Proc. Nat. Acad. Sci. USA
Vol. 71, No. 11, pp. 4371–4375, November 1974

Characterization of the End Groups of RNA of Rous Sarcoma Virus

(3′-terminal nucleoside analysis/5′-phosphate analysis)

J. KEITH, M. GLEASON, AND H. FRAENKEL-CONRAT

Department of Molecular Biology and Virus Laboratory, University of California, Berkeley, Calif. 94720

Contributed by H. Fraenkel-Conrat, August 19, 1974

ABSTRACT **The 3′ terminus of the 30–35S RNA of Rous sarcoma virus is adenosine. Its amount indicates an average molecular weight for that RNA of about 3×10^6. The 5′ terminus of 30–35S RNA of Rous sarcoma virus was not di- or triphosphorylated, whether isolated by the standard procedure or from virus collected within 3 min of its release from the cells.**

The chemical structure of the RNA of the oncornaviruses is still in doubt. Although much evidence favors the existence of subunit RNA molecules of 30–35 S held together by base-paired segments to form the 65–70S genome of these viruses, this concept, as well as the size of the presumed component RNA molecules and the question of whether they are identical or different, are not yet firmly established.

It appeared to us that a study of the end groups of the subunits would contribute significant answers to these questions. We previously reported preliminary results (at a meeting at the University of California, Los Angeles, in 1973) that adenosine appeared to be the predominant 3′-terminal residue of the RNA of Rous sarcoma virus (RSV), in contrast to earlier reports that various oncornaviral RNAs terminated in uridine (1–3). Adenosine has been found to be the predominant terminal nucleoside of the RNA of avian myeloblastosis virus (4–6), and the finding that the oncornaviruses contained poly(A) (7), 3′-terminal in avian myeloblastosis virus (5), made this the preferred and presumed end group also for RSV, notwithstanding the reports to the contrary. We have now refined our preparative and analytical procedures sufficiently to yield consistent and reproducible results for the ratio of 3′-terminal nucleosides to internal nucleotides. The results of these analyses are that there is one 3′-unphosphorylated adenosine per average molecular weight of about 3×10^6, with the values for guanosine, uridine, and cytidine varying from 40 to 10% of the adenosine under different experimental conditions. It appears, therefore, that the 30–35S component of RSV RNA is predominantly uniform in regard to its 3′ end. Separate studies of the nature of the end group of the poly(A) isolated from RSV have shown that this material also terminates only in unphosphorylated adenosine. It thus appears that poly(A) is located at the 3′ end of many, if not all, RSV RNA subunits.

Concerning the 5′ ends, we have confirmed the surprising finding of Silber *et al.* (8) that these are unphosphorylated, and have shown this to be true even for virus harvested at 3-min intervals and, thus, probably not a result of secondary extracellular dephosphorylation.

Abbreviations: RSV, Rous sarcoma virus; TMV, tobacco mosaic virus.

MATERIALS AND METHODS

Secondary cultures of chick embryo fibroblasts were prepared according to published procedures (9, 10). Original PR-RSV-B virus stock was a gift of P. Duesberg. For 3′ end group analyses, virus was labeled with a mixture of the four ^{3}H-labeled nucleosides (Schwarz Mann) in 6.0 ml (per 10-cm plastic petri dish) of Dulbecco's modified Eagle's medium (11) with bicarbonate buffer, supplemented with 1% (v/v) calf and 1% (v/v) chick serum. Typically the label was administered to 40 dishes in a ratio of 2 mCi of adenosine (labeled in the 2 position) to 1 mCi of each of guanosine, cytidine, and uridine; 1 mCi of total label was used per 10-cm petri dish. Medium containing radioactive virus was removed after 4 hr and replaced with 6 ml of the same medium but without addition of more labeled nucleosides. Collections were continued at 4-hr intervals for a period of one or two days.

Virus was purified and viral RNAs were extracted as described (10). The 30–35S RNA used was obtained by heating viral RNAs at 100° for 45 sec, followed by another sucrose gradient fractionation. The RNA, together with carrier RNA of tobacco mosaic virus (TMV) used as a 30S marker on sucrose gradients, was hydrolyzed in 50 μl of 0.3 M KOH for 18 hr at 37° in a sealed capillary tube, and then spotted directly with absorbance markers (nucleosides and nucleotides) 44.5 cm from the anode side, and 8 cm from one edge of a sheet of Whatman 3 MM paper of 95 × 48 cm. Electrophoresis was in a refrigerated two-phase electrophoresis system, with pyridine–acetate buffer at pH 3.5 (5% v/v glacial acetic acid, 0.5% v/v pyridine, and 1 mM EDTA) at about 3000 V for 4 hr. After the chromatograms were dried overnight in an air stream, the nucleosides and nucleotides were made visible under a UV lamp and the spots were marked. The cathode portion of the paper was then paper chromatographed in solvent A (BuOH–EtOH–water 80:10:25). An example of the resulting pattern is shown (Fig. 1).

The spots were eluted from the paper and in some instances again chromatographed with solvent B (76% ethanol) on Whatman 1 paper, previously soaked in a 10% saturated solution of ammonium sulfate and dried. All spots were eluted with water; small aliquots of the nucleotide eluates and large aliquots or all of the nucleoside eluates were counted in 10 ml of Bray's scintillation solution.

For 5′ end group analyses, RSV was labeled with ^{32}P in 6 ml (per 10-cm plastic petri dish) of phosphate-free medium 199, supplemented with 1% calf and 1% chick serum by addition of 0.5 mCi of carrier-free $H_3{}^{32}PO_4$ (New England Nuclear Corp.) per dish. Radioactive virus was collected at 4-hr intervals up to 2 days. After each collection, 6 ml of medium 199 supple-

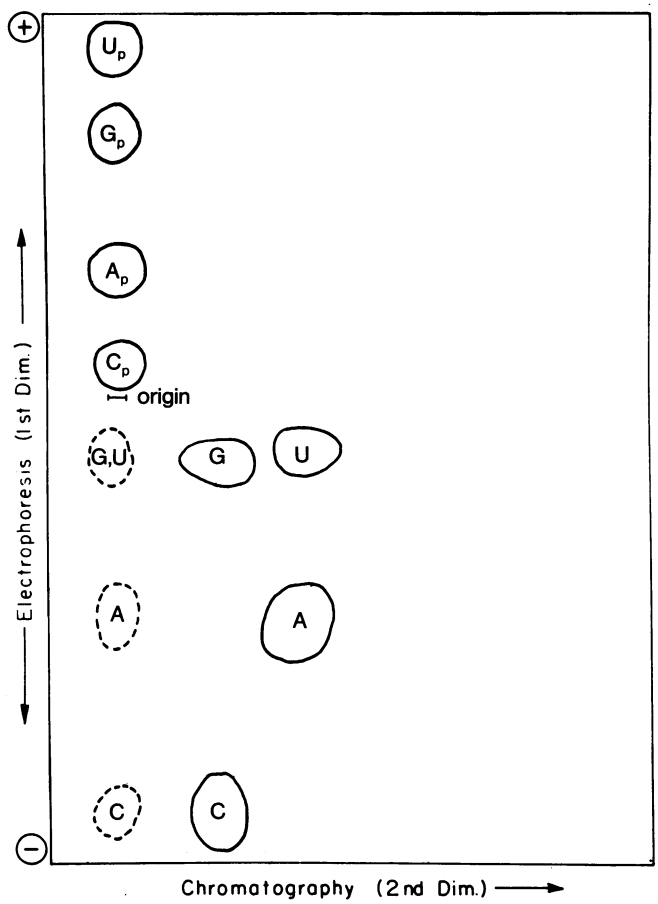

FIG. 1. Electrophoresis-chromatography pattern of nucleotides and nucleosides. Alkaline digests of RNA and markers were subjected to electrophoresis at pH 3.5, and the lower cathodic half of the paper was then chromatographed in solvent A. Spurious radioactivity appears in and between the guanosine and uridine areas, and in the right bottom corner of the graph. Longer chromatography was often used to increase the separation and diminish the contamination of guanosine and uridine.

mented with 2% tryptose phosphate broth and 1% calf and 1% chick serum was added to each dish.

To obtain freshly released virus, labeling with ^{32}P was done for 14 hr in 12.5 ml (per 15-cm plastic petri dish) of phosphate-free medium 199 supplemented with 1% calf and 1% chick serum and 2% tryptose broth, by addition of 10 mCi of carrier-free $H_3{}^{32}PO_4$ (New England Nuclear Corp.) per dish. Radioactive virus was collected at 3-min intervals for 4 hr with a vacuum collection device and modified petri dish covers. After each collection, 5 ml of medium 199 supplemented with 0.25% calf and 0.25% chick serum was added to each dish with a pipet with a series of three-way valves. The plates remained in the incubator during the entire process and were rocked slowly with an automatic rocking device of our own design. The virus and its RNA were isolated by the usual procedure; the yield of heated and recentrifuged 30–35S RNA was about 600,000 cpm.

RNA was hydrolyzed in 0.2 ml of KOH for 18 hr at 37°. The solution was neutralized with Dowex 50WX-8 resin in hydrogen form. Sample volumes were reduced by lyophilization or in an airstream at 37°. The two-dimensional thin-layer chromatography procedures were used as described (12); the 2 M Na formate buffer (pH 3.4) was made by adjusting 2 M Na formate to pH 3.4 with 2 M formic acid. Polyethylenimine-impregnated cellulose, anion-exchange, precoated plastic sheets were used (Brinkmann Instruments). They are washed with water just before use.

Oligo(dT)–cellulose T-2 was purchased from Collaborative Research Inc. Affinity chromatography was as described (13), except that both high- and low-salt eluting buffers also contained 0.5% Sarkosyl and 0.1% diethylpyrocarbonate.

RESULTS

Analysis for Terminal Nucleosides of RSV RNA. Our procedure is a variant of that used by Sugiyama with TMV RNA for the separation and purification of a small amount of labeled nucleoside from the great amount of radioactive nucleotides (14). The present procedure avoids repeated elution, but rather achieves primary separation of the four nucleotides and nucleosides by two-dimensional electrophoresis chromatography. It requires no correction for mechanical losses, since the calculation of the chain length relies entirely on the ratio of ^3H found in the nucleoside as compared to the corresponding nucleotide area of the same sheet of paper, and the known base composition of RSV RNA. The method was validated with TMV RNA.

In principle, the one-step electrophoretic separation of all nucleotides and nucleosides should suffice for the purpose at hand, but it became evident that spurious radioactivity streaking between and beyond the nucleosides vitiated the valid determination of the low radioactivity actually associated with the nucleosides. When these areas were therefore purified by chromatography in the second dimension (in solvent A), the radioactivity associated with guanosine, uridine, and cytidine was often greatly diminished. The former two, however, continued to be contaminated to a variable extent, which was seen when all four nucleosides were eluted and rechromatographed in solvent B. Under these conditions the counts in the adenosine and cytidine areas diminished only to the extent of the expected operational losses (20%), but those of uridine and particularly of guanosine often decreased significantly. Most of the data listed in Table 1 for guanosine and uridine were obtained without this second chromatographic purification step and the values can, thus, be regarded as possibly too high. The problem with spurious radioactivity was greater with RSV RNA than when TMV [^3H]RNA was used as a methodological standard, since with the latter adequate end-group results were obtained without chromatography in solvent B (see Table 1).

The conclusion from these data is that only adenosine occurs in stoichiometrically significant and quantitatively reproducible amounts, corresponding to the end group of an RNA of about 10,000 nucleotides, and a molecular weight of 3.2×10^6 (see Table 1).

Terminal Nucleoside Analysis of Poly(A)-rich and Poly(A)-poor Fraction of RSV RNA. Upon several instances, RSV RNA was fractionated on oligo(dT)-cellulose columns into a fraction bound by the column at high salt concentration and a fraction that was not bound. The ratio of these varied in different experiments from 2:1 to 1:1. This observation accords with the finding by Ihle *et al.* (15) that mammalian oncornaviral RNA contains fractions with short or no poly(A) sequences.

TABLE 1. *3′ End groups of RNA of Rous sarcoma virus**

	Nucleotides/nucleoside				Average	MW × 10^{-6}
			RSV RNA			
A	2,800	2,280	2,260	2,570	2,480	3.2
G	6,300	5,600	5,000	8,100	6,330	(10.5)
U	11,500	5,300	11,200	17,300	11,300	(22)
C	18,000	19,200	21,400	25,600	21,000	27
	RSV RNA not bound to oligo(dT)-cellulose†					
A	1,540					2.3
G	3,100					(3.8)
U	5,400					(6.6)
C	4,800					5.9
	RSV RNA digest freed from terminal poly(A) on oligo(dT)-cellulose†					
A	3,600					5.2
G	5,200					(6.5)
U	13,000					(16)
C	13,000					15
	Poly(A) from RSV RNA (single experiment)					
A	130 (6.8% of label, only A labeled)					2.5
	TMV RNA					
A	1,610					1.9
G	10,000					(13)
U	4,900					(6)
C	12,000					19

* The data are for individual experiments on different RSV RNA isolates. They are generally based on the radioactivity found upon two-dimensional separation (see Fig. 1). Those for adenosine and cytidine are quantitatively valid, as shown by rechromatography in solvent B; those for guanosine and uridine are subject to variable contamination, and the lowest nucleoside values (highest ratios) are thus the most valid. Molecular weight (MW) estimates, based on the highest ratios in the case of RSV RNA, are therefore given in parentheses.

† Molecular weight calculations, assuming that these samples lacking the poly(A) (10% of total A) have the composition: A/G/U/C = 22/26/26/26.

Terminal nucleoside analysis of the RSV RNA that was not bound indicated, in one experiment, that adenosine was also the predominant end group in this fraction, and the chain length, based on this analysis, was possibly slightly smaller than that of the bound RNA. The oligo(dT)-bound poly(A)-rich fraction was, after low-salt elution, degraded with pancreatic plus T1 RNase and again fractionated on the same type of column into the viral poly(A) and the bulk of the degradation products. These two fractions were analyzed for 3′-terminal nucleosides only once each and, thus, these results must also be regarded as preliminary. They indicated that the poly(A) fraction contained one adenosine per 130 adenylylic acid molecules, in amount equivalent to a molecular weight for the 30–35S RNA of 2.7×10^6; and that there were, as expected, less than stoichiometric amounts of the nucleosides in the unbound fraction of the digest that contained most of the labeled nucleotides (2.8×10^6 cpm). (See Table 1.)

Analysis of Phosphorylated 5′-Terminal Nucleotides. Since the RNAs of all animal and bacterial viruses that have been studied carried 5′ di- or triphosphate groups, we applied the thin-layer technique of Hayashi and Hayashi (12), which permits the separation of all nucleoside di-, tri-, and tetraphosphates from the mixture of nucleotides obtained upon alkaline degradation of ^{32}P-labeled RNA.

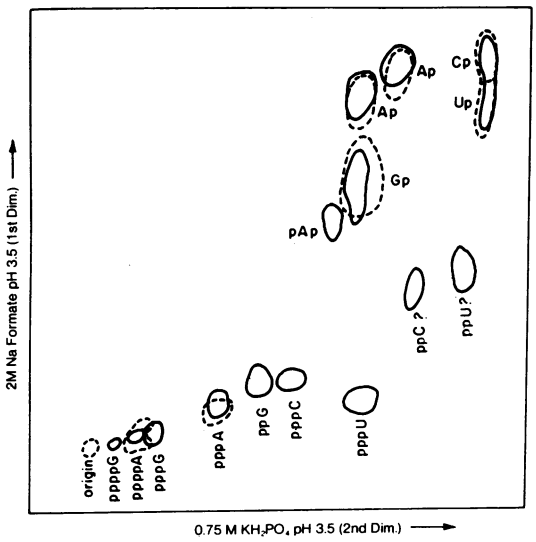

FIG. 2. Anion-exchange thin-layer chromatography of RNA digests and markers (12). The separation of tetra-, tri-, and diphosphorylated nucleotides is illustrated. *Solid lines* outline the added markers as indicated; *dashed outlines* are from the radioautogram, due to degradation products of influenza virus (WSN) [^{32}P]RNA. The previously reported pppAp and ppAp (17) chromatograph similarly to the ppppA and pppA markers. In later experiments resolution in that area was improved by running the plates, with an attached wick (12), longer in the second dimension. This leads to everything on the right of pppU and pAp to run off the plate but gives better separation of, and more space between, the polyphosphorylated purines.

TMV RNA, as other plant viral RNAs, is unphosphorylated at the 5′ terminus (16). This result was confirmed by this new technique, which was further validated by analysis of influenza virus RNA (strains WSN and PR-8) and low-molecular-weight RSV RNA fractions, largely tRNA, and finding of spots on radioautograms in the di-, tri-, and/or tetraphosphate areas in these cases (see Fig. 2). In contrast, digests of 30–35S RSV RNA showed no detectable 5′-phosphorylated degradation products (less than 0.5 per 10^7 daltons of RNA) in the areas of or near the markers (ppppG, ppppA, pppG, pppA, ppG, pAp, ppC, and ppU).

It was recently shown that considerable conformational changes occur in the RNA of extracellular RSV during hours of contact with the cells; RSV RNA, separated from the cells at 3-min intervals, was 30–35 S without heating, and formed the 70S component only secondarily (18). It appeared to us that terminal dephosphorylation might also take place during the customary hours of incubation of released virus in the presence of cells and their medium, and that different results might be obtained if virus for end-group analysis was also collected at 3-min intervals.

The earlier findings, that RSV collected at 3-min intervals contained predominantly 30–35S RNA, were confirmed. However, this type of 30–35S viral RNA, if freed from small associated primer or tRNA molecules by the usual melting procedure, also proved to be free from 5′ di- or triphospho-

rylated terminal nucleotides. The faint streak of radioactivity from the origin in and between these markers (ppppG, ppppA, etc.) amounted to about 7 cpm per marker, when over 700,000 cpm were applied and about 300 cpm would be expected per terminal triphosphate. In contrast, the 10–4S material released upon heating of this 30–35S RNA and separated on a sucrose gradient, as well as unmelted RNA, showed distinct radioactive spots in the guanosine di- and tetraphosphate areas amounting to about 0.05 and 0.1% of the applied radioactivity for the former, and one order of magnitude less for the latter type of RNA. TMV RNA showed no spots (<30 cpm) in the polyphosphate marker areas when 1.7×10^6 cpm was analyzed.

DISCUSSION

The choice of methodology for end-group analysis of nucleic acids available only in small amounts merits discussion. The first data on viral RNA end groups and terminal sequences, those of TMV RNA, were obtained by alkaline degradation of generally ^{14}C-labeled RNA, and analysis for nucleosides formed (14). Later, several methods were advocated that required no labeled RNA, but relied on labeled end-group reagents. Of the methods for the 3′ terminals used in that manner, the oxidation-reduction method (periodate followed by [^{3}H]borohydride) tends to give nonquantitative data for the end groups of large RNAs and erroneous identifications of the trialcohols in the hands of several investigators; the blocking of the oxidized 3′ end with ^{14}C-labeled reagents, such as dimedone, may be more generally reliable, but suffers from the difficulty in obtaining sufficiently highly labeled reagents and, in some instances, from the limited stability of the aldehyde condensation product.

Particularly when the RNA to be analyzed is scarce, there are distinct advantages in its isolation in radioactive form (^{3}H), and, thus, procedures relying on end-group labeling lose their advantages. We have for this reason returned to the original method for 3′-terminal nucleoside analysis, assuming that end to be unphosphorylated. The original procedure of freeing the terminal nucleoside from all spurious radioactivity by multiple electrophoretic and chromatographic runs was simplified by a two-dimensional process in which the nucleosides were first subjected to electrophoresis in opposite direction to the nucleotides. The cathodic part of the paper was then subjected to chromatography for further purification, and particularly for separation of guanosine and uridine and of both of these from spurious coelectrophoresing radioactivity. The technique was validated by means of TMV RNA. However, it was found that with RSV RNA the latter two nucleosides had to be purified by an additional chromatographic step to assure freedom from all spurious radioactivity.

Our data indicate that unfractionated 30–35S RSV RNA contains one terminal adenosine per about 9500 nucleotides or 2.7 to 3.2×10^6 daltons. All data are within the range of chain length and molecular weight (2.2 to 3.4×10^3) proposed for the subunit (19), and also agree with the number of 5′-terminal adenosines (8). Most of our data seem to support the molecular weight derived from the complexity of oligonucleotide patterns (20). It would seem *a priori* difficult to rule out the alternate interpretation of our data, that the actual 3′-end group of an 11×10^6-dalton molecule was guanosine (or uridine); however, the presence in the alkaline digests of about four times as much terminal adenosine is not explainable by that hypothesis, and the melting procedure used before the 30–35S RNA was isolated seems to rule out the possibility that the adenosines were derived from contaminating small molecules.

Our preliminary results with oligo(dT) columns suggest that RSV RNA is not homogeneous in regards to its poly(A) content, as has recently been shown to be the case for mammalian oncornaviruses (15). It appears possible that only part of the molecules are polyadenylated, while others carry only one or a few terminal adenosines. The poly(A) isolated by us is somewhat shorter than previously reported for RSV RNA (7), but much longer than that reported for AMV RNA (5). However, the fact that the latter data were obtained with somewhat degraded (<30S) RNA makes those conclusions somewhat questionable.

Our aim had been to characterize both ends of RSV RNA, expecting to find about as many tri- or tetraphosphorylated 5′ ends (pppNp or ppNp) as there were 3′-terminal nucleosides. This expectation, based on analogy in reference to the available data for 5′ termini of all other animal and bacterial RNA viruses, appeared in doubt when a recent report indicated that RSV terminated, like TMV RNA and other plant viruses, in 5′-unphosphorylated adenosine, in numbers approximately equivalent to the expected chain length (8). We have now confirmed, by a different experimental approach and again using [^{32}P]RNA rather than labeled end-group reagent, the conclusion by Silber *et al.;* we were unable to find any significant amounts of nucleoside di-, tri-, or tetraphosphate of the purines in alkaline digests of RSV by sufficiently sensitive methodology. The absence of pCp and pUp is not firmly established, since there was some streaking in these areas.

We considered the possibility that this lack of terminal 5′-phosphates might be an artifact occurring during the hours before virus collection. We, therefore, tested this hypothesis by analyzing freshly released virus, collected at 3-min intervals, for 5′-terminal phosphates. It appears that the 30–35S molecules of such freshly released virus are also free of 5′ di- or triphosphate groups.

This finding suggests the possibility that the subunits of RSV RNA result from post-transcriptional cleavage of 5′-phosphoester bonds of a larger molecule. If that were the case, one would expect the 3′ ends of the resultant molecules to carry phosphate groups, and thus yield no nucleosides upon alkaline degradation. One is tempted to suggest the following explanation for this paradox. If viral as other messenger RNA precursors, generally resulting from cleavage, are thus 3′-phosphorylated, this end may represent a signal for the onset of polyadenylation. Thus, -Np may become transformed, possibly by one enzyme, to -NpAp(Ap)$_n$A, and molecules unphosphorylated on both ends may result.

On the other hand, RSV RNA resembles the RNA of plant viruses in lacking phosphorylated 5′ ends, and it may well be that both types of RNA are dephosphorylated enzymatically, or partly degraded from the 5′ end, as soon as they are made.

Aided by Grant ORG-1253 from the Damon Runyon Memorial Fund for Cancer Research. We are grateful to Dr. E. E. Penhoet for the generous gift of influenza virus (WSN stain).

1. Maruyama, H. B., Hatanaka, M. & Gilden, R. V. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1999–2001.
2. Erikson, R. L., Erikson, E. & Walker, T. A. (1971) *Virology* **47**, 527–528.
3. Lewandowski, L. J., Content, J. & Leppla, S. H. (1971) *J. Virol.* **8**, 701–707.

4. Stephenson, M. L., Wirthlin, L. S., Scott, J. F. & Zamecnik, P. C. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 1176–1180.
5. Stephenson, M. L., Scott, J. F. & Zamecnik, P. C. (1973) *Biochem. Biophys. Res. Commun.* **55**, 8–16.
6. Ahmad, M. S., Markham, P. D. & Glitz, D. G. (1972) *Biochim. Biophys. Acta* **281**, 554–563.
7. Lai, M. M.-C. & Duesberg, P. H. (1972) *Nature* **235**, 383–386.
8. Silber, R., Malathi, V. G., Schulman, L. H., Hurwitz, J. & Duesberg, P. H. (1973) *Biochem. Biophys. Res. Commun.* **50**, 467–472.
9. Rubin, H. (1960) *Proc. Nat. Acad. Sci. USA* **46**, 1105–1119.
10. Duesberg, P. H. & Vogt, P. K. (1970) *Proc. Nat. Acad. Sci. USA* **67**, 1673–1680.
11. Dulbecco, R. & Freeman, G. (1959) *Virology* **8**, 396–397.
12. Hayashi, M. N. & Hayashi, M. (1972) *J. Virol.* **9**, 207–215.
13. Aviv, H. & Leder, P. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 1408–1412.
14. Sugiyama, T. & Fraenkel-Conrat, H. (1961) *Proc. Nat. Acad. Sci. USA* **47**, 1393–1397.
15. Ihle, J. N., Lee, K.-L. & Kenney, F. T. (1974) *J. Biol. Chem.* **249**, 38–42.
16. Fraenkel-Conrat, H. & Singer, B. (1962) *Biochemistry* **1**, 120–128.
17. Young, R. J. & Content, J. (1971) *Nature New Biol.* **230**, 140–142.
18. Canaani, E., Helm, K.v.D. & Duesberg, P. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 401–405.
19. Duesberg, P. & Vogt, P. K. (1973) *J. Viol.* **12**, 594–599.
20. Beemon, K., Duesberg, P. & Vogt, P. (1974) *Proc. Nat. Acad. Sci. USA*, **71**, 4254–4258.