarcoma virions harvested at 3-h intervals (5 μg of RNA) (B) were heated to 80°C for 2 min in TE buffer containing 0.25% SDS and centrifuged in 5 to 30% glycerol gradients at 24,000 rpm in an SW41 rotor and a Beckman L5-65 centrifuge for 18 h. The gradient fractions were precipitated with 2 volumes of 95% ethanol. The precipitates were dissolved in 200 μl of hybridization buffer. Hybridization of 20-μl samples of the RNA solutions with [^{3}H]dCMP-labeled 5'-cDNA was carried out for 18 h at 68°C according to the procedures described in the text. The percent resistance to S1 nuclease was determined and is plotted versus fraction number. Radioactive virion 38S genome RNA and chicken embryo fibroblast 28S and 18S rRNA's were centrifuged in a parallel gradient, and their positions are marked on the figure.*

gradients, and tested fractions from the gradient for complementarity to [^{3}H]dCMP-labeled 5′-cDNA (Fig. 2A).

Three species of virus-specific RNAs were present, with sedimentation coefficients of approximately 38S, 28S and 21S. This profile was similar to that described previously for the total RNA from cells infected by ASV (16, 32). We conclude that the subgenomic RNA species containing at least part of the 5′ sequences of the genome RNA are associated with polyribosomes and are presumably functioning as mRNA's. It has been shown that a sequence of 21 nucleotide bases at the 5′-terminal end of the avian retrovirus genome RNA is repeated at the 3′ end (24, 27). This terminally redundant sequence contains only 1 of the 23 guanosine residues in the 5′-terminal 101 bases (14, 15). Consequently, in the redundant sequence, the 5′-cDNA contains only 1 of the 23 total dCMP residues, or 4% of the total [^{3}H]dCMP-labeled 5′-cDNA radioactivity. The percentage of the 5′-cDNA that hybridized to each of the subgenomic RNA species was much greater than this, and, therefore, we concluded that the common sequences must include at least part of the nonredundant part of the 5′-cDNA.

We also tested the possibility that significant amounts of subgenomic mRNA's containing the genome 5′ sequences are assembled into virions. The poly(A)-containing genome RNA isolated from purified B77 virions was heat denatured, subjected to sedimentaion on glycerol gradients, and analyzed for 5′ sequences by hybridization (Fig. 2B). In contrast to the polysomal RNA, the only significant peak of hybridizable material sedimented at 38S. We concluded that intracellular 38S, 28S, and 21S mRNA's contained some of the same sequences as those present at the 5′ terminus of the genome RNA. Only 38S RNA, however, was packaged to a significant extent.

It was possible that the discrimination necessary for 38S RNA to be packaged into virions is dependent on the ability of 38S RNA molecules to form a dimer structure. Since it appears from electron microscopic observations of the structure of 60S to 70S retrovirus RNAs that two 38S genome RNA subunits contain a stable noncovalent linkage near their 5′ ends (3), we investigated whether the entire 101 bases of the 5′-terminal sequence were present in the smaller mRNA's and whether alterations in this sequence might determine the selection of 38S RNA for packaging.

Extent of sequence homology of subgenomic virus-specific RNAs with 5′ genome sequences in B77 sarcoma virus-infected cells. [^{3}H]uridine-labeled poly(A)-containing RNA was isolated from B77 ASV-infected cells and separated into three different size classes by sedimentation on glycerol gradients (Fig. 3). The RNA in each of these pools was then subjected to a second cycle of oligo(dT)-cellulose chromatography to remove 5′-terminal fragments produced by cleavage from larger RNA species. The kinetics of hybridization of RNA from each size class with ^{32}P-labeled 5′-cDNA was followed and compared with the kinetics of hybridization of 60S to 70S RNA isolated from purified virions with 5′-cDNA (Fig. 4). Note that the final extent of hybridization of RNA in pools I, II, and III to 5′-cDNA was 90 to 95% that of the 60S to 70S genome RNA. This indicates that, within the limits of this type of analysis, sequences complementary to most, if not all, the 5′-cDNA sequence were present in each of the three RNA size classes.

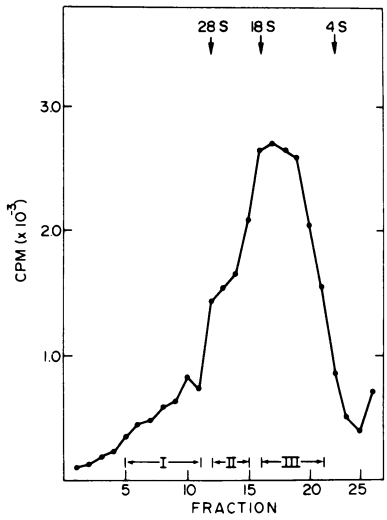

FIG. 3. *Size distribution of poly(A)-containing RNA from B77 sarcoma virus-infected cells. Poly(A)-containing RNA labeled with [^{3}H]uridine was isolated from whole infected cells by procedures described in the text. The RNA was sedimented on 5 to 30% glycerol gradients and fractionated as described in legend to Fig. 2. Samples from each fraction were counted for radioactivity, and the fractions were combined into three pools as shown on the figure.*

It was possible that in infected cells there are two or more populations of RNAs, some of which contain the entire 5′ genome sequence and some of which contain only part of this sequence. Since the RNA was present in large excess in the experiment shown in Fig. 4, the RNA-DNA hybridization reaction should follow pseudo-first order kinetics. If C_0 is the initial concentration of single-stranded cDNA, C is the concentration of single-stranded cDNA at time t, and C_r is the

concentration of RNA, a plot of ln C_0/C versus C_rt should yield a straight line if all of the 5' sequences are present in equal abundance. Such plots for fractions I, II, and III are shown in Fig. 5. Linear plots with a single slope were obtained in each case. This evidence was, thus, consistent with the hypothesis that all mRNA's contain the entire 101-base genome 5' sequence.

These kinetic data, as well as the data below, also imply that, under the hybridization conditions we used, there was no significant annealing of 5'-cDNA to the 21-base terminal repetition at the 3' terminus of the RNA molecules. If this were the case, those regions of the cDNA contained in the reiterated sequence would anneal at a faster rate than would those in the unique sequence, and the ln C_0/C versus C_rt plot would be expected to deviate from linearity. Furthermore, this would result in the production of a 21-base DNA fragment after S1 digestion and alkali denaturation. As is shown below, this was not the case.

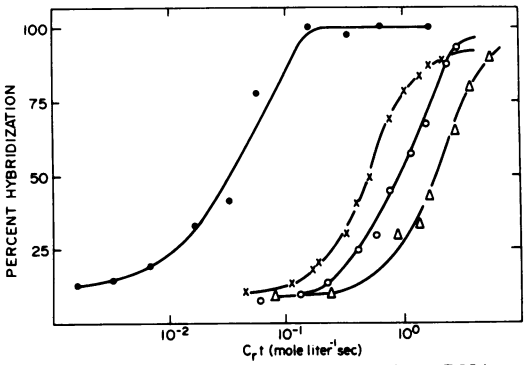

FIG. 4. *Kinetics of hybridization of 5'-cDNA to various size classes of poly(A)-containing RNA from B77-infected cells. The RNA from the three pools shown in Fig. 3 was subjected to a second cycle of oligo(dT)-cellulose chromatography, as described in the text, and the poly(A)-containing RNA was dissolved in 200 μl of sterile distilled water. Various dilutions of 70S genome RNA (●) or RNA from pools I (×), II (○), and III (△) were hybridized with ^{32}P-labeled 5'-cDNA as described in the text. C_r is the molar concentration of nucleotides in RNA, and t is the time of hybridization in seconds. The maximum hybridization of the 5'-cDNA to 70S RNA was 85% in this experiment, and all other values have been normalized to this value.*

As a further test of the hypothesis that the majority of the intracellular viral mRNA contains the 5' genome sequence, we compared the fraction of total molecules in each RNA size class containing genome 5'-terminal sequences with the fraction of the molecules containing sequences complementary to the 3'-terminal 200 to 300 bases of the genome RNA. Since it has been shown that 3'-terminal genome sequences are present in each size class of intracellular RNA (16, 32), the determination of the fraction of molecules containing sequences complementary to the 3'-cDNA should be equivalent to an assay for the concentration of the bodies of the subgenomic mRNA's. Any large excess in the fraction of molecules containing 3' sequences would indicate the possibility that there exists a population of molecules that do not contain the 5'-terminal genome sequences. Therefore, ^{32}P-labeled DNA complementary to the 3' end of the viral genome 38S RNA was prepared by an oligo(dT)-primed reverse transcriptase reaction. This 3'-cDNA was then used to quantitate, by an experiment similar to that shown in Fig. 4,

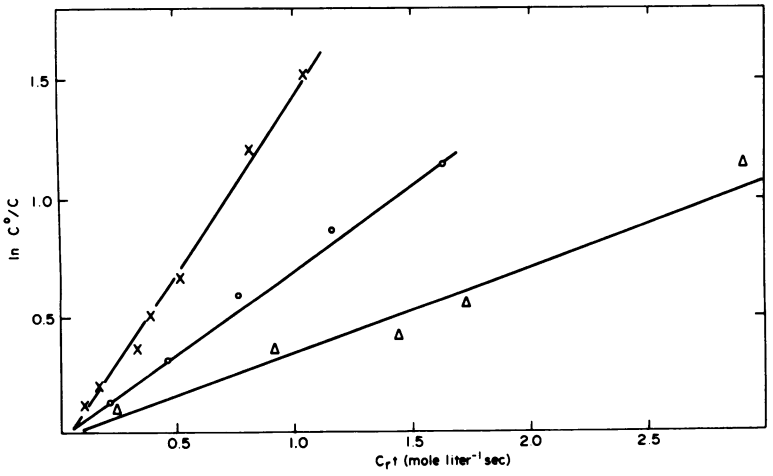

FIG. 5. *Plot of ln C_0/C versus C_rt. Replot of the data given in Fig. 4. C_0, initial concentration of cDNA (C_0); C, concentration of RNA-DNA hybrids at time t; C_r, concentration of RNA.*

TABLE 1. *Determination of the fraction of virus-specific RNA in B77-infected cell poly(A)$^+$ RNA*[a]

RNA pool	Size class	$C_rt_{1/2}$ (mol·s/liter)		Fraction of virus-specific RNA[b]		Ratio of 5′ to 3′[c]
		5′-cDNA	3′-cDNA	5′-cDNA	3′-cDNA	
Expt 1						
Genome		0.034	0.0075	1.0	1.0	
I	31S–43S	0.53	0.20	0.064	0.038	1.7
II	23S–28S	0.94	0.35	0.018	0.011	1.7
III	6S–20S	1.9	0.34	0.005	0.007	0.8
Expt 2						
Genome		0.030	0.012	1.0	1.0	
I	35S–42S	0.50	0.35	0.060	0.034	1.8
II	25S–32S	0.85	0.60	0.018	0.010	1.7
III	14S–22S	2.15	0.80	0.005	0.005	1.0
IV	7S–12S	8.5	0.35	0.00024	0.002	0.1

[a] Two independent experiments were carried out, using different preparations of ^{32}P-labeled 5′- and 3′-cDNA and of [^{3}H]uridine-labeled cellular RNA.

[b] Fraction of virus-specific RNA = [($C_rt_{1/2}$ for genome RNA)/($C_rt_{1/2}$ for RNA pool)] × [(complexity of RNA in pool)/(complexity of genome RNA)]. The molecular complexities of the genome RNA and the RNA in pools I, II, III, and IV were assumed to be 10, 10, 5, 3, and 0.6 kilobases, respectively.

[c] Ratio = (fraction of virus-specific RNA for 5′-cDNA)/(fraction of virus-specific RNA for 3′-cDNA).

the fraction of virus-specific RNA in each size class (Table 1). Note that for each size class of RNA, except for pool IV in experiment 2, the fraction of RNA containing 5′ ends appears to be the same or somewhat greater than the fraction of RNA containing 3′ ends. This suggests that molecules containing 5′-terminal genome RNA sequences probably represent the bulk of the intracellular virus-specific RNA molecules of the 21S to 38S size class. The reason for the 10-fold excess of 3′ ends in RNA of the 7S to 12S size class (Table 1) is not yet understood. It is possible that this reflects the presence of 3′-terminal fragments derived from larger virus-specific RNA molecules which contaminate pool IV.

It should be stated here that the fractions of virus-specific RNA determined in Table 1 were probably underestimated. The concentration of poly(A)$^+$ RNA was calculated from the specific radioactivities determined for the total RNA after a 12-h labeling with [^{3}H]uridine. Since the specific activity of the poly(A)$^+$ fraction was certainly greater than that of the rRNA, although the latter constituted the bulk of the total RNA, it follows that the calculated C_rt values were also probably overestimated. For the purpose of these experiments, however, it was only the relative C_rt values which were of significance, and these values are not dependent on absolute concentrations.

Integrity of 5′-cDNA-RNA hybrids. The previous experiments indicated that a majority of virus-specific RNA molecules contain most of the 5′-terminal 101-nucleotide sequence common to the genome RNA. The data, however, did not exclude the possibility that there might be small deletions or rearrangements of this sequence which occur during the process of formation of the subgenomic mRNA's. Such deletions, for instance, could create initiation sites for protein synthesis.

To test for this possibility, ^{32}P-labeled 5′-cDNA was annealed to the different size classes of RNA isolated from infected cells. The resulting hybrids were then treated with S1 nuclease to remove any unpaired DNA seqeunces, and then the digested hybrids were denatured with alkali. The size of the denatured DNA was determined by gel electrophoresis, using appropriate single-stranded marker DNAs. Under the digestion conditions used, a nick is produced in the DNA at sites where the RNA sequences are spliced out (5). Any splicing of this type should result in production of fragments smaller than 101 bases.

The results of this experiment are shown in Fig. 6. As expected, hybridization of 5′-cDNA to 60S to 70S genome RNA followed by S1 digestion resulted in no detectable change in the size distribution of the 5′-cDNA when compared with the untreated control (Fig. 6A and B). There was no evidence of any material migrating faster than the major peak of 101 bases. In the absence of added cellular or genome RNA, the cDNA was completely digested (Fig. 6C). When the 5′-cDNA was hybridized to each of the three different size classes of cellular RNA, a peak of radioactivity corresponding to the size of the untreated 5′-cDNA was obtained (Fig. 6C through E). However, in contrast to the results obtained with the genome RNA, some labeled

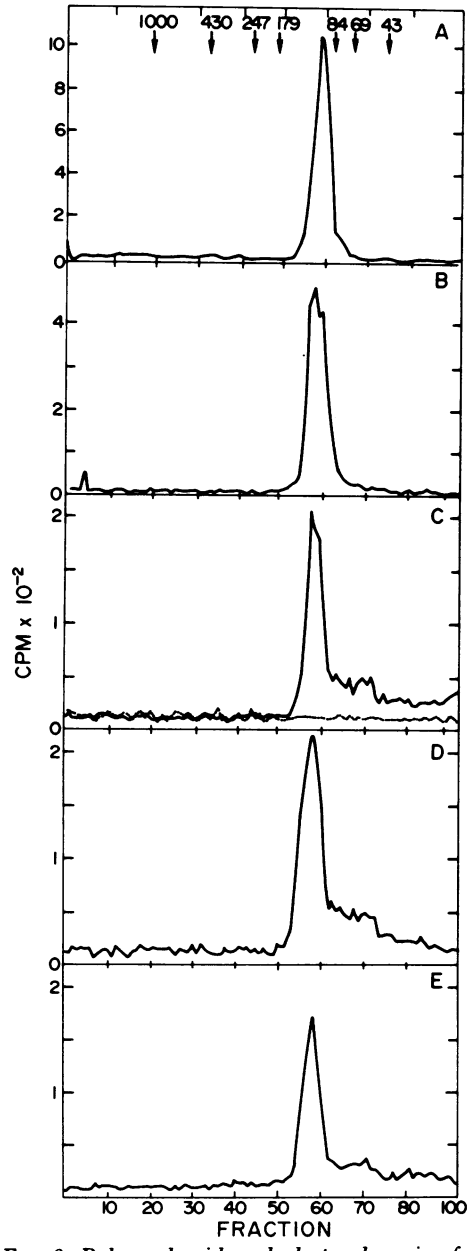

FIG. 6. *Polyacrylamide gel electrophoresis of S1 nuclease-treated and alkali-denatured cDNA-RNA hybrids. ^{32}P-labeled 5'-cDNA was hybridized with 70S virion genome RNA and with RNA in pools I, II, and III (Fig. 3) according to procedures given in the text, except that the time of hybridization was 44 h. After hybridization, the samples were digested with S1 nuclease at 43°C for 30 min. The hybrids were then precipitated with 2 volumes of 95% ethanol at −20°C. The precipitates were dissolved in 10% glycerol-0.3 M NaOH-2 mM EDTA. Electrophoresis of the samples was carried out on 10% polyacrylamide gels for 3.5 h at 100 V. The gels were fractionated into 1-mm slices, eluted with 0.3 M sodium acetate containing 50% formamide, and counted for radioactivity. Fragments of DNA derived by HaeIII restriction endonuclease cleavage of ColE1 DNA were treated with alkali similarly and electrophoresed on a parallel gel. Their positions are shown on the figure. (A) Untreated 5'-cDNA. (B) 5'-cDNA hybridized with 70S genome RNA. 5'-cDNA hybridized with RNA in pools I (C), II (D), and III (E). The 5'-cDNA hybridized with no added RNA is shown in the dashed line in (C).*

material migrated faster than the major peak at 101 bases. Although there are no deletions or rearrangements of the 5'-terminal 101 bases in the majority of the intracellular mRNA's, the possibility remains open that a small fraction of the mRNA does contain such deletions or rearrangements.

DISCUSSION

In this paper we first established that subgenomic RNAs containing the sequences from the 5' terminus of the genome RNA are associated with polysomes in infected cells and presumably are translated (Fig. 2A). This indicates that such RNAs are functional mRNA's. Our data further suggest that the bulk of these subgenomic mRNA molecules are restricted to infected cells and are not found in mature virions (Fig. 2B). We cannot, however, rule out the presence of small amounts of these smaller mRNA's in virions.

In view of this selection for 38S RNA molecules in virions, we investigated the extent of similarity of the 5'-terminal sequence more thoroughly. The results of hybridization kinetics experiments established, within the limits of this type of analysis, that most, if not all, of the 5'-terminal 101 bases of the genome RNA were present in the smaller mRNA's (Fig. 4). The data shown in Table 1 are also consistent with the hypothesis that most intracellular viral RNAs in the different size classes contain the 5'-terminal 101-base sequence. The possibility remains open, however, that within the 7S to 12S size class there exists a population of virus-specific RNAs that lack this sequence (Table 1).

The calculated fraction of virus-specific RNA in pools I and II when the 5'-cDNA was used as probe was somewhat greater than the fraction when the 3'-cDNA was used as a probe (Table 1). The reason for this difference is not yet understood but appeared to be reproducible in a number of experiments. The presence in the RNA preparation of 3'-terminal poly(A)-containing fragments derived from larger mRNA's would be expected to give the opposite result; i.e., the fragments would hybridize to the 3'-cDNA but not to the 5'-cDNA. Further study will be required to determine whether the ob-

served excess of mRNA's containing the 5'-terminal sequence is due to a systematic error of the technique or whether it is due to the presence of some species of virus-specific RNAs that either lack the 3'-terminal genome sequences or contain redundant 5' sequences.

The data in Fig. 6 indicated that there appear to be no deletions or other sequence rearrangements of the 5'-terminal 101 bases of most of the subgenomic mRNA's. Therefore, the leader sequence appears to extend for at least 101 bases with no interruptions. However, we do not know the minimum size of an RNA deletion that could be detected by the S1 mapping technique. It has been shown that S1 nuclease cleaves deletion loops of less than 15 bases in heteroduplex DNA (25). Furthermore, S1 nuclease specifically cleaves duplex DNA containing single mismatched base pairs. The cleavage at these sites, however, was less effective than that with deletions or insertion DNA-DNA heteroduplexes (26).

If the subgenomic mRNA's are generated by splicing, our data suggest that the first splice point lies to the 3' side of the initiation site for viral DNA synthesis (30). Rothenberg et al. have reported that no intervening sequences were visible by electron microscopy in the 5'-terminal 500-base leader sequence of murine leukemia virus RNA when subgenomic intracellular viral RNAs were hybridized to full-length copies of viral cDNA (23). Electron microscopy, however, cannot be used to detect short, unpaired regions of less than 100 bases. Analyses of the type carried out in this study must be done in conjunction with electron microscopy in order to demonstrate conclusively the absence of additional splicing in the leader sequence.

We have not proved in this paper whether or not the 5'-terminal genome sequence is actually present at the 5' termini of the smaller mRNA's. The data of Mellon and Duesberg, however, strongly suggest that a 21-base oligonucleotide which contains the cap structure and is found in the genome RNA is also present in the subgenomic mRNA's (19). These results have been confirmed and extended by Cordell et al., who showed that this cap-containing oligonucleotide was present in RNA sequences protected by hybridization to 5'-cDNA (12). Taken together, then the evidence is consistent with the hypothesis that most of the subgenomic mRNA's of ASV contain at their 5' termini a sequence of at least 101 bases identical that at the 5' end of the genome RNA. Whether these subgenomic RNAs are formed by RNA splicing or by some other mechanism, such as integration of subgenomic DNA molecules, has not yet been established.

The function of the shared 5'-terminal leader sequence is not yet clear. It has been shown recently that ribosomes from a reticulocyte lysate bind strongly to the ASV genome RNA between residues 9 and 53 from the 5' end (13). Translation of the *gag* precursor polypeptide, however, does not appear to begin at this site but occurs at a site distal to the 5'-terminal 101-base sequence. It is possible, therefore, that the sequence between residues 9 and 53 is required in the subgenomic mRNA's in order for ribosome binding to occur.

ACKNOWLEDGMENTS

We thank R. Bird for ColE1 DNA and *Hae*III restriction endonuclease, T. Walseth and R. Johnson for assistance in the preparation of [α-^{32}P]dATP, and J. Beard and the Office of Program Resources and Logistics, Viral Cancer Program, for supplying us with avian myeloblastosis virus reverse transcriptase. We thank B. Cordell, S. R. Weiss, H. E. Varmus, and J. M. Bishop for sending us a manuscript of a paper before its publication.

This work was supported by Public Health Service grant CA-16262 from the National Cancer Institute. C.M.S. is the recipient of Public Health Service Career Development Award 1 KO4 CA-00205 from the National Cancer Institute.

LITERATURE CITED

1. **Aloni, Y., R. Dhar, O. Laub, M. Horowitz, and G. Khoury.** 1977. Novel mechanism for RNA maturation: the leader sequences of simian virus 40 mRNA are not transcribed adjacent to the coding sequences. Proc. Natl. Acad. Sci. U.S.A. **74**:3686–3690.
2. **Beeman, K., P. Duesberg, and P. Vogt.** 1974. Evidence for crossing over between avian tumor viruses based on analysis of viral RNAs. Proc. Natl. Acad. Sci. U.S.A. **71**:4253–4258.
3. **Bender, W., and N. Davidson.** 1976. Mapping of poly(A) sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules. Cell **7**:595–607.
4. **Berget, S. M., C. Moore, and P. A. Sharp.** 1977. Spliced segments at the 5'-terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. U.S.A. **74**:3171–3175.
5. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus . Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
6. **Billeter, M. A., J. T. Parson, and J. M. Coffin.** 1974. The nucleotide sequence complexity of avian tumor virus RNA. Proc. Natl. Acad. Sci. U.S.A. **71**:3560–3564.
7. **Breathnach, R., J. L. Mandel. and P. Chambon.** 1977. Ovalbumin gene is split in chicken DNA. Nature (London) **270**:314–319.
8. **Bronson, D. L., A. Y. Elliott, and D. Ritzi.** 1975. Concentration of Rous sarcoma virus from tissue culture fluids in polyethylene glycol. Appl. Microbiol. **30**:464–471.
9. **Bronson, D. L., A. Y. Elliott, and D. Ritzi.** 1976. A comparison of four methods used to concentrate Rous sarcoma virus from tissue culture fluids. J. Gen. Virol. **33**:403–410.
10. **Brugge, J. S., A. F. Purchio, and R. L. Erickson.** 1977. Virus-specific RNA species present in the cytoplasm of Rous sarcoma virus-infected chicken cells. Virology **83**: 16–26.
11. **Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts.** 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell **12**:1–8.

12. **Cordell, B., S. R. Weiss, H. E. Varmus, and J. M. Bishop.** 1978. At least 104 nucleotides are transposed from the 5'-terminus of the avian sarcoma virus genome to the 5'-termini of the smaller mRNAs. Cell **15**:79–91.
13. **Darlix, J.-L., P.-F. Spahr, P. A. Bromley, and J.-C. Jaton.** 1979. In vitro, the major ribosome binding site on Rous sarcoma virus RNA does not contain the nucleotide sequence coding for the N-terminal amino acids of the *gag* gene product. J. Virol. **29**:597–611.
14. **Friedrich, R., H.-J. Kung. B. Baker, H. E. Varmus, H. M. Goodman, and J. M. Bishop.** 1977. Characterization of DNA complementary to nucleotide sequences at the 5'-terminus of the avian sarcoma virus genome. Virology **79**:198–215.
15. **Haseltine, W. A., A. M. Maxam, and W. Gilbert.** 1977. Rous sarcoma genome is terminally redundant: the 5' sequence. Proc. Natl. Acad. Sci. U.S.A. **74**:989–993.
16. **Hayward, W.** 1977. Size and genetic content of viral RNAs in avian oncornavirus-infected cells. J. Virol. **24**: 47–63.
17. **Jeffreys, A. J., and R. A. Flavell.** 1977. The rabbit β-globin gene contains a large insert in the coding sequence. Cell **12**:1097–1108.
18. **Joho, R. H., M. A. Billeter, and C. Weissman.** 1975. Mapping of biological functions on RNA of avian tumor viruses: location of regions required for transformation and determination of host range. Proc. Natl. Acad. Sci. U.S.A. **72**:4772–4776.
19. **Mellon, P., and P. H. Duesberg.** 1977. Subgenomic cellular Rous sarcoma virus RNAs contain oligonucleotides from the 3' half and the 5'-terminus of the virion RNA. Nature (London) **270**:631–634.
20. **Opperman, H., J. M. Bishop, H. E. Varmus, and L. Levintow.** 1977. A joint product of the genes "gag" and "pol" of avian sarcoma virus: a possible precursor of reverse transcriptase. Cell **12**:993–1005.
21. **Pawson, T., R. Harvey, and A. E. Smith.** 1977. The size of Rous sarcoma virus mRNA's active in cell-free translation. Nature (London) **268**:416–421.
22. **Pawson, T., G. S. Martin, and A. E. Smith.** 1976. Cell-free translation of virion RNA from nondefective and transformation-defective Rous sarcoma viruses. J. Virol. **19**:950–967.
23. **Rothenberg, E., D. J. Donoghue, and D. Baltimore.** 1978. Analysis of a 5' leader sequence on murine leukemia virus 21S RNA: heteroduplex mapping with long reverse transcriptase products. Cell **13**:435–451.
24. **Schwartz, D. E., P. C. Zamecnik, and H. L. Weith.** 1977. Rous sarcoma virus is terminally redundant: the 3' sequence. Proc. Natl. Acad. Sci. U.S.A. **74**:994–998.
25. **Shenk, T. E., J. Carbon, and P. Berg.** 1976. Construction and analysis of viable deletion mutants of simian virus 40. J. Virol. **18**:664–671.
26. **Shenk, T. E., C. Rhodes, P. W. J. Rigby, and P. Berg.** 1975. Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **72**:989–993.
27. **Stoll, E., M. A. Billeter, A. Palmerberg, and C. Weissman.** 1977. Avian myeloblastosis virus RNA is terminally redundant: implications for the mechanism of retrovirus replication. Cell **12**:57–72.
28. **Stoltzfus, C. M., and P. N. Snyder.** 1975. Structure of B77 sarcoma virus RNA: stabilization of RNA after packaging. J. Virol. **16**:1161–1170.
29. **Tal, J., H.-J. Kung, H. E. Varmus, and P. M. Bishop.** 1977. Characterization of DNA complementary to nucleotide sequences adjacent to poly(A) at the 3'-terminus of the avian sarcoma virus genome. Virology **79**: 183–197.
30. **Taylor, J. M., and R. Illmensee.** 1975. Site on the RNA of an avian sarcoma virus at which the primer is bound. J. Virol. **16**:553–558.
30a. **Walseth, T. M., and R. A. Johnson.** 1979. The enzymatic preparation of [α-^{32}P]nucleoside triphosphates, cyclic [^{32}P]AMP, and cyclic [^{32}P]GMP. Biochim. Biophys. Acta **562**:11–31.
31. **Wang, L.-H., P. Duesberg, K. Beemon, and P. K. Vogt.** 1975. Mapping RNase T_1-resistant oligonucleotides of avian tumor virus RNAs: sarcoma-specific oligonucleotides are near the poly(A) end and oligonucleotides common to sarcoma and transformation-defective viruses are at the poly(A) end. J. Virol. **16**:1051–1070.
32. **Weiss, S. R., H. E. Varmus, and J. M. Bishop.** 1977. The size and genetic composition of virus-specific RNAs in the cytoplasm of cells producing avian sarcoma-leukosis viruses. Cell **12**:983–992.